Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890] Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

Contents:

The following documents are made available to stakeholders:

Access the **final scope** and **final stakeholder list** on the NICE website.

Pre-technical engagement documents

1. **Company submission from CSL Vifor**

2. **Clarification questions and company responses**

3. **Patient group, professional group, and NHS organisation submissions** from:
   - Kidney Research UK

4. **External Assessment Report** prepared by Kleijnen Systematic Reviews Ltd

5. **External Assessment Report – factual accuracy check**

Post-technical engagement documents

6. **Technical engagement response from company**

7. **Technical engagement responses and statements from experts:**
   - Faizan Awan- patient expert nominated by Kidney Research UK

8. **Technical engagement responses from stakeholders:**
   a. Kidney Research UK
   b. UK Renal Pharmacy Group
   c. NHSE

9. **External Assessment Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews Ltd
   a. Critique
   b. Appendix
   c. EAG response to questions following the lead team meeting

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis

[ID3890]

Document B

Company evidence submission

File name	Version	Contains confidential information	Date
ID3890 Document B	1.0	Yes	10/08/2022

Abbreviations list

AE	Adverse event
ANCOVA	Analysis of covariance
BAD	British Association of Dermatologists
CHW	Cui, Hung, Wang
CI	Confidence interval
CKD	Chronic Kidney Disease
DB	Double-Blind
DOPPS	Dialysis Outcomes and Practice Patterns Study
DSMB	Data Safety Monitoring Board
ECG	Electroencephalograph
EOT	End of treatment
ESRD	End-stage renal disease
HD	Haemodialysis
HRQoL	Health-related quality of life
ICF	Informed consent form
IDMC	Independent data monitoring committee
IP	Investigational product
ITT	Intent-to-treat
IVRS	Interactive Voice Response Systems
IWRS	Interactive Web Response Systems
LS	Least squares
MAR	Missing-at-random
MedDRA	Medical Dictionary for Regulatory Activities
MHRA	Medicines and Healthcare products Regulatory Agency
MI	Multiple imputation
MMRM	Mixed model with repeated measures
MNAR	Missing-not-at-random
MOS	Medical Outcomes Study
ND	Non-dialysis-dependent

NICE	National Institute for Health and Care Excellence
NRS	Numerical Rating Scale
OLE	Open-label long-term extension
OOWS	Objective Opiate Withdrawal Scale
PbR	Payment-by-results
PGIC	Patient Global Impression of Change
PRO	Patient-reported outcome
PT	Preferred Terms
PTH	Parathyroid hormones
QALY	Quality-adjusted life year
QoL	Quality of life
SAE	Serious adverse event
SD	Standard deviation
ShOWS	Short Opiate Withdrawal Scale
SLR	Systematic literature review
SmPC	Summary of product characteristics
TEAE	Treatment-emergent adverse events
US	United States
UV	Ultraviolet

Contents

Tables and figures

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

The marketing authorisation for difelikefalin (Kapruvia ®) is for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients on haemodialysis. This submission covers the full marketing authorisation.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with moderate-to-severe pruritus receiving haemodialysis.	For the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis, including where established clinical management is insufficient in reducing pruritus.	An update was made as difelikefalin is restricted for in-centre haemodialysis use only.
Intervention	Difelikefalin	Difelikefalin	No change from scope.
Comparator(s)	Established clinical management without difelikefalin, including gabapentin and pregabalin	Established clinical management without difelikefalin, including gabapentin and pregabalin.	No change from scope.
Outcomes	The outcome measures to be considered include: • Itching intensity • Adverse effects of treatment • Health-related quality of life.	As per NICE final scope.	No change from scope.
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	As per NICE final scope.	No change from scope.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Subgroups to be considered	None specified.	• People with anti-pruritic medication use at baseline • People without anti-pruritic medication use at baseline • People with severe or very severe CKD-aP at baseline	Currently, there are no approved treatments for CKD-aP. The KALM trials did not directly include any comparator treatments, although Patients using anti-itch medication at baseline were allowed to continue doing so. It was deemed relevant to analyse subgroups based on use of anti-pruritic medication at baseline. The third subgroup was included to examine the impact of difelikefalin in the most severe CKD-aP category.
Special considerations including issues related to equity or equality	People in lower socio-economic groups are more likely to develop chronic kidney disease (CKD), progress towards kidney failure, and die earlier with CKD. People from black, Asian, and minority ethnic populations are more likely to progress to kidney failure faster and less likely to receive a transplant. Women are more likely to be diagnosed with CKD, but less likely to start dialysis. Older people with CKD are	As per NICE final scope	No change from scope

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	less likely to receive a kidney transplant than their younger counterparts. These populations are at greater risk of developing CKD-associated pruritus (CKD-aP) and experiencing symptoms for longer while on dialysis. Therefore, guidance on the use of difelikefalin could have a different impact on people with protected characteristics than on the wider population (1). Difelikefalin is restricted for in-centre haemodialysis use only. This may be considered to represent a barrier to some patients for whom in-centre haemodialysis is not accessible.		

B.1.2 Description of the technology being evaluated

Please see Appendix C for Summary of Product Characteristics (SmPC) and UK Assessment Report

Table 2 presents an overview of the drug being evaluated (difelikefalin).

Table 2: Technology being evaluated

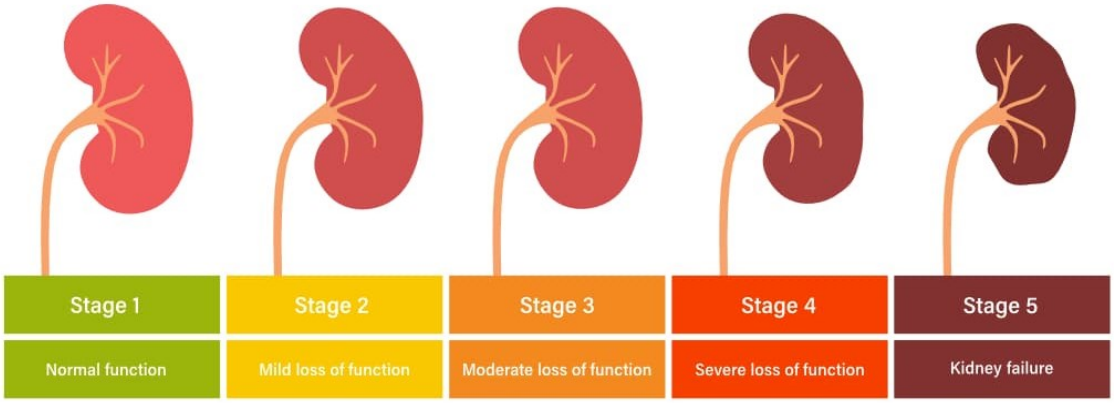
UK approved name and brand name	Difelikefalin (Kapruvia).
Mechanism of action	Difelikefalin is a selective kappa opioid receptor agonist with low central nervous system penetration. The pathophysiology of chronic kidney disease-associated pruritus is thought to be multifactorial, including systemic inflammation and an imbalance in the endogenous opioid system (e.g., overexpression of mu opioid receptors and concomitant downregulation of kappa opioid receptors). Opioid receptors are known to modulate itch signals and inflammation, with kappa opioid receptor activation reducing itch and producing immunomodulatory effects. The activation of kappa opioid receptors on peripheral sensory neurons and immune cells by difelikefalin are considered mechanistically responsible for its antipruritic and anti-inflammatory effects.
Marketing authorisation/CE mark status	UK market authorisation was granted by the Medicines and Healthcare products Regulatory Agency (MHRA) on 29th April 2022.
Indications and any restriction(s) as described in the SmPC	Difelikefalin is Kapruvia is indicated for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients on haemodialysis. Difelikefalin should be restricted to in-centre haemodialysis use only.
Method of administration and dosage	Difelikefalin is administered three times per week by intravenous bolus injection into the venous line of the dialysis circuit at the end of haemodialysis treatment, during rinse back or after rinse back. The recommended dose of difelikefalin is 0.5 micrograms/kg dry body weight (i.e., the target post-dialysis weight). The total dose volume (mL) required from the vial should be calculated as

	follows: 0.01 × dry body weight (kg), rounded to the nearest tenth (0.1 mL). For patients with a dry body weight equal to or above 195 kg the recommended dose is 100 micrograms (2 mL).
Additional tests or investigations	No additional tests or investigations required.
List price and average cost of a course of treatment	£35.00 per 1mL vial (50µg/mL) of difelikefalin. £420.00 for a 12 x 1mL vial pack
Patient access scheme (if applicable)	A simple discount patient access scheme is under fast-track review by NHS England.

B.1.3. Health condition and position of the technology in the treatment pathway

CKD is a common and progressive disease with a major global health burden associated with high morbidity and mortality. It is characterised by abnormalities of kidney function or structure that have been present for more than three months (2, 3). CKD can be categorised into five stages dependent on functionality of the kidney (as shown in **Error! Reference source not found.**), with stages 1 to 5 affecting an estimated 13.4% of the population worldwide, and stages 3 to 5 affecting 10.6%. Stage 3 CKD was shown to be most prevalent (7.6% of the population) with prevalence increasing with age ($p<0.001$) (3). The World Health Organization ranked kidney disease as the tenth most common cause of death in 2019, accounting for 2.4% of global mortality (4).

Figure 1: Levels of kidney function



Source: (5)

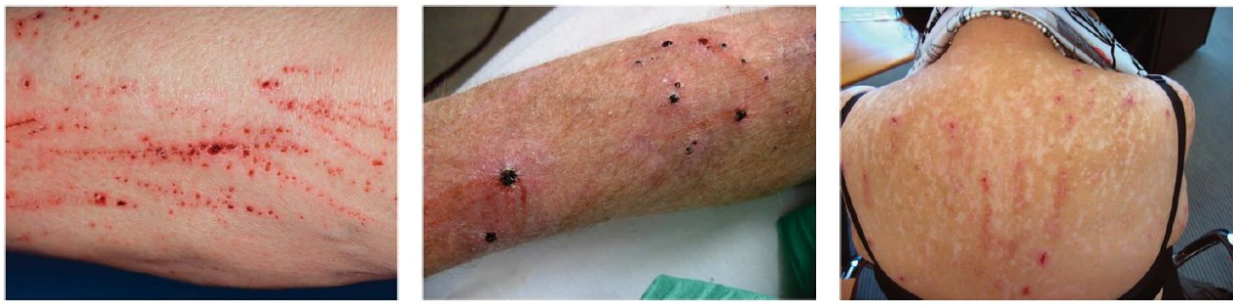
CKD-associated pruritus (CKD-aP), previously commonly referred to as uremic pruritus, is a serious, systemic itch comorbidity that occurs in CKD patients, particularly those undergoing dialysis, and is common among kidney failure patients (6). It is associated with poor quality of life (QoL), sleep disturbance, anxiety, and depression, as well as increased risks of infection, hospitalisation, and mortality (6, 7).

CKD-aP tends to present with symmetrical distribution and can be either generalised, affecting the entire skin, or localised, affecting only specific areas of the

body such as the scalp, face, upper back, arms (particularly the dialysis access arm), or buttocks (7). A 2015 review found that around 50% of CKD-aP patients report generalised pruritus, while the remainder mostly report itch localised to the back, face, and shunt arm (8). The severity of CKD-aP can change over time in individual patients, ranging from sporadic discomfort to complete restlessness during both day and night that significantly reduces QoL (8). Around 25% of patients report that severity is at its highest during or immediately after dialysis (8).

Error! Reference source not found. shows some of the typical skin manifestations of CKD-aP, including excoriations, prurigo nodularis, and scarring as a result of scratching (8). It is important to note, however, that some patients have no skin manifestations of CKD-aP, and that the severity of the itch is not correlated to observable skin damage.

Figure 2: Skin manifestations of CKD-aP



Source: (8)

CKD-aP is a very common condition in patients with CKD. When analysed using data from CKD Outcomes and Practice Patterns Study (CKDopps) (n=5,658 from France, United States, and Brazil), the overall prevalence of moderate-to-extreme CKD-aP was 24% (23% in France, 24% in Brazil, and 29% in the United States) (9). CKD-aP is found to occur in both non-dialysis-dependent (ND) and dialysis-dependent patients. However, analysis from the same study found it to be more prevalent in CKD patients undergoing dialysis (9, 10).

The most comprehensive data on the prevalence of CKD-aP in patients undergoing dialysis is from the international observational Dialysis Outcomes and Practice Patterns Study (DOPPS). DOPPS collected data on the prevalence of CKD-aP in 35,452 CKD patients on haemodialysis (HD) over six phases: 1996 to 2001, 2002 to 2004, 2005 to 2008, 2009 to 2011, 2012 to 2015, and 2015 to 2018. Analysis of the

data from phases 1 through 5 demonstrated that although there has been a decline in those who are ‘very much or extremely’ bothered by their itch, from 28% in 1996 to 18% in 2015, pruritus continues to be common among patients on HD. The analysis also revealed that although mild pruritus may inflict annoyance, severe pruritus has a major negative impact on patient lives (10). Additionally, it was found that in all 21 DOPPS countries analysed (including the UK), nephrologists tended to underestimate the prevalence of itching in their patients, and that many patients’ itch went unreported (10). Even amongst patients who were nearly always or always bothered by their itch, 17% had not reported any symptoms to any healthcare provider (10).

In the DOPPS Phase 5 (2012–2015), the proportion of patients at least moderately bothered by itchy skin ranged from 26% in Germany to 48% in the United Kingdom; 13% (Germany) to 26% (the United Kingdom) were very much or extremely bothered (10). A more recent analysis of DOPPS phases 4 to 6 demonstrates that overall prevalence is unchanged, and that the UK still has the highest proportion of patients in all 21 DOPPS countries who are moderately to extremely bothered by their itch (11). Overall, CKD-aP remains an under-reported and burdensome condition, with the most recent DOPPS data suggesting that approximately 47% of UK CKD patients on HD have moderate-to-severe CKD-aP, (Phase 4-6: 2009-2018) (11, 12).

Compared to those with normal renal function, patients with chronic kidney failure have a lower QoL, including reduced physical, psychological, and social functioning (13, 14). CKD-aP further decreases the QoL of these patients. Ramakrishnan et al. (2014) used the Kidney Disease Quality of Life scoring system tool to assess QoL in a population of >70,000 CKD-aP patients undergoing dialysis, finding a statistically significant association between increased itch severity and lower physical component summary and mental component summary scores (both $p<0.0001$) (6). A more recent study analysed 2,978 dialysis patients who completed patient reported outcome measures between 2018 and 2020, a sample taken from the Dutch RENINE/PROMs registry (15). It found that itching was associated with a lower physical and mental health-related quality of life (HRQoL) compared with patients without itching, as measured with the 12-item Short Form Health Survey (SF-12) ($p<.001$) (15). Moderate-to-severe itching also showed a larger decrease in physical and mental HRQoL compared to no or mild itching (15). These patients were

monitored for a 2-year period; no change in physical or mental HRQoL was seen in the overall population throughout the follow-up (15).

CKD-aP patients often report restless and poor quality sleep as a result of their itch, with CKD-aP patients reporting a loss of up to 2.2 hours of sleep per night (7). Poor sleep quality is closely associated with the severity of a patient's itch. In an analysis of DOPPS Phase 5 (2012–2015, n = 6025), of patients who were extremely bothered by itchy skin (n = 425), 66% frequently experienced restless sleep (10). It has been found that patients who were extremely bothered by their itch were four times as likely to find themselves awake at night compared with those who were not bothered by itch (16).

A patient-reported outcomes study conducted between 2009-2018 by Sukul et al. (2021), analysing 23,264 haemodialysis patients from 21 countries in the DOPPS phases 4 to 6. It was found that 37% of these patients were at least moderately bothered by itch, with pruritus being associated with poor sleep quality (32.1% of eligible patients reported ≥3 nights/week of restless sleep), depression (44% of eligible patients reported a 'Center for Epidemiologic Studies-Depression' score of ≥ 10) and mortality. There was also a strong association between itch severity and withdrawal from appointments or missed dialysis sessions and decreased employment rates.

Depression and anxiety are also estimated to affect up to 25% of CKD patients, considerably higher than the lifetime prevalence of 7% in the general population (17). The impact of CKD-aP further increases the risk of depression in the CKD population: in a DOPPS analysis, patients who were moderately to extremely bothered by CKD-aP were significantly more likely to have physician-diagnosed depression than those with no or mild CKD-aP (16).

CKD-aP patients on HD often have visible signs of scratching, as shown in **Error! Reference source not found.**. Significant associations have been found between itching intensity and feelings of embarrassment and frustration, as well as negative feelings about skin appearance in CKD-aP patients on HD (all $p<0.001$) (7). Further, CKD-aP patients on HD have a higher mortality rate than those with CKD alone - a 24-months prospective observational study showed that dialysis patients with CKD-aP were six times more likely to die than CKD patients without pruritus (18). A more

recent analysis of mortality in 23,264 haemodialysis patients in DOPPS phases 4-6 showed that compared with patients who reported being not at all bothered by itchy skin, patients who were extremely bothered had a higher rate of all-cause mortality (11). It was found by Sukul et al. (2020) that patients extremely bothered by itching also had higher rates of cardiovascular-related and infection-related mortality. Additionally, adjusted rates of all-cause, cardiovascular-related, and infection-related hospitalisations were all 20% greater for those extremely bothered versus not bothered at all by itchy skin (11). A previous study found that pruritus in haemodialysis patients was associated with a 17% increase in mortality risk ($P < 0.0001$) (16). It was also noted that this increase was no longer significant after adjusting for sleep quality measures, leading to the conclusion that the pruritus/mortality association may be attributed to poor sleep quality (16). Overall, a relationship exists between pruritus and mortality rate, as well as severity of pruritus and mortality rate. Although the exact cause of these relationships is not clear, it highlights one of the poor outcomes associated with CKD-aP that underlie the need for advances in treatment.

Although the pathophysiology is not well understood, there is increasing evidence that the cause of CKD-aP is multifactorial and involves immune system dysfunction (including elevated proinflammatory activity) and an imbalance in the endogenous opioid system (with overexpression of mu opioid receptors in dermal cells and lymphocytes and concomitant downregulation of kappa opioid receptors [KORs]) (19, 20).Despite the clear burden of CKD-aP, there is a lack of effective treatments for the comorbidity and there no approved drugs in Europe apart from difelikefalin. Current CKD-aP treatments (e.g., antihistamines, gabapentin, pregabalin) are used off-label, and these interventions, as well as UV phototherapy, are only supported by limited and low-grade clinical evidence (e.g., small sample sizes, high risk of bias, study heterogeneity). Consequently, there is a lack of robust treatment recommendations with no established standard of care (Simonsen et al., 2017). This results in a high level of unmet need among HD patients with moderate-to-severe CKD-aP.

From March to April 2022, a modified Delphi panel was conducted to collect expert opinion from eight consultant nephrologists from across England who treat patients with CKD-aP (Appendix N: Clinical opinion and consensus report).). [redacted]

[redacted].

In April-July 2022 a systematic literature review (SLR) was conducted to identify cost-effectiveness studies and treatment pathway guidelines for CKD-aP with a focus on the UK and Europe, see Appendix D.1.1 for more information. Figure 3 describes the treatment pathway for treating CKD-aP in the UK based off these SLR results, looking specifically at those published by the British Association of Dermatologists (BAD), while considering the results from the Delphi panel. The BAD provides guidelines on the investigation and management of uremic pruritus in adults without underlying dermatosis (21). As with other guidelines, most information relates to non-CKD-aP conditions. Where CKD-aP is mentioned, the guidelines recommend ensuring adequate dialysis, normalising the calcium-phosphate balance, controlling parathyroid hormones (PTH) to acceptable levels, correcting any anaemia and using simple emollients before employing other treatment strategies. If a patient is still suffering from pruritus the next stage is to use best supportive care, including creams and emollients, antihistamines, gabapentin and in some cases ultraviolet therapy or antidepressants. If a patient has failed on best supportive care this is when difelikefalin will be offered for the duration of dialysis, as long as a sufficient reduction in itch score has been achieved within the first 12 weeks of treatment.

Figure 3: Treatment pathway for pruritus patients, including the proposed positioning of difelikefalin

B.1.4 Equality considerations

Vifor aims to comply fully with all legal obligations to: promote race and disability equality and equality of opportunity between men and women, eliminate unlawful discrimination on grounds of race, disability, age, sex, gender reassignment, marriage and civil partnership, pregnancy and maternity (including women post-delivery), sexual orientation, and religion or belief (these are protected characteristics under the Equality Act 2010).

People in lower socio-economic groups are more likely to develop chronic kidney disease, progress towards kidney failure, and die earlier with CKD. People from

black, Asian and minority ethnic populations are more likely to progress to kidney failure faster and less likely to receive a transplant. Women are more likely to be diagnosed with CKD, but less likely to start dialysis. Older people with CKD are less likely to receive a kidney transplant than their younger counterparts. These populations are at greater risk of developing CKD-aP and experiencing symptoms for longer while on dialysis. Therefore, guidance on the use of difelikefalin could have a different impact on people with protected characteristics compared to the wider population (1). Difelikefalin is also restricted for in-centre haemodialysis use only, which may be considered to represent a barrier to some patients for whom in-centre haemodialysis is less accessible.

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

See Appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being evaluated.

B.2.2 List of relevant clinical effectiveness evidence

The efficacy and safety of difelikefalin as a treatment for moderate-to-severe CKD-aP in patients on HD has been demonstrated in a series of clinical studies as described in the following tables:

Table 3 KALM 1 Study Overview

Study	CLIN3102 (KALM-1)
Study design	Phase 3 randomised, 12-week, double-blind, placebo-controlled study
Population	Adults (≥18 years of age) with end-stage renal disease who had been on HD at least three times per week for at least three months and who had moderate-to-severe CKD-aP, defined as a weekly mean score of >4 points on the 24-hour WI-Numerical Rating Scale (NRS) (Worst Itching Intensity Numerical Rating Scale). A total of 378 patients were enrolled between February 2018 and December 2018. There was one patient exclusion due to not meeting the entry requirements, therefore a total of 377 patients progressed to the double-blind treatment period.
Intervention	Drug: Difelikefalin 0.5 mcg/kg
Comparator	Drug: Placebo
Indicate if study supports application for marketing authorisation	Yes (22, 23)
Indicate if study used in the economic model	Yes

Reported outcomes specified in the decision problem	Primary efficacy outcome: • Proportion of patients achieving ≥3-point reduction from baseline in weekly mean WI-NRS score (Week 12) Secondary efficacy outcome: • Change from baseline in HRQoL measured using the Skindex-10 scale total score (Week 12) • Change from baseline in HRQoL measured using the 5-D Itch scale total score (Week 12) • Proportion of patients achieving ≥4-point reduction from baseline in weekly mean WI-NRS score (Week 12) Safety: • Severity and seriousness of AEs and their relationship to study drug
All other reported outcomes	All reported outcomes are listed in Appendix M

Table 4: KALM-1 Open-Label Long-Term Extension study overview

Study	CLIN3102 (KALM-1) Open-Label Long-Term Extension
Study design	Phase 3, open-label, multicentre, long-term extension safety study
Population	To be eligible for the open-label extension phase of the study, a subject had to have received at least 30 doses of study drug (either placebo or active) during the 12-week double-blind treatment period and had to continue to meet other eligibility criteria listed below.
Intervention	Drug: Difelikefalin 0.5 mcg/kg
Comparator	None
Indicate if study supports application for marketing authorisation	Yes (22, 23)
Indicate if study used in the economic model	Yes
Reported outcomes specified in the decision problem	Severity and seriousness of adverse events (AEs) and their relationship to study drug. Change in total score and change by domain score from baseline in 5-D Itch scale.

All other reported outcomes	Clinical laboratory tests. Vital signs. 12-lead electrocardiogram (ECG) (first dialysis of Week 53 or early termination/End of Treatment). Inflammatory biomarkers. Use of concomitant and antipruritic medications. Number and reason(s) for missed dialysis. Use of concomitant ESAs. Use of concomitant iron medicine. Please see Appendix M for a full list of outcomes.

Table 5: CLIN3103 (KALM-2) study overview

Study	CLIN3103 (KALM-2)
Study design	Phase 3 randomised, 12-week, double-blind, placebo-controlled study
Population	Eligible patients were adults (18 – 85 years of age) with end-stage renal disease (ESRD) who had been on HD at least three times per week for at least three months, and who had moderate-to-severe CKD-aP (defined as a weekly mean score >4 on the 24-hour WI-NRS). A total of 473 patients were enrolled between July 2018 and February 2020. There were two patient exclusions, therefore a total of 471 patients progressed to the double-blind treatment period.
Intervention	Drug: Difelikefalin 0.5 mcg/kg
Comparator	Drug: Placebo
Indicate if study supports application for marketing authorisation	Yes (22, 23)
Indicate if study used in the economic model	Yes
Reported outcomes specified in the decision problem	Primary efficacy outcome: • Proportion of patients achieving at least a 3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at Week 12 Secondary efficacy outcome: • Proportion of patients achieving ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 4 and Week 8 • Proportion of patients achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 4 and Week 8 and Week 12 • Change from baseline in itch-related QoL at the end of Week 12, as assessed by the Skindex-10 scale total score • Change from baseline in itch-related QoL at the end of Week 12, as assessed by the 5-D Itch scale total score • Severity and seriousness of AEs and their relationship to study drug
All other reported outcomes	All reported outcomes are listed in Appendix M

Table 6 KALM-2 Open-Label Long-Term Extension study overview

Study	CLIN3103 (KALM-2) Open-Label Long-Term Extension
Study design	Phase 3, open-label, multicentre, long-term extension safety study
Population	Patients who received at least 30 doses of study drug (either active or placebo) during the 12-Week treatment period and continued to meet other eligibility criteria were eligible to receive open-label difelikefalin for an additional 52 weeks (52-week open-label extension (OLE) phase).
Intervention	Drug: Difelikefalin 0.5 mcg/kg
Comparator	None
Indicate if study supports application for marketing authorisation	Yes (22, 23)
Indicate if study used in the economic model	Yes
Reported outcomes specified in the decision problem	• Severity and seriousness of AEs and their relationship to study drug • Clinical laboratory tests • Vital signs • 12-lead ECG (first dialysis of Week 53 or early termination/end of treatment) • Inflammatory biomarkers • Change in total score and change by domain score from baseline in 5-D Itch scale
All other reported outcomes	• Use of concomitant and antipruritic medications • Number and reason(s) for missed dialysis • Use of concomitant ESAs • Use of concomitant iron medicine Please see Appendix M for a full list of outcomes.

CLIN3105 was a global multicentre open-label study conducted in the US and Eastern Europe to evaluate the safety and effectiveness of difelikefalin at a dose of 0.5 mcg/kg IV administered after each haemodialysis session to subjects with moderate-to-severe CKD-aP. Eligible patients were aged 18 years to 85 years, had ESRD, and had been on haemodialysis 3 times per week for at least 3 months prior to the star of screening (as well as other inclusion and exclusion criteria, which are

listed in Appendix O). In total, 222 patients received study treatment in CLIN3105. The primary objective of the study was to evaluate the safety of difelikefalin. The secondary objectives were to evaluate the effectiveness of IV difelikefalin at a dose of 0.5 mcg/kg in reducing the intensity of itch and improving the itch-related quality of life and quality of sleep measures in haemodialysis patients with moderate-to-severe pruritus. Please see Appendix M for full list of reported outcomes for CLIN3105.

CLIN3105 is being submitted as additional supporting evidence. CLIN3105 was not used to populate the economic model but is included in sections 2.2 to 2.6. The results of this study, in which all patients knowingly receive active treatment, provide insight into the expected real-world effectiveness of difelikefalin. This study was not included in the economic model because it did not contain a relevant comparator arm.

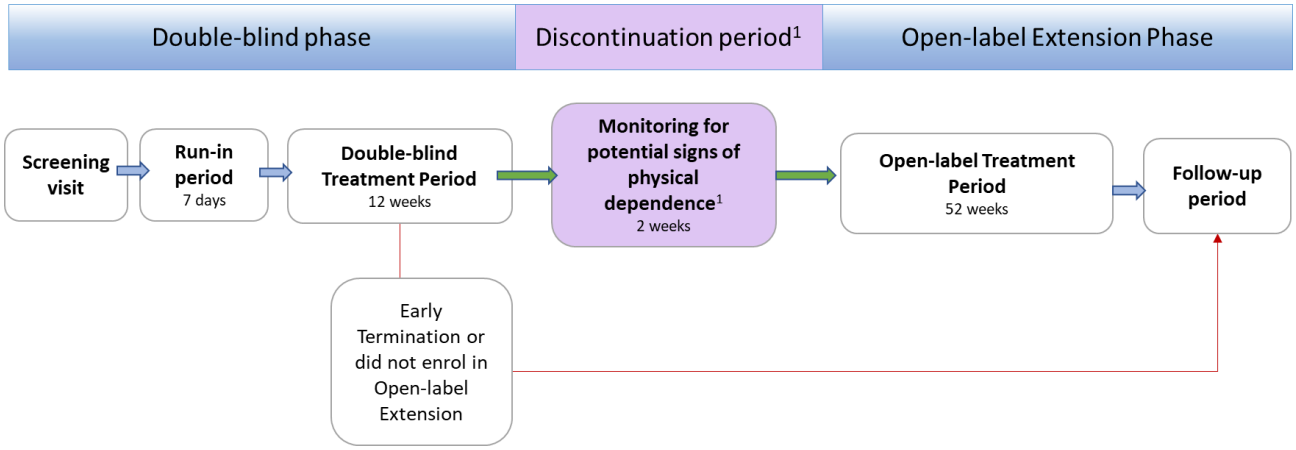
B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

The trial design and methodology of KALM-1, KALM-1 OLE, KALM-2, KALM-2 OLE, and CLIN3105 are described below:

KALM-1 and KALM-2 Trial designs (24, 25)

KALM-1 and KALM-2 were multicentre, randomised, double-blind, placebo-controlled studies to evaluate the safety and efficacy of difelikefalin at a dose of 0.5 mcg/kg administered after each haemodialysis session (3 times a week) in subjects with moderate-to-severe pruritus. Both studies included a double-blind phase and OLE phase. The double-blind phase consisted of a screening visit, a 7-day run-in period during the week prior to randomisation and a 12-week double-blind treatment period where difelikefalin was evaluated relative to placebo. For KALM-1 the double-blind treatment period was followed by a 2-week discontinuation period, during which no study drug was administered and subjects were monitored for potential signs or symptoms of physical dependence, before advancing to the OLE phase (**Error! Reference source not found.**).

Figure 4 Trial design of KALM-1, KALM-2 and the OLE studies



[1]The discontinuation period is only applicable to KALM-1 and not KALM-2

The purpose of the 7-day run-in period was to confirm that each subject did have moderate-to-severe pruritus, and to establish baseline itch intensity. The criteria for eligibility were not communicated to the subjects. The run-in period was also used to record each subject's use of anti-itch medications: subjects were stratified according to their use or non-use of concomitant medications to treat pruritus during the week prior to randomisation, as well as the presence or absence of specific medical conditions.

If subjects continued to meet all inclusion criteria and no exclusion criteria at the end of the 7-day run-in period, they were randomised in a 1:1 ratio to receive either difelikefalin 0.5 mcg/kg or placebo as an IV bolus after the end of each haemodialysis session during the 12-week double-blind treatment period. This meant that each subject received difelikefalin or placebo three times weekly, for a total of up to 36 doses.

Eligibility criteria

Please see Table 12 for more information.

Settings and locations

Please see Table 12 for more information.

Trial drugs and concomitant medications

Please see Table 12 for more information

Outcomes used in the economic model or specified in the scope

Please see Table 12 for more information

Subject baseline characteristics

Table 7 summarises the baseline characteristics of the KALM-1 double-blind safety population.

Table 7 Baseline characteristics of KALM-1 (double-blind safety population)

Baseline characteristic	Difelikefalin	Placebo
Number of participants	(n= 189)	(n= 188)
Mean age, years (SD)	58.2 (11.16)	56.8 (13.89)
Sex – n (%)		
Male	112 (59.3%)	118 (62.8%)
Female	77 (40.7%)	70 (37.2%)
Ethnicity – n (%)		
Hispanic or Latino	64 (33.9%)	68 (36.2%)
Not Hispanic or Latino	123 (63.8%)	120 (63.8%)
Unknown	2 (1.1%)	0
Race		
American Indian or Alaska Native	6 (3.2%)	5 (2.7%)
Asian	6 (3.2%)	7 (3.7%)
Black or African American	82 (43.4%)	75 (39.9%)
Native Hawaiian or Other Pacific Islander	4 (2.1%)	4 (2.1%)
White	91 (48.1%)	94 (49.5%)
Unknown	1 (0.5%)	2 (1.1%)
Other	1 (0.5%)	2 (1.1%)
Mean prescription dry body weight, kg (SD)	85.91 (20.264)	84.98 (21.084)
Baseline worst itching NRS, mean (SD)	7.06 (1.439)	7.25 (1.606)
Baseline anti-itch medication use? [1] – n (%)		
Yes	72 (38.1%)	78 (41.5%)
No	117 (61.9%)	110 (58.5%)
Specific medical condition? [1] – n (%)		
Yes	25 (13.2%)	28 (14.9%)
No	164 (86.8%)	160 (85.1%)
Mean duration of pruritus, years (SD)	3.19 (3.244)	3.45 (3.369)

Baseline characteristic	Difelikefalin	Placebo
Mean years since diagnosis of ESRD, years (SD)	4.66 (3.898)	5.66 (5.178)
Years since diagnosis of CKD		
n	187	189
Mean (SD)	6.92 (5.926)	7.03 (5.739)
Years on chronic haemodialysis, mean (SD)	4.37 (3.982)	4.73 (4.219)
Aetiology of CKD [2]		
Diabetes	107 (56.6%)	94 (50.0%)
Hypertension	129 (68.3%)	139 (73.9%)
Large vessel disease	4 (2.1%)	4 (2.1%)
Glomerulonephritis	7 (3.7%)	8 (4.3%)
Vasculitis	0	0
Interstitial nephritis	1 (0.5%)	0
Pyelonephritis	0	0
Cystic	1 (0.5%)	2 (1.1%)
Hereditary	1 (0.5%)	2 (1.1%)
Congenital	0	0
Neoplasms	1 (0.5%)	1 (0.5%)
Tumours	2 (1.1%)	0
Urologic	0	0
Nephrotic syndrome	2 (1.1%)	4 (2.1%)
Unknown	7 (3.7%)	6 (3.2%)
Other	11 (5.8%)	16 (8.5%)

CKD = chronic kidney disease; ESRD = end-stage renal disease; max = maximum; min = minimum; NRS = numerical rating scale; SD = standard deviation.

[1] Observed stratum values.

[2] More than one item may have been checked.

Source: (24)

Table 8 summarises the baseline characteristics of the KALM-2 double-blind safety population.

Table 8: Baseline characteristics KALM-2 (double-blind safety population)

Baseline characteristic	Difelikefalin	Placebo
Number of participants	236	235
Mean age, years (SD)	59.7 (13.11)	59.6 (13.07)
Sex – n (%)		
Male	135 (57.4%)	139 (58.9%)

Baseline characteristic	Difelikefalin	Placebo
Female	100 (42.6%)	97 (41.1%)
Ethnicity – n (%)		
Hispanic or Latino	68 (28.9%)	68 (28.8%)
Not Hispanic or Latino	163 (69.4%)	166 (70.3%)
Not reported	2 (0.9%)	2 (0.8%)
Unknown	2 (0.9%)	0
Race		
American Indian or Alaska Native	1 (0.4%)	1 (0.4%)
Asian	12 (5.1%)	20 (8.5%)
Black or African American	53 (22.6)	38 (16.1%)
Native Hawaiian or Other Pacific Islander	1 (0.4%)	3 (1.3%)
White	162 (68.9%)	169 (71.6%)
Other	6 (2.6%)	5 (2.1%)
Mean prescription dry body weight, kg (SD)	81.56 (19.731)	79.95 (19.450)
Baseline Worst Itching NRS, Mean (SD)	7.27 (1.358)	7.12 (1.363)
Baseline anti-itch medication use? [1] – n (%)		
Yes	87 (37.0%)	85 (36.0%)
No	148 (63.0%)	151 (64.0%)
Specific medical condition? [1] – n (%)		
Yes	41 (17.4%)	37 (14.7%)
No	194 (82.6%)	199 (84.3%)
Mean duration of pruritus, years (SD)	3.21 (4.567)	3.20 (3.184)
Mean years since diagnosis of ESRD, years (SD)	5.23 (4.677)	5.46 (4.509)
Years since diagnosis of CKD		
n	234	232
Mean (SD)	9.28 (7.638)	9.76 (7.009)
Years on chronic haemodialysis, mean (SD)	4.83 (4.588)	5.09 (4.327)
Aetiology of CKD [2]		
Diabetes	118 (50.2%)	112 (47.5%)
Hypertension	121 (51.5%)	114 (48.3%)
Large vessel disease	4 (1.7%)	3 (1.3%)
Glomerulonephritis	14 (6.0%)	17 (7.2%)
Vasculitis	3 (1.3%)	2 (0.8%)
Interstitial nephritis	2 (0.9%)	1 (0.4%)

Baseline characteristic	Difelikefalin	Placebo
Pyelonephritis	3 (1.3%)	1 (0.4)
Cystic	18 (7.7%)	16 (6.8%)
Hereditary	13 (5.5%)	6 (2.5%)
Congenital	1 (0.4%)	3 (1.3%)
Neoplasms	0	2 (0.8%)
Tumours	1 (0.4%)	1 (0.4%)
Urologic	6 (2.6%)	9 (3.8%)
Nephrotic syndrome	3 (1.3%)	6 (2.5%)
Unknown	8 (3.4%)	14 (5.9%)
Other	26 (11.1%)	28 (11.0%)

CKD = chronic kidney disease; ESRD = end-stage renal disease; max = maximum; min = minimum; NRS = Numerical Rating Scale; SD = standard deviation.

[1] Observed stratum values.

[2] More than one item may have been checked.

Source: (25)

KALM-1 and KALM-2 OLE trial designs (26, 27)

Both KALM-1 and KALM-2 included an OLE phase. The study protocols for KALM-1 and KALM-2 consisted of both a double-blind phase and an OLE phase; no separate objectives were specified for the OLE phase. The open-label part of the study was designed to evaluate the safety of difelikefalin at a dose of 0.5 mcg/kg administered intravenously after each dialysis session (generally 3 times per week) during long-term use (for up to 52 weeks) in subjects who had completed the 12-week double-blind treatment period. It also evaluated the maintenance of treatment effect during long-term use.

The OLE phase consisted of the open-label treatment period and the follow-up period. The first visit and first dosing for the OLE phase occurred during the week following the discontinuation period in KALM-1. For KALM-2 the dose is given either on the day of the last visit of the double-blind treatment period or on the next visit up to 1 week following the double-blind treatment period. The last dose of the study drug was administered at the last haemodialysis treatment of Week 52. A final safety follow-up visit was conducted 7 to 10 days after the end of treatment/early termination visit.

Eligibility criteria:

Please see Table 12 for more information.

Settings and locations:

Please see Table 12 for more information.

Trial drugs and concomitant medications for OLEs

Please see Table 12 for more information.

Outcomes used in the economic model or specified in the scope:

Please see Table 12 for more information.

Subject baseline characteristics

Baseline characteristics for the KALM-1 open-label safety population are presented in Table 9. These values were collected during the screening visit for the double-blind treatment phase.

Table 9: Baseline characteristics KALM-1 OLE (open-label safety population)

Baseline characteristic	Pbo/DFK	DFK/DFK
Number of participants	162	151
Mean age, years (SD)	57.1 (13.79)	58.0 (11.45)
Sex – n (%)		
Male	102 (63.0%)	87 (57.6%)
Female	60 (37.0%)	64 (42.4%)
Ethnicity – n (%)		
Hispanic or Latino	56 (34.6%)	52 (34.4%)
Not Hispanic or Latino	106 (65.4%)	98 (64.9%)
Unknown	0	1 (0.7%)
Race		
American Indian or Alaska Native	4 (2.5%)	3 (2.0%)
Asian	7 (4.3%)	5 (3.3%)
Black or African American	68 (42.0%)	67 (44.4%)
Native Hawaiian or Other Pacific Islander	3 (1.9%)	1 (0.7%)
White	76 (46.9%)	74 (49.0%)
Unknown	2 (1.2%)	0
Other	2 (1.2%)	1 (0.7%)
Mean prescription dry body weight, kg (SD)	84.53 (20.885)	85.87 (20.905)
Baseline Worst Itching NRS, Mean (SD)	7.20 (1.586)	7.00 (1.440)

Baseline characteristic	Pbo/DFK	DFK/DFK
Baseline anti-itch medication use? [1] – n (%)		
Yes	64 (39.5%)	54 (35.8%)
No	98 (60.5%)	97 (64.2%)
Specific medical condition? [1] – n (%)		
Yes	23 (14.2%)	22 (14.6%)
No	139 (85.8%)	129 (85.4%)
Mean duration of pruritus, years (SD)	3.53 (3.439)	3.29 (3.492)
Mean years since diagnosis of ESRD, years (SD)	5.77 (5.272)	4.67 (4.011)
Years since diagnosis of CKD		
n	161	151
Mean (SD)	7.0 (5.829)	6.97 (5.995)
Years on chronic haemodialysis, mean (SD)	4.85 (4.404)	4.44 (4.131)
Aetiology of CKD [2]		
Hypertension	120 (74.1%)	107 (70.9%)
Diabetes	82 (50.6%)	82 (54.3%)
Other	13 (8.0%)	8 (5.3%)
Glomerulonephritis	8 (4.9%)	4 (2.6%)
Unknown	5 (3.1%)	5 (3.3%)
Large Vessel Disease	3 (1.9%)	4 (2.6%)
Nephrotic Syndrome	2 (1.2%)	2 (1.3%)
Cystic	2 (1.2%)	1 (0.7%)
Hereditary	2 (1.2%)	1 (0.7%)
Neoplasms	1 (0.6%)	1 (0.7%)
Tumours	0	2 (1.3%)
Interstitial Nephritis	1 (0.6%)	0
Congenital	0	0
Pyelonephritis	0	0
Urologic	0	0
Vasculitis	0	0

CKD = chronic kidney disease; ESRD = end-stage renal disease; max = maximum; min = minimum; NRS = numerical rating scale; SD = standard deviation.

[1] Observed stratum values.

[2] More than one item may have been checked.

Note: Baseline characteristics were recorded during the screening visit for the double-blind treatment phase.

Baseline characteristics for the KALM-2 open-label safety population are presented in Table 10. These values were collected during the screening visit for the double-blind treatment phase.

Table 10: Baseline characteristics KALM-2 OLE (open-label safety population)

Baseline characteristic OLE	Pbo/DFK	DFK/DFK
Number of participants	210	189
Mean age, years (SD)	59.4 (13.13)	59.7 (12.88)
Sex – n (%)		
Male	124 (59.0%)	110 (58.2%)
Female	86 (41.0%)	79 (41.8%)
Ethnicity – n (%)		
Hispanic or Latino	56 (26.7%)	58 (30.7%)
Not Hispanic or Latino	152 (72.4%)	127 (67.2%)
Unknown	0	2 (0.5%)
Not reported	2 (1.0%)	4 (1.0%)
Race		
American Indian or Alaska Native	1 (0.5%)	1 (0.5%)
Asian	16 (7.6%)	9 (4.8%)
Black or African American	33 (15.7%)	39 (20.6%)
Native Hawaiian or Other Pacific Islander	3 (1.4%)	1 (0.5%)
White	153 (72.9%)	135 (71.4%)
Other	4 (1.9%)	4 (2.1%)
Mean prescription dry body weight, kg (SD)	79.67 (19.227)	81.75 (20.326)
Baseline Worst Itching NRS, Mean (SD)	7.07 (1.352)	7.24 (1.396)
Baseline anti-itch medication use? [1] – n (%)		
Yes	75 (35.7%)	65 (34.4%)
No	135 (64.3%)	124 (65.6%)
Specific medical condition? [1] – n (%)		
Yes	35 (16.7%)	30 (15.9%)

Baseline characteristic OLE	Pbo/DFK	DFK/DFK
No	175 (83.3%)	159 (84.1%)
Mean duration of pruritus, years (SD)	3.31 (3.258)	2.92 (2.837)
Mean years since diagnosis of ESRD, years (SD)	5.61 (4.668)	5.19 (4.848)
Years since diagnosis of CKD		
n	206	188
Mean (SD)	10.04 (7.254)	9.29 (7.949)
Years on chronic haemodialysis, mean (SD)	5.23 (4.488)	4.82 (4.797)
Aetiology of CKD [2]		
Hypertension	99 (47.1%)	100 (52.9%)
Diabetes	96 (45.7%)	93 (49.2%)
Other	26 (12.4%)	21 (11.1%)
Cystic	15 (7.1%)	14 (7.4%)
Glomerulonephritis	17 (8.1%)	12 (6.3%)
Unknown	13 (6.2%)	7 (3.7%)
Hereditary	5 (2.4%)	12 (6.3%)
Urologic	8 (3.8%)	5 (2.6%)
Nephrotic Syndrome	6 (2.9%)	3 (1.6%)
Large Vessel Disease	3 (1.4%)	4 (2.1%)
Vasculitis	2 (1.0%)	2 (1.1%)
Pyelonephritis	1 (0.5%)	2 (1.1%)
Congenital	2 (1.0%)	0
Interstitial Nephritis	1 (0.5%)	1 (0.5%)
Neoplasms	2 (1.0%)	0
Tumours	1 (0.5%)	1 (0.5%)

CKD = chronic kidney disease; ESRD = end-stage renal disease; max = maximum; min = minimum; NRS = numerical rating scale; SD = standard deviation.

[1] Observed stratum values.

[2] More than one item may have been checked.

Note: Baseline characteristics were recorded during the screening visit for the double-blind treatment phase.

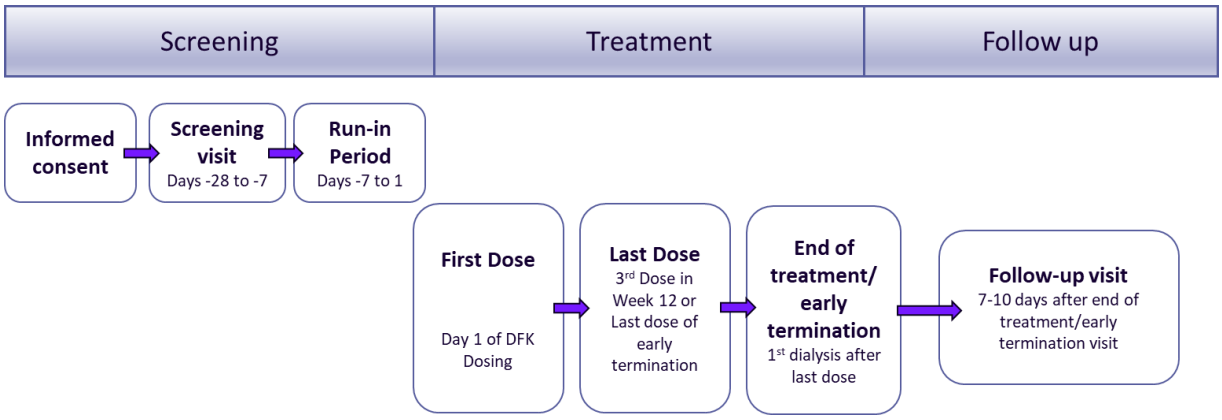
CLIN3105 (28)

Trial design

CLIN3105 was an open-label, multicentre, Phase III study conducted in the United States and Europe. It was designed to evaluate the safety and efficacy of IV difelikefalin at a dose of 0.5 mcg/kg moderate-to-severe CKD-aP patients

undergoing haemodialysis. Patients received difelikefalin as an IV bolus after the end of their dialysis during a treatment period of up to 12 weeks, so that each patient received difelikefalin three times per week for a total of up to 36 doses. End of treatment (EOT) was defined as the first day of dialysis following the last dose of the drug. The EOT procedures were conducted during the dialysis visit following the last dose of the study drug. A final safety follow-up visit was conducted 7 to 10 days after the EOT or early termination visit (Figure 5).

Figure 5: CLIN3105 study design



The screening period to assess eligibility occurred within 28 days prior to treatment, and consisted of a screening visit and a run-in period. The purpose of the run-in period was to confirm that each subject had moderate-to-severe pruritus. The screening period was also used to record each subject's use of antipruritic medications.

If subjects continued to meet all inclusion criteria and no exclusion criteria at the end of the run-in period, they could start the treatment period and begin treatment with IV difelikefalin 0.5 mcg/kg.

Eligibility criteria

Please see Table 12 for more information.

Settings and locations

Please see Table 12 for more information.

Trial drugs and concomitant medications:

Please see Table 12 for more information.

Outcomes used in the economic model or specified in the scope

No outcomes of CLIN3105 were used in the economic model.

Please see Table 12 for more information.

Subject baseline characteristics

Baseline characteristics of subjects in CLIN3105 are presented in Table 11.

Table 11: Baseline characteristics CLIN3105 (safety population)

Baseline characteristic	Difelikefalin
Number of participants	222
Mean age, years (SD)	58.1 (12.81)
Sex – n (%)	
Male	121 (54.5%)
Female	101 (45.5%)
Ethnicity – n (%)	
Hispanic or Latino	48 (21.6%)
Not Hispanic or Latino	173 (77.9%)
Not reported	1 (0.5%)
Race	
American Indian or Alaska Native	2 (0.9%)
Asian	7 (3.2%)
Black or African American	110 (49.5%)
Native Hawaiian or Other Pacific Islander	3 (1.4%)
White	96 (43.2%)
Other	4 (1.89%)
Mean target dry body weight at baseline, kg (SD)	86.64 (23.548)
Baseline worst itching NRS, mean (SD)	7.57 (1.331)
Baseline anti-itch medication use? – n (%)	
Yes	70 (31.5%)
No	152 (68.5%)
Mean duration of pruritus, years (SD)	3.89 (3.312)
Mean years since diagnosis of ESRD (SD)	5.87 (4.690)
Mean years since diagnosis of CKD (SD)	8.51 (6.878)
Mean years on chronic haemodialysis (SD)	5.42 (4.413)
Aetiology of CKD [1]	
Hypertension	135 (60.8%)

Baseline characteristic	Difelikefalin
Diabetes	110 (49.5%)
Other	25 (11.3%)
Glomerulonephritis	11 (5.0%)
Large vessel disease	4 (1.8%)
Urologic	3 (1.4%)
Pyelonephritis	2 (0.9%)
Cystic	2 (0.9%)
Unknown	2 (0.9%)
Interstitial nephritis	1 (0.5%)
Nephrotic syndrome	1 (0.5%)
Tumours	1 (0.5%)
Vasculitis	1 (0.5%)

max = maximum; min = minimum; SD = standard deviation; CKD = chronic kidney disease; ESRD = end-stage renal disease; WI-NRS = Worst Itching Intensity Numerical Rating Scale

Percentages were based on the number of subjects in the safety population and noted parenthetically.

Vital signs baseline was defined as the last measurement taken on or prior to the first day of dosing.

[1] - More than one item may have been checked.

Summary of methodologies (26, 27, 29-31)

The table below summarises the methodologies of all trials included in this submission:

Table 12: Summary of methodology

Trial Number	KALM-1	KALM-2	KALM-1 OLE	KALM-2 OLE	CLIN3105
Location	57 centres in the United States	93 centres in the United States, Australia, Canada, Czech Republic, Germany, Hungary, South Korea, New Zealand, Poland, Taiwan, and the United Kingdom	57 centres in the United States	93 centres in the United States, Australia, Canada, Czech Republic, Germany, Hungary, South Korea, New Zealand, Poland, Taiwan, and the United Kingdom	43 centres across the United States, Czech Republic, Hungary, and Poland
Trial design	Multicentre, randomised, double-blind, placebo-controlled studies		52-week OLE phase to KALM-1	52-week OLE phase to KALM-2	Global, multicentre, open-label study
Eligibility criteria	Adults (≥18 years of age) with ESRD who had been on HD at least three times per week for at least three months and who had moderate-to-severe CKD-aP. For full list of eligibility criteria please see Appendix O.	Adults (18-85 years of age) with ESRD who had been on HD at least three times per week for at least three months and who had moderate-to-severe CKD-aP. For full list of eligibility criteria please see Appendix O.	Adults (≥18 years of age) with ESRD who had been on HD at least three times per week for at least three months, who had moderate-to-severe CKD-aP, and who had received at least 30 doses of difelikefalin in the double-blind phase of KALM-1. For full list	Adults (18-85 years of age) with ESRD who had been on HD at least three times per week for at least three months, who had moderate-to-severe CKD-aP, and who had received at least 30 doses of difelikefalin in the double-blind phase of KALM-2. For full list	Adults (18-85 years of age) with ESRD who had been on HD at least three times per week for at least three months and who had moderate-to-severe CKD-aP. For full list of eligibility criteria please see Appendix O.

Trial Number	KALM-1	KALM-2	KALM-1 OLE	KALM-2 OLE	CLIN3105
			of eligibility criteria please see Appendix O.	of eligibility criteria please see Appendix O.	
Trial drugs	Subjects were randomised 1:1 to receive either IV difelikefalin (0.5 mcg/kg) (N= 189), or placebo (N=189).	Subjects were randomised 1:1 to receive either IV difelikefalin (0.5 mcg/kg) (N= 237), or placebo (N=236).	All patients received IV difelikefalin (0.5 mcg/kg).		
	The study drug was dispensed by qualified staff members who had received training on study drug handling and administration.				
	Subjects received difelikefalin at a dose of 0.5 mcg/kg or placebo after each haemodialysis session, generally 3 times per week for up to 12 weeks. Treatment was administered as an IV bolus into the venous line of the haemodialysis circuit either during or after rinse back at the end of each haemodialysis session.		Subjects received difelikefalin at a dose of 0.5 mcg/kg after each haemodialysis session, generally 3 times per week for up to 52 weeks. This was in addition to the treatments received during the double-blind phase (0.5 mcg/kg after each haemodialysis session, generally 3 times per week for up to 12 weeks). Treatment was administered as an IV bolus into the venous line of the haemodialysis circuit either during or after rinse back at the end of each haemodialysis session.		Subjects received difelikefalin three times per week for up to 12 weeks, for a total of up to 36 doses. Difelikefalin was administered as a 0.5 mcg/kg IV bolus into the venous line at the end of the subject's haemodialysis, either during rinse back or after rinse back.
	If a subject received additional haemodialysis during a given week for any reason, an additional dose of difelikefalin or placebo was administered following haemodialysis. A maximum of four doses per week was allowed. No additional doses were given to subjects receiving an additional unscheduled ultrafiltration treatment. If a subject missed a haemodialysis visit and the planned dose of difelikefalin or placebo for that visit, dosing is resumed at the next haemodialysis visit with no additional doses given.				

Trial Number	KALM-1	KALM-2	KALM-1 OLE	KALM-2 OLE	CLIN3105
Permitted and disallowed concomitant medication	Concomitant medication during the treatment period was restricted as follows: **Investigational drug (other than the study drug)** – Not allowed **Ultraviolet light-B treatments** – Not allowed **Naloxone, naltrexone, or mixed agonist-antagonists (e.g., buprenorphine and nalbuphine)** - Not allowed from the start of dosing of the double-blind treatment period to the end of the open-label treatment period (or from screening to the end of the treatment period for CLIN3105), unless needed for acute treatment of an adverse event or emergent medical condition. **Antihistamines (oral, IV, or topical), corticosteroids (oral, IV, or topical), opioids, gabapentin, or pregabalin** - Changes to current prescription were to be avoided from screening to the end of the treatment period, unless for the acute treatment of an adverse event or emergent medical condition (in this case, the study Medical Monitor was to be notified and, as appropriate, the adverse event(s) were to be reported). No new medication to treat itch was to be initiated.				
Primary outcomes	Proportion of patients achieving ≥3-point reduction from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at Week 12		The following assessments were used to evaluate the safety of difelikefalin in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus: • AEs • Vital signs • Electrocardiograms • Clinical laboratory values		
Other outcomes used in the economic model	WI-NRS total score at baseline, Week 4, Week 8 and Week 12. Proportion of patients achieving ≥3-point reduction from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at Week 4, Week 8, and Week 12.	WI-NRS total score at baseline, Week 4, Week 8 and Week 12. Proportion of patients achieving ≥3-point reduction from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at Week 4, Week 8, and Week 12	5-D Itch total score at baseline (Week 12 of double-blind phase) and at Week 52. Adverse events.	5-D Itch total score at baseline (Week 12 of double-blind phase) and at Week 52. Adverse events.	No outcomes used in the economic model

Trial Number	KALM-1	KALM-2	KALM-1 OLE	KALM-2 OLE	CLIN3105
	5-D Itch total score at baseline, Week 4, Week 8 and Week 12.	5-D Itch total score at baseline, Week 4, Week 8 and Week 12.			
	Please see Appendix M for full list of outcomes in all studies				
Pre-planned subgroups	Interim analysis subjects and post-interim analysis. By stratification factor: • Use of anti-itch medication at baseline • Presence of specific medical conditions at baseline (Please see Section Appendix E for more information)		No pre-planned subgroups		
	-	By region By dialysis type (haemodialysis or haemodiafiltration)			

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

Please see Appendix D for numbers of participants eligible to enter the studies.

Description of study populations

Sources for this section: (32-36)

Table 13 provides a description of all study populations across the trials included in this submission.

Table 13: Overview of study populations

Study	Study population	Description
KALM-1	Enrolled population	Subjects who signed informed consent.
	Intent-to-treat (ITT) population	Subjects who were randomised to a treatment group. Subjects in the ITT population were analysed according to their randomised treatment, regardless of the actual treatment received. The ITT population was used to analyse all efficacy endpoints collected during the double-blind phase.
	Double-blind safety population	Randomised subjects who received at least one dose of double-blind study drug during the double-blind treatment period subjects in the double-blind safety population were analysed according to the actual treatment received. This population was used to analyse all safety endpoints collected during the double-blind phase.
	Double-blind discontinuation safety population	Subset of subjects in the double-blind safety population who had at least one visit in the discontinuation period. The double-blind discontinuation safety population was used to analyse all safety endpoints collected during the discontinuation period.
	Double-blind discontinuation population	Subset of subjects in the double-blind safety population who completed 12 weeks of treatment, received at least six doses in the 2 weeks prior to the start of the discontinuation period, and had at least one visit in the discontinuation period.

Study	Study population	Description
		The double-blind discontinuation population was the primary population used to analyse the endpoints related to drug withdrawal. These analyses were also conducted for the double-blind discontinuation safety population.
	Per protocol population	Subset of subjects in the ITT population who did not have any major protocol deviations that could have affected the efficacy analyses of the double-blind data. The per protocol population was defined as subjects who: Received at least 80% of the planned study drug doses while in the study Received at least one study dose in each of Week 11 and 12 of the double-blind treatment period, if present through Week 12 Did not receive a different treatment than the one to which they were randomised Had a mean baseline WI-NRS score >4.0 Had a non-missing average 24-hour weekly WI-NRS score available for at least 75% of study weeks while in the study (weeks with >3 missing daily values were considered missing) Did not have significant amounts of restricted and prohibited medications Did not have other major protocol violations that would have impacted efficacy outcomes The per protocol population was used to analyse the primary and secondary efficacy variables in a supportive fashion.
KALM-1 OLE	Open-label safety population	Subjects who received at least one dose of open-label study drug during the open-label treatment period. Subjects in the open-label safety population were analysed according to the sequence of treatments received during the double-blind treatment period and the open-label treatment period.
KALM-2	Enrolled population	See description in KALM-1
	ITT population	See description in KALM-1
	Double-blind safety population	See description in KALM-1
	Per protocol population	Same as in KALM-1, except that subject needed a mean baseline WI-NRS score of ≥5.0 instead of >4.0
KALM-2 OLE	Open-label safety population	See description in KALM-1 OLE
CLIN3105	Enrolled population	See description in KALM-1

Study	Study population	Description
	Safety population	Subjects who received at least one dose of difelikefalin in the study. All summaries and analyses of safety, effectiveness, and additional endpoints were conducted using the safety population

Statistical analysis

Sources for this section: (32-36)

The objectives of KALM-1, KALM-2 and their OLEs were to evaluate the efficacy of IV difelikefalin (0.5 mcg/kg) compared to placebo in reducing the intensity of itch in haemodialysis subjects with moderate-to-severe pruritus. A summary of the statistical analysis of these trials is displayed in Table 14:

Table 14: Summary of statistical analyses

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
KALM-1	Primary efficacy outcomes: • The percentage of patients who had an improvement of ≥3 points from baseline at Week 12 in the weekly mean score on the daily WI-NRS. Prespecified secondary efficacy outcomes were: • Mean change from baseline at Week 12 in the 5-D Itch scale total score • Mean change from baseline at Week 12 in the Skindex-10 scale total score	In the primary analysis, for each imputed data set, the difference between placebo and difelikefalin were analysed using a logistic regression model containing terms for trial group, baseline WI-NRS score, baseline use of antipruritic medication, and history of prespecified medical conditions. The multiple imputation process was implemented separately for patients contributing to the interim assessment and those who underwent randomisation after the interim assessment.	A total of 378 patients underwent randomisation. It was calculated that, assuming a response in 30% of the placebo group, a planned sample of 350 patients would result in a 79%-90% or greater power to detect a difference of 15%-20% in the primary outcome, on the basis of a two-sided Chi square continuity corrected test at a significance level of 0.05. An interim analysis for sample size re-estimation was conducted by an independent data monitoring committee after 50% of the first 350	In the primary analysis, missing weekly mean WI-NRS scores were estimated with the use of multiple imputation, under a missing-at-random assumption. WI-NRS scores reported when patients were no longer receiving difelikefalin or placebo after the completion or discontinuation of the trial regimen were censored and treated as missing data.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
	• Percentage of patients who had a decrease of at least four points from baseline at Week 12 in the weekly mean WI-NRS score.	The final p-value was calculated with the use of the Cui–Hung–Wang weighted test statistic. Testing of the primary outcome was two-sided at an alpha level of 0.05. Secondary outcomes were analysed according to a prespecified hierarchy (first 5-D Itch scale, then Skindex-10 scale, and percentage of patients with a decrease of ≥4 points from baseline to Week 12 in the weekly mean WI-NRS score). The changes in scores on the 5-D and Skindex-10 scales at Week 12 were analysed with the use of an analysis of covariance (ANCOVA) model, with trial group as a fixed effect and baseline score and stratification factors as covariates. The percentage of patients who had a decrease of ≥4 points from baseline to Week 12 in the weekly mean WI-NRS score was analysed with the use of the method described for the primary outcome.	patients either completed the 12-week intervention period or discontinued the trial regimen. No change was made to the original enrolment target of 350 subjects.	

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		To control the type I error, a gatekeeping strategy was implemented. Testing of the secondary outcomes was to proceed only if the primary efficacy analysis was significant at the 5% level. Testing of the secondary outcomes was two-sided and performed sequentially with an alpha value of 0.05. All the efficacy analyses were conducted in the ITT population, which was defined as all the patients who underwent randomisation.		
KALM-1 OLE	Primary efficacy outcome: The change in total 5-D Itch score and change by domain score from baseline. Secondary efficacy outcomes are: • To evaluate the efficacy of IV difelikefalin at a dose of 0.5 mcg/kg compared to placebo in improving itch-related QoL measures in haemodialysis subjects	The 5-D Itch scale was the only measured used to evaluate efficacy in the OLE phase. The 5-D Itch scale scores will be analysed using a mixed model with repeated measures (MMRM). The model will contain treatment sequence, week, and treatment-by-week interaction as fixed effects, and baseline score and the randomisation stratification variables as covariates.	The sample size for the open-label extension phase was not defined a priori: all subjects who were eligible and willing to continue into the open-label extension phase were enrolled.	The scoring manual for 5-D Itch scale does not give specific direction regarding scoring when some questions are missing; therefore, each domain and the total score will be set to missing when any of their individual components are missing, with the exception of the disability domain. The maximum of any items present for disability will be used for that domain. Missing data will be handled implicitly in the MMRM model.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
	with moderate-to-severe pruritus. • To evaluate the safety of IV difelikefalin at a dose of 0.5 mcg/kg in haemodialysis subjects with moderate-to-severe pruritus.	Two independent analyses will be presented using different time points for the baseline values and changes from baseline using each of those baselines. In the first analysis, all visits in both the double-blind and the open-label treatment periods will be included; the baseline will be the 5-D Itch scale total score collected on Day 1, prior to randomisation. In the second analysis, only the visit in the open-label treatment period will be included; the baseline will be the last 5-D Itch scale total score in the double-blind treatment period. An unstructured covariance matrix will be used to model the within-subject errors. Should the model fail to converge, a compound symmetric covariance matrix will be used instead. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Missing scores will not be imputed. Assuming that the data are		Assuming that the data are MAR, the estimates calculated from the MMRM described in the statistical analysis.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		missing-at-random (MAR), the estimates calculated from the MMRM described above are unbiased. Standard descriptive statistics will be reported for each time point on the values and changes from baseline along with the least squares (LS) means, standard errors, 95% confidence intervals (CIs), and differences from baseline within each treatment sequence reported with LS means, standard errors, and 95% CIs. Plots will also be created. The above analyses for the 5-D Itch scale total score will be repeated for each of its domain scores.		
KALM-2	Please see KALM-1 hypothesis objective for primary and secondary efficacy outcomes	The efficacy of difelikefalin 0.5 mcg/kg compared to placebo in pivotal Phase 3 study KALM-2 will be evaluated based on one primary and seven secondary efficacy endpoints. • The proportion of subjects who have an improvement from	The planned sample size for this study was 350 (175 per treatment group) male and female haemodialysis subjects with chronic moderate-to-severe pruritus (mean baseline 24-hour WI-NRS score ≥5), randomised at approximately 95 clinical sites. The sample size	In the primary efficacy analysis, missing NRS data at the end of Week 12 will be imputed using a multiple imputation (MI) approach, assuming that subjects who discontinue double-blind treatment early would have similar Worst Itching Intensity NRS scores as other subjects

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score ≥3 points will be calculated for each imputed dataset. Differences between difelikefalin 0.5 mcg/kg and placebo with respect to the primary endpoint will be compared using a logistic regression model containing terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and presence of specific medical conditions. • The observed number and proportion of subjects with ≥ 3-point improvement among the non-imputed data will be reported along with the imputed data logistic regression model-based estimates of the proportions of	calculation was based on results of the completed Phase 2 double-blind, placebo-controlled study CR845-CLIN2101, which evaluated difelikefalin in subjects with ESRD and moderate-to-severe pruritus undergoing haemodialysis. Assuming a true response rate of 30% for the placebo group and a true response rate of 50% for the difelikefalin group (defining response as an improvement from baseline ≥3 points with respect to the WI-NRS at Week 12), a 2-sided continuity corrected Chi square would have 96% power to detect a treatment difference. The power of this test statistic would be ≥84% for differences from placebo as low as 0.16. Based on the results of a planned interim assessment conducted when approximately 50% of the 350 patients had either completed the 12-week double-blind treatment period or had	in their respective treatment arm who have complete data: • Intermittent missing NRS scores will first be imputed using the Markov Chain Monte Carlo method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data. • For each stage, MI will be performed within treatment group with covariates for baseline NRS score, both randomisation stratification factors, region and the non-missing NRS scores for each week. Should convergence issues occur due to small cell size for the categorical covariates corresponding to strata (at either stage), those specific covariates will be removed from the model.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		responders, odds ratio, 95% CIs, and p-value.	discontinued from treatment early, the size of the study was increased by approximately 20%, to 430 subjects.	
KALM-2 OLE		Please see KALM-1 OLE statistical analysis. Additionally: As a separate analysis, the number and percentage of subjects who have a 5-point or greater improvement will be reported by visit and treatment sequence. This will be repeated as above for each baseline.	The sample size for the open-label extension phase was not defined a priori: all subjects who were eligible and willing to continue into the open-label extension phase were enrolled.	The scoring manual does not give specific direction regarding scoring when some questions are missing; therefore, each domain and the total score will be set to missing when any of their individual components are missing, with the exception of the disability domain. The maximum of any items present for disability will be used for that domain. Missing data will be handled implicitly in the MMRM model. Missing scores will not be imputed. Assuming that the data are MAR, the estimates calculated from the MMRM described in the statistical analysis section are unbiased.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
CLIN3105	Primary objective: • To evaluate the safety of difelikefalin at a dose of 0.5 mcg/kg IV in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus. Secondary objectives: • To evaluate the effectiveness of difelikefalin at a dose of 0.5 mcg/kg IV in reducing the intensity of itch in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus • To evaluate the effectiveness of difelikefalin at a dose of 0.5 mcg/kg IV in improving itch-related QoL and quality of sleep measures in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus.	This study uses the following five instruments to assess effectiveness: • WI-NRS • Sleep quality questionnaire • 5-D Itch Scale • Skindex-10 • EQ-5D-5L-P No primary efficacy endpoint was defined. All effectiveness analyses were performed on the safety population. For the WI-NRS, Sleep Quality Questionnaire, 5-D itch scale, and the Skindex-10 scale, summary statistics (n, mean, SD, minimum, maximum) for the respective baseline and Week 12 score were produced, along with the change from baseline. For the WI-NRS and Sleep Quality Questionnaire, the count and percentage of subjects with an improvement in WI-NRS from baseline of >0, ≥1, ≥2, ≥3, ≥4, ≥5, and ≥6-points at Week 12 were	Approximately 200 male and female with moderate-to-severe pruritus undergoing haemodialysis were to be enrolled in this study at approximately 50 US and non-US clinical sites. No sample size calculation was performed to select this sample size.	Missing data will not be imputed. Data from subjects who terminated prematurely will be included in any analyses for which their data is available, unless otherwise specified. Please see Section 8.2 of the SAP for further details.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		reported. The count and percentage of subjects with an improvement from baseline of ≥3 and ≥4-points at Week 12 were also reported and stratified by region. Quantitative laboratory parameters were summarised using descriptive statistics for observed values and the changes from baseline to each time point (when applicable), including the designation of last post-baseline treatment visit. Observed measurements of vital signs and the changes from baseline were summarised using descriptive statistics (n, mean, SD, median, minimum, and maximum) for baseline, each post-baseline assessment, and the last post-baseline treatment visit.		

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

The tables below assess the relevant clinical effectiveness evidence, using criteria taken from the NICE User Guide. Please see Appendix D.1.3 for full quality assessment and Section B.2.12 for further discussion on the strengths and limitations of the clinical evidence base.

Table 15: Quality assessment of KALM-1

Trial number (acronym)	Response
Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes
Were there any unexpected imbalances in drop-outs between groups?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes

Table 16: Quality assessment of KALM-2

Question	Response
Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes
Were there any unexpected imbalances in drop-outs between groups?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes

Table 17: Quality assessment of KALM-1 OLE

Question	Response
Was the cohort recruited in an acceptable way?	Yes

Question	Response
Was the exposure accurately measured to minimise bias?	Yes
Was the outcome accurately measured to minimise bias?	Yes
Have the authors identified all important confounding factors?	Yes
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes
Was the follow-up of patients complete?	Yes
How precise (for example, in terms of confidence interval and p values) are the results?	See Appendix D.1.3

Table 18: Quality assessment of KALM-2 OLE

Question	Response
Was the cohort recruited in an acceptable way?	Yes
Was the exposure accurately measured to minimise bias?	Yes
Was the outcome accurately measured to minimise bias?	Yes
Have the authors identified all important confounding factors?	Yes
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes
Was the follow-up of patients complete?	No
How precise (for example, in terms of confidence interval and p values) are the results?	See Appendix D.1.3

Table 19 Quality assessment of CLIN3105

Question	Response
Was the cohort recruited in an acceptable way?	Yes
Was the exposure accurately measured to minimise bias?	Yes
Was the outcome accurately measured to minimise bias?	Yes
Have the authors identified all important confounding factors?	Yes
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes
Was the follow-up of patients complete?	Yes
How precise (for example, in terms of confidence interval and p values) are the results?	See Appendix D.1.3

B.2.6 Clinical effectiveness results of the relevant studies

KALM-1 (24)

Primary efficacy endpoint

The primary efficacy endpoint was the proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period. The results for the ITT poulation based on the combined imputed data from interim and post-interim analysis subjects have been summarised in Table 20. At Week 12, the LS mean percentage of subjects with at least a 3-point improvement from baseline in the WI-NRS was 51.0% in the difelikefalin group, compared with 27.6% in the placebo group. The odds ratio for a ≥3-point improvement from baseline with difelikefalin versus placebo was 2.72 (95% CI, 1.72 to 4.30), which was statistically significant (p<.001).

Table 20: Primary analysis: subjects with a ≥3-point improvement from baseline at Week 12 with respect to the Worst Itching Intensity NRS score – MI with MAR assumption (population:ITT)

Combined estimates at Week 12	Placebo (N=189)	DFK (N = 189)
Observed ≥3-point NRS improvement [1] - n (%)		
Yes	51 (30.9%)	82 (52.2%)
No	114 (69.1%)	75 (47.8%)
Missing	24	32
LS means estimate of percent with improvement [2]		
Percent (95% CI)	27.6% (20.2%, 36.6%)	51.0% (42.9%, 58.9%)
LH odds ratio (95% CI)	-	2.72 (1.72, 4.30)
CHW p-value	-	<.001

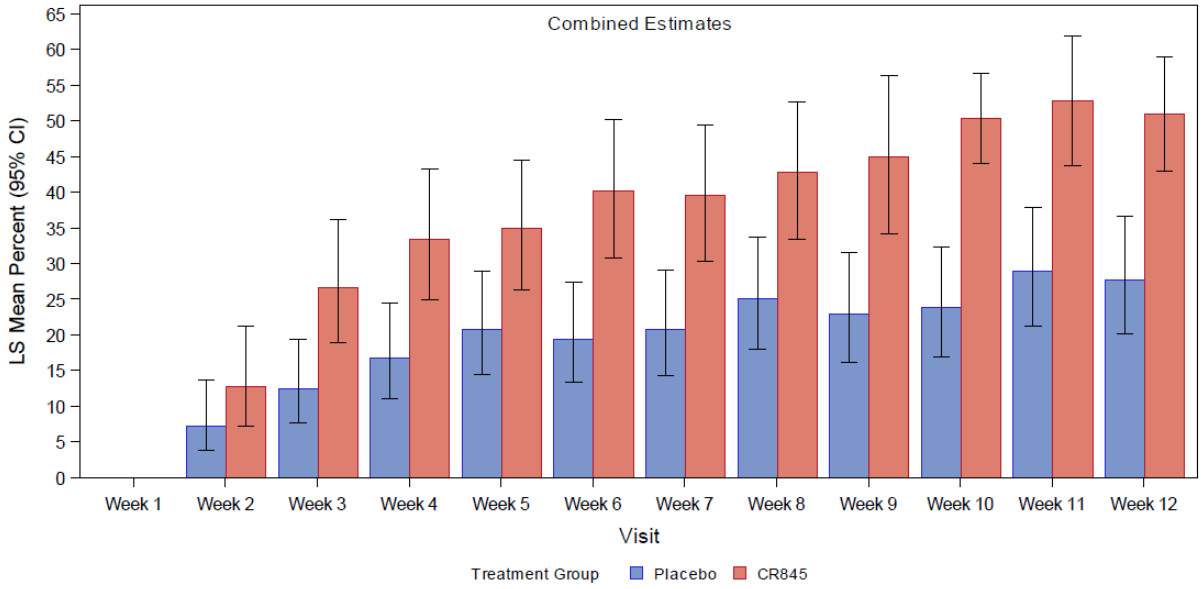
CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = numerical rating scale. [1] Counts and percentages were based on non-missing data.
[2] Estimated percent, odds ratio, and p-value used a logistic regression model with terms for treatment group, baseline Worst Itching Intensity NRS score, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately. Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the LH/CHW methodology.

When the primary analysis was conducted separately for interim analysis subjects and post-interim analysis subjects, the results were consistent with the combined analysis presented (Table 20). In interim analysis subjects, the odds ratio for

achieving a ≥3-point improvement from baseline in the WI-NRS at Week 12 with difelikefalin versus placebo was 3.31 (95% CI, 1.67 to 6.57; p<.001); in post-interim subjects, the odds ratio was 2.20 (95% CI, 1.21 to 3.99; p=.009).

Error! Reference source not found. depicts the LS mean percentage of ITT subjects with a ≥3-point improvement from baseline in WI-NRS by study week (Week 12 being the primary efficacy time point). A statistically significant treatment group difference favouring difelikefalin was observed as early as Week 3 (p<.001); this was maintained throughout the remainder of the double-blind treatment period. At Week 4, the LS mean percentage of subjects in the difelikefalin group with a ≥3-point improvement from baseline in WI-NRS was 33.5%, versus 16.7% for the placebo group (p<.001). At Week 8, the respective percentages were 42.7% versus 25.1% (p<.001)

Figure 6: Percentage of subjects with a ≥3-point improvement in Worst Itching Intensity Numerical Rating Scale Score by Week (primary efficacy imputation) (Population: ITT)



CI = confidence interval; ITT = Intent-to-treat; LS = least squares. Note: difelikefalin is referred to as its previous name 'CR845' in this diagram. Note: Estimated percentages and CIs used a logistic regression model with terms for treatment group, baseline Worst Itching Intensity NRS score, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under a MAR missing data assumption for interim subjects and post-interim subjects separately. Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the Lawrence, Hung/Cui, Hung, Wang methodology.

Various supportive and sensitivity analyses showed efficacy with difelikefalin consistent with the efficacy shown in the primary efficacy analysis (Table 21). The

findings support the robustness of the results across multiple statistical methodologies, i.e., significant improvement in WI-NRS at Week 12 with difelikefalin treatment compared with placebo treatment.

Sensitivity and supportive analyses of the primary efficacy analysis

The analysis of the primary endpoint was repeated for the per protocol population as a supportive analysis (Table 21). The results matched those of the primary analysis for the ITT population. At Week 12, the LS mean percentage of subjects (interim and post-interim combined) with at least a 3-point improvement from baseline in the WI-NRS was 50.4% in the difelikefalin group, compared with 27.0% in the placebo group. The odds ratio with difelikefalin versus placebo was 2.74 (95% CI, 1.71 to 4.41), which was statistically significant (p <.001). Table 21 summarises the analysis of the primary efficacy endpoint without using the Cui, Hung, Wang adjustment procedure and without splitting of the data with respect to interim versus post-interim status. This analysis also showed statically significant (P <.001) results favouring difelikefalin at Week 12, with an odds ratio (difelikefalin versus placebo) for a ≥3-point improvement from baseline in WI-NRS score of 2.62 (95% CI, 1.68 to 4.09).

Table 21 Key results of supportive and sensitivity analyses of the primary efficacy endpoint – percentage of subjects with a ≥3-point improvement in Worst Itching Intensity NRS at Week 12 (population: ITT and per protocol)

Analysis statistic	Placebo	DFK
Sensitivity analyses		
Subjects who discontinued early as non-responders [1]	-	-
N	189	189
LS mean percent with improvement (95% CI)	26.0% (19.0%, 34.5%)	44.6% (35.4%, 54.2%)
LH odds ratio (95% CI)	-	2.29 (1.46, 3.60)
CHW p-value	-	<.001
MI with missing-not-at-random (MNAR) assumption [1]	-	-
N	189	189
LS mean percent with improvement (95% CI)	27.6% (20.2%, 36.4%)	44.6% (35.4%, 54.2%)
LH odds ratio (95% CI)	-	2.33 (1.47, 3.71)
CHW p-value	-	<.001

Analysis statistic	Placebo	DFK
Tipping point [1]	-	-
N	189	189
Highest shift parameter without tipping	6.50	6.50
Percent with improvement (95% CI)	29.1% (21.5%, 38.1%)	42.8% (33.7%, 52.4%)
LH odds ratio	-	1.82 (1.16, 2.86)
CHW p-value	-	.009
Additional analysis		
Per Protocol Population [1]		
N	169	163
LS mean percent with improvement (95% CI)	27.0% (19.1%, 36.6%)	50.4% (47.1%, 53.6%)
LH odds ratio (95% CI)		2.74 (1.71, 4.41)
CHW p-value		<.001
No CHW adjustment for interim analysis [1]		
N	189	189
LS mean percent with improvement (95% CI)	28.3% (21.0%, 37.1%)	50.9% (41.6%, 60.2%)
LH odds ratio (95% CI)		2.62 (1.68, 4.09)
P-value		<.001

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; MNAR = missing-not-at-random

[1] Analysis based on interim and post-interim subjects combined.

Key Secondary Efficacy Endpoints

Itch-related QoL – change from baseline in total 5-D Itch scale score at end of Week 12

The first secondary efficacy endpoint was the change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the 5-D Itch scale (total score). The five dimensions of itch assessed are degree, duration, direction, disability, and distribution (37). Each domain is scored 1-5 with a total score range of 5-25 (5=no pruritus, 25=most severe pruritus). A 5-point change is considered clinically significant (37); thus, the observed improvement in the difelikefalin group (-5.0) was found to be clinically significant, whereas mean improvement for placebo group (-3.7) was not (Table 22). Compared with the

placebo group, the difelikefalin group again showed a statistically significant (p<.001) reduction in total 5-D Itch scale score at the end of Week 12, with a LS mean treatment group difference of -1.3 (95% CI, -2.0 to -0.5) (Table 22). The findings for the ANCOVA analysis in the per protocol population were also in favour of difelikefalin and statistically significant (p<.001), with an LS mean treatment group difference of -1.4 (-2.2, -0.6).

Table 22: ANCOVA analysis of change from baseline in total 5-D Itch score at Week 12 - multiple imputation (population: ITT)

	Placebo (N=189)	DFK (N=189)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-3.7	-5.0	-1.3	<.001
(SE)	(0.33)	(0.33)	(0.38)	-
95% CI	(-4.4, -3.1)	(-5.7, -4.4)	(-2.0, -0.5)	-

Abbreviations: ANCOVA = analysis of covariance; CI = confidence interval; ITT = Intent-to-treat; LS = least squares; SE = standard error. Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption.

An MMRM sensitivity analysis was performed on the change from baseline in the total 5-D Itch scale score by time point with no data imputation. At the end of Week 12, the LS mean change from baseline in total 5-D Itch scale score was -4.9 (95% CI, -5.6 to -4.3) in the difelikefalin group and -3.6 (95% CI, -4.2 to -2.9) in the placebo group. The LS mean treatment group difference (difelikefalin minus placebo) of -1.3 (95% CI, -2.2 to -0.5) was statistically significant (p=.002) in favour of difelikefalin.

Itch-related QoL – change from baseline in Skindex-10 scale score at end of Week 12

The second secondary efficacy endpoint was the change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the total Skindex-10 scale score. The Skindex-10 scale is a patient-reported measurement of itch and its impact on QoL in the last week, and has been specifically developed for CKD-aP. It consists of 10 questions across three domains (disease, mood/emotional stress, and social functioning). Each of the 10 questions is scored from 0–6 (0=never bothered; 6=always bothered), meaning the total score varies from 0-60. A 15-point change in score is regarded as clinically significant (7). At the end of Week 12, the LS mean change in total Skindex-10 scale score was

greater in the difelikefalin group than in the placebo group (-17.2 versus -12.0) (Table 23). The change from baseline was considered clinically significant for difelikefalin, but not for placebo; a statistically significant LS mean difference was also noted: -5.1 (95% CI, -8.0 to -2.3); p<.001 (Table 23). The findings for the per protocol population were also in favour of difelikefalin and statistically significant (p<.001).

Table 23: ANCOVA Analysis of change from baseline in total Skindex-10 scale at Week 12 - MI (Population: ITT)

	Placebo (N=189)	DFK (N=189)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-12.0	-17.2	-5.1	<.001
(SE)	(1.24)	(1.26)	(1.44)	-
95% CI	(-14.5, -9.6)	(-19.6, -14.7)	(-8.0, -2.3)	-

Abbreviations: ANCOVA = analysis of covariance; CI = confidence interval, LS = least squares; SE = standard error; ITT = Intent-to-treat. Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption.

The results of the MMRM sensitivity analysis (no imputation) of total Skindex-10 scale at Week 12 for the ITT population were similar to ANCOVA with MI. At the end of Week 12, the LS mean change from baseline in total Skindex-10 scale score was -17.4 (95% CI, -19.9 to -14.8) in the difelikefalin group and -12.2 (95% CI -14.7 to -9.6) in the placebo group. The treatment group difference (difelikefalin minus placebo) of -5.2 (95% CI -8.3 to -2.1) was in favour of difelikefalin and statistically significant (p<0.001).

≥4-Point improvement in weekly 24-Hour Worst Itching Intensity NRS

The third key secondary efficacy endpoint was the proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period. Table 24 summarises the analysis of this endpoint for the ITT population, which was conducted identically to the primary analysis of the primary endpoint. At Week 12, the LS mean percentage of subjects with a ≥4-point improvement in WI-NRS from baseline was 38.9% in the difelikefalin group and 18.0% in the placebo group; the odds ratio with difelikefalin was 2.89 (95% CI, 1.75 to 4.76), which was statistically significant (p<.001). When the analysis was conducted separately for interim and post-interim

subjects, the Week 12 results were consistent with the combined analysis, in favour of difelikefalin, and statistically signifiacnt (p=.006 for interim subjects; p=.002 for post-interim subjects).

The same analysis was conducted using the per protocol population and showed a percentage of difelikefalin subjects than placebo subjects achieving a ≥4-point improvement from baseline in WI-NRS, a result which was statistically significant (p<.001).

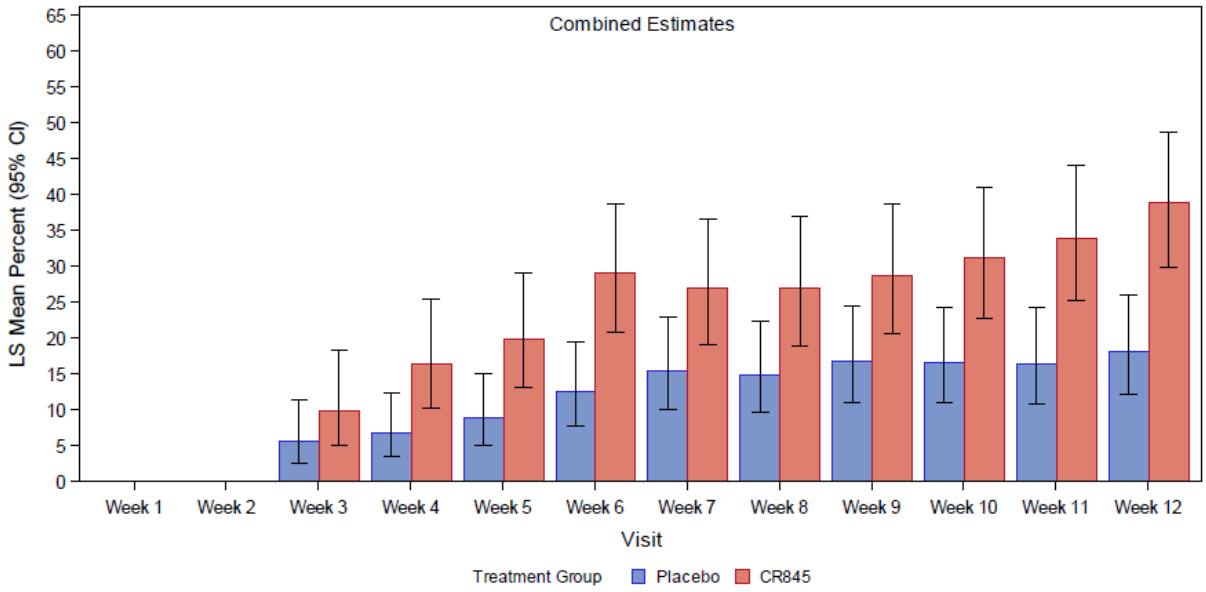
Table 24*:* Subjects with a ≥4-point improvement from baseline at Week 12 in Worst Itching Intensity NRS Score – MI with MAR assumption (population: ITT)

Combined assessments (Week 12)	Placebo (n=189)	DFK (n=189)
Observed ≥4-point NRS improvement [1] – n (%)		
Yes	35 (21.2%)	64 (40.8%)
No	130 (78.8%)	93 (59.2%)
Missing	24	32
LS means estimate of percent with improvement [2]		
Percent (95% CI)	18.0% (21.1%, 26.0%)	38.9% (29.8%, 48.7%)
LH odds ratio (95% CI)		2.89% (1.75, 4.76)
CHW p-value		<.001

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = numerical rating scale. [1] Counts and percentages were based on non-missing data. [2] Estimated percent, odds ratio, and p-value use a logistic regression model with terms for treatment group, baseline Worst Itching Intensity NRS score, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.

Error! Reference source not found. depicts the percentage of ITT subjects with a ≥4-point improvement in the WI-NRS by study week. A statistically significant (p≤.05) treatment group difference favouring difelikefalin was observed by Week 4 (p=.003), which was maintained throughout the remainder of the double-blind treatment period. At Week 4, the LS mean percentage of subjects in the difelikefalin group with a ≥4-point improvement from baseline in WI-NRS was 16.4% versus 6.6% for the placebo group (p=.003), and at Week 8, the respective percentages were 26.9% versus 14.9% (p=.005).

Figure 7: Percentage of subjects with a ≥4-Point improvement in Worst Itching Intensity NRS Score by week (MI with MAR assumption) (population: ITT)



CI = confidence interval; ITT= Intent-to-treat; LS = least squares. Note: difelikefalin is referred to as its previous name, 'CR845', in this diagram Note: Estimated percentages and CIs used a logistic regression model with terms for treatment group, baseline. Worst Itching Intensity NRS score, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under a MAR missing data assumption for interim subjects and post-interim subjects separately. Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the Lawrence, Hung/Cui, Hung, Wang methodology

A complete responder was a subject with ≥80% of the non-missing 24-hour WI-NRS scores equal to 0 or 1 on Week 12. Subjects who reported fewer than 4 WI-NRS scores and subjects who dropped out prior to Week 12 were considered non-responders.

The Patient Global Impression of Change (PGIC) is a global PRO measure which assesses the overall change in itch (no change, improvement, or worsening) relative to the start of the study (38). The scale has only one item: each subject was asked to mark the category that best described the change in itch, ranging from "Very Much Improved" to "Very Much Worse". [redacted]

[redacted]

KALM-1 OLE (26)

In the first analysis of treatment effect, all visits in both the double-blind treatment period and open-label treatment period were included; the baseline was the 5-D Itch scale total score collected on Day 1 of the double-blind treatment period, prior to randomisation (i.e., double-blind baseline). In the second analysis, only the visits in the open-label treatment period were included; the baseline was the last 5-D Itch scale total score collected in the double-blind treatment period (i.e., open-label baseline).

[redacted]

When subjects randomised to placebo during the double-blind treatment period transitioned to active treatment (between double-blind Week 12 and open-label Week 4), [redacted] Over the course of long-term treatment through Week 52, the total number of subjects [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

	[redacted]		[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

At the beginning of the open-label treatment period (end of open-label Week 4), [redacted] The number of double-blind difelikefalin subjects with an improvement of at least five points from the double-blind baseline [redacted] Throughout the open-label treatment period, [redacted]

[redacted]

[redacted]

The mean (SD) open-label baseline 5-D Itch score (i.e. the value from the last 5-D Itch scale total score collected in the double-blind treatment period) was [redacted] for subjects who were randomised to placebo and difelikefalin in the double-blind treatment period, respectively. During the open-label treatment period,

[redacted]

The percentage of subjects with an improvement of at least five points from the open-label baseline was [redacted]

The same analyses performed for the total 5-D Itch were also performed for the individual 5-D Itch domains of degree, duration, direction, disability, and distribution. The individual 5-D Itch domains followed the same trends as the total 5-D Itch score over double-blind and open-label visits. [redacted]

KALM-2 (25)

Primary efficacy endpoint

The primary efficacy endpoint was the proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period. Table 25 summarises these results for the ITT population, based on the combined data from interim and post-interim analysis subjects. At Week 12, the LS mean percentage of subjects with at

least a 3-point improvement from baseline in the WI-NRS was 54.0% in the difelikefalin group, compared with 42.2% in the placebo group. The estimated odds ratio for a ≥3-point improvement from baseline with difelikefalin versus placebo was 1.61 (95% CI, 1.08 to 2.41), which was statistically significant (p=.020).

Table 25: Analysis: Subjects with ≥3-point improvement from baseline at Week 12 with respect to the Worst Itching Intensity NRS score – multiple imputations with MAR assumption (population: ITT)

Combined estimates (Week 12)	Placebo (n=236)	DFK
Observed ≥3-point NRS improvement [1] - n (%)		
Yes	77 (33.2%)	95 (49.7%)
No	130 (62.8%)	96 (50.3%)
Missing	29	46
LS means estimate of percent with improvement [2]		
Percent (95% CI)	42.2% (32.5%, 52.5%)	54.0% (43.9%, 63.9%)
LH odds ration (95% CI)		1.61 (1.08, 2.41)
CHW p-value		0.020

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = Numerical Rating Scale
[1] Counts and percentages were based on non-missing data.
[2] Estimated percentage, odds ratio and p-value used a logistic regression model with terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.
Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate

Analyses of the primary efficacy endpoint were conducted separately for interim analysis subjects and post-interim analysis subjects and are consistent with the combined analysis presented above. For interim analysis subjects, the odds ratio for achieving a ≥3-point improvement from baseline in the WI-NRS at Week 12 with difelikefalin versus placebo was 1.88 (95% CI, 0.97 to 3.65); in post-interim subjects, the odds ratio was 1.42 (95% CI, 0.88 to 2.30).

Sensitivity and supportive analyses of the primary efficacy analysis

Table 26 summarises key results of the three sensitivity analyses of the primary efficacy endpoint, conducted to evaluate the robustness of the study results under different assumptions and imputation algorithms. Additionally, results of the analysis on the per protocol population and without the Cui, Hung, Wang adjustment are presented.

Table 26 Key results of sensitivity and supportive analyses of the primary efficacy endpoint – percentage of subjects with a ≥3-point improvement in Worst Itching Intensity NRS at Week 12 (population: ITT and per protocol)

Analysis statistic	Placebo	DFK
Sensitivity analyses		
Subjects who discontinued early as non-responders [1]		
N	236	237
LS mean percent with improvement (95% CI)	37.2% (27.8%, 47.6%)	47.7% (33.4%, 54.7%)
LH odds ratio (95% CI)	-	1.31 (0.89, 1.94)
CHW p-value	-	0.168
MI with MNAR assumption [1]		
N	236	237
LS mean percent with improvement (95% CI)	39.9% (30.6%, 50.1%)	50.7% (41.2, 60.1%)
LH odds ratio (95% CI)	-	1.55 (1.05, 2.28)
CHW p-value	-	0.029
Tipping point [1]		
N	236	237
Highest shift parameter without tipping	0.75	0.75
Percent with improvement (95% CI)	41.9% (32.0%, 52.4%)	52.1% (42.5%, 61.5%)
LH odds ratio	-	1.51 (1.01, 2.35)
CHW p-value	-	0.044
Additional analysis		
Per protocol population [1]		
N	213	205
LS mean percent with improvement (95% CI)	39.7 (29.7%, 50.7%)	52.0% (43.8%, 60.2%)
LH odds ratio (95% CI)	-	1.65 (1.08, 2.51)
CHW p-value	-	0.019
No CHW adjustment for interim analysis [1]		
N	236	237
LS mean percent with improvement (95% CI)	42.6% (33.4%, 52.3%)	53.4% (43.7%, 62.8%)
LH odds ratio (95% CI)	-	1.54 (1.05, 2.27)
P-value	-	0.027

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; MNAR = missing-not-at-random. [1] Analysis based on interim and post-interim subjects combined.

Key secondary endpoints

The key secondary efficacy endpoints were analysed in a hierarchical testing order and are summarised in Table 27. If an endpoint did not reach statistical significance, then each subsequent endpoint was not considered significant.

Table 27: Hierarchical testing order of key secondary endpoints

Secondary endpoints	Nominal p-value	Significant? (yes or no)
Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period	0.010	Yes
Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 8 of the double-blind treatment period	0.010	Yes
Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 4 of the double-blind treatment period	0.002	Yes
Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 8 of the double-blind treatment period	0.010	Yes
Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 4 of the double-blind treatment period	0.036	Yes
Change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the total Skindex-10 scale score	0.171	No
Change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the 5-D Itch scale score	0.002	No

QoL = quality of life; WI-NRS = Worst Itching Intensity numerical rating scale

≥4-Point improvement in weekly mean 24-Hour Worst Itching Intensity NRS at week 12

The first key secondary efficacy endpoint was the proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period. Table 28 summarises the analysis of this endpoint for the ITT population, which was conducted in a manner identical to that employed in the primary analysis of the primary endpoint.

At Week 12, the LS mean percentage of subjects with a ≥4-point improvement in WI-NRS from baseline was 41.2% in the difelikefalin group and 28.4% in the placebo group; the odds ratio was 1.77 (95% CI, 1.14 to 2.74), which was statistically significant (p=.010).

Table 28 Subjects with a ≥4-point improvement from baseline at Week 12 in Worst Itching Intensity NRS Score – MI with MAR assumption (population: ITT)

Combined estimates (Week 12)	Placebo (n=236)	DFK (n=237)
Observed ≥4-point NRS improvement [1] - n(%)		
Yes	52 (25.1%)	72 (37.7%)
No	155 (74.9%)	119 (62.3%)
Missing	29	46
LS means estimate of percent with improvement [2]		
Percent (95% CI)	28.4% (21.3%, 37.7%)	41.2% (33.0%, 50.0%)
LH odds ratio (95% CI)	-	1.77 (1.14, 2.74)
CHW p-value	-	0.010

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = numerical rating scale. [1] Counts and percentages were based on non-missing data.
[2] Estimated percent, odds ratio, and p-value use a logistic regression model with terms for treatment group, baseline Worst Itching Intensity NRS score, region, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately. Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the LH/CHW methodology.

When the analysis was conducted separately for interim and post-interim subject, the Week 12 results were consistent with the combined analysis.

≥3-Point improvement in weekly mean 24-Hour Worst Itching Intensity NRS at Weeks 8 and 4

The second and third key secondary efficacy endpoints were the proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Weeks 8 and 4 of the double-blind treatment period, respectively. Table 29 summarises the analysis of these secondary endpoints for the ITT population, which was conducted in a manner identical to that employed in the primary analysis of the primary endpoint. At Week 8, the LS mean percentage of subjects with a ≥3-point improvement in WI-NRS from baseline was 49.0% in the difelikefalin group and 36.2% in the placebo group; the odds ratio was 1.69 (95% CI, 1.13 to 2.53), which was statistically significant (p=.010). At Week 4, the LS mean percentage of subjects with a ≥3-point improvement in WI-NRS from

baseline was 38.3% in the difelikefalin group and 23.8% in the placebo group; the odds ratio was 1.99 (95% CI, 1.29 to 3.06), which was statistically significant (p=.002).

Table 29: Subjects with a ≥3-point improvement from baseline at Weeks 8 and 4 in Worst Itching Intensity NRS Score – MI with MAR Assumption (population: ITT)

Combined estimates	Placebo (n=236)	Difelikefalin (n=237)
Week 8		
Observed ≥3-point NRS improvement [1] – n (%)		
Yes	73 (33.0%)	93 (44.5%)
No	148 (67.0%)	116 (55.5%)
Missing	15	28
LS means estimate of percent with improvement [2]		
Percent (95% CI)	36.2% (27.3%, 46.2%)	49.0% (38.3%, 59.9%)
LH odds ratio (95% CI)	-	1.69 (1.13, 2.53)
CHW p-value	-	0.010
Week 4		
Observed ≥3-point NRS improvement [1] - n (%)		
Yes	50 (22.2%)	75 (35.0%)
No	175 (77.8%)	139 (65.0%)
Missing	11	23
LS means estimate of percent with improvement [2]		
Percent (95% CI)	23.8% (16.6%, 32.8%)	38.3% (28.5%, 49.1%)
LH odds ratio (95% CI)	-	1.99 (1.29, 3.06)
CHW p-value	-	0.002

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = Numerical Rating Scale. [1] Counts and percentages were based on non-missing data.

[2] Estimated percentage, odds ratio, and p-value used a logistic regression model with terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.

Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the LH/CHW methodology

≥4-Point improvement in weekly mean 24-Hour Worst Itching Intensity NRS at Weeks 8 and 4

The fourth and fifth secondary efficacy endpoints were the proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Weeks 8 and 4 the double-blind treatment period, respectively. Table 30 summarises the analysis of these secondary endpoints for the

ITT population, which was conducted in a manner identical to that employed in the primary analysis of the primary endpoint. At Week 8, the LS mean percentage of subjects with a ≥4-point improvement in WI-NRS from baseline was 36.1% in the difelikefalin group and 23.7% in the placebo group; the odds ratio was 1.82 (95% CI, 1.16 to 2.86), which was statistically significant (p=.010). At Week 4, the LS mean percentage of subjects with a ≥4-point improvement in WI-NRS from baseline was 26.1% in the difelikefalin group and 16.7% in the placebo group; the odds ratio was 1.76 (95% CI, 1.04 to 2.98), which was statistically significant (p=.036).

Table 30: Subjects with a ≥4-point improvement from baseline at Weeks 8 and 4 in Worst Itching Intensity NRS Score – MI with MAR assumption (population: ITT)

Combined estimates	Placebo (n=236)	Difelikefalin
Week 8		
Observed ≥4-point NRS improvement [1] – n (%)		
Yes	45 (20.4%)	64 (30.6%)
No	176 (79.6%)	145 (69.4%)
Missing	15	28
LS means estimate of percent with improvement [2]		
Percent (95% CI)	23.7% (17.2%, 31.8%)	36.1% (28.0%, 45.1%)
LH odds ratio (95% CI)	-	1.82 (1.16, 2.86)
CHW p-value	-	0.010
Week 4		
Observed ≥4-point NRS improvement [1] - n (%)		
Yes	30 (13.3%)	43 (20.1%)
No	195 (86.7%)	171 (79.9%)
Missing	11	23
LS means estimate of percent with improvement [2]		
Percent (95% CI)	16.7% (11.4%, 23.9%)	26.1% (18.8%, 34.9%)
LH odds ratio (95% CI)	-	1.76 (1.04, 2.98)
CHW p-value	-	0.036

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = numerical rating scale. [1] Counts and percentages were based on non-missing data. [2] Estimated percent, odds ratio, and p-value use a logistic regression model with terms for treatment group, baseline Worst Itching Intensity NRS score, region, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately. Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the LH/CHW methodology.

Itch-related QoL – change from baseline in total Skindex-10 scale score at end of Week 12

The next key secondary efficacy endpoint was the change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the total Skindex-10 scale score. Table 31 summarises the change from baseline in total Skindex-10 scale score at the end of Week 12 for the ITT population, using ANCOVA with MI under the MAR assumption. Compared with the placebo group, the difelikefalin group showed a numerically greater reduction in LS mean total Skindex-10 scale score (-16.6 versus -14.8) at the end of Week 12, with a LS mean treatment group difference of -1.8 (95% CI, -4.3 to 0.8), which was not statistically significant (p=.171). The findings for the per protocol population also showed a numerically greater reduction in total Skindex-10 scale score favouring the difelikefalin group compared with placebo (-17.4 versus -14.8, respectively); the LS mean treatment group difference of -2.6 (95% CI, -5.3 to 0.2), although larger in absolute terms than for the ITT population, was not statistically significant (p=.064)

Table 31 ANCOVA Analysis of change from baseline in total Skindex-10 scale at Week 12 - MI under MAR assumption (population: ITT)

End Of Week 12 change from baseline	Placebo (N=236)	Difelikefalin (N=237)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-14.8	-16.6	-1.8	0.171
(SE)	(1.32)	(1.35)	(1.29)	-
95% CI	(-17.4, -12.2)	(-19.3, -14.0)	(-4.3, 0.8)	-

ANCOVA = analysis of covariance; CI = confidence interval, LS = least squares; SE = standard error; ITT = Intent-to-treat Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption.

MMRM and ANCOVA sensitivity analysis were performed for the Skindex-10 scale. All analyses were consistent with the key analysis.

Itch-related QoL – change from baseline in total 5-D Itch scale score at end of Week 12

The final secondary efficacy endpoint was the change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the total score of the 5-D Itch scale. Table 32 summarises the change from baseline in total 5-D Itch scale score at the end of Week 12, using ANCOVA with MI of missing data under a MAR assumption. Compared with the placebo group, the difelikefalin group showed a greater reduction in total 5-D Itch scale score at the end of Week

12, with a LS mean treatment group difference of -1.1 (95% CI, -1.7 to -0.4). Although the nominal p-value was 0.002, this difference could not be declared to be statistically significant based on the hierarchical testing order, as the prior secondary endpoint (Skindex-10 at Week 12) was not statistically significant. Additionally, the findings for the ANCOVA analysis in the per protocol population also showed a greater reduction in total 5-D Itch scale score at the end of Week 12 for the difelikefalin group, with a LS mean treatment group difference of -1.3 (95% CI, -2.0 to -0.6).

Table 32: ANCOVA analysis of change from baseline in total 5-D Itch score at Week 12 - MI (population: ITT)

End of Week 12 change from baseline	Placebo (n=236)	Difelikefalin (n=237)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-3.8	-4.9	-1.1	0.002
(SE)	(0.36)	(0.36)	(0.35)	-
95% CI	(-4.5, -3.1)	(-5.6, -4.2)	(-1.7, -0.4)	-

ANCOVA = analysis of covariance; CI = confidence interval; ITT = Intent-to-treat; LS = least squares; MAR = missing-at-random; SE = standard error. Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score, region, and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption.

Sensitivity analyses were performed for the 5-D Itch scale secondary efficacy outcome. One such analysis was an MMRM sensitivity analysis of the change from baseline in the total 5-D Itch scale score by time point with no imputation for missing data. At the end of Week 12, the LS mean change from baseline in total 5-D Itch scale score was -5.1 (95% CI, -5.8 to -4.4) in the difelikefalin group and -3.9 (95% CI, -4.6 to -3.3) in the placebo group. The LS mean treatment group difference (difelikefalin minus placebo) of -1.2 (95% CI, -1.9 to –0.5) was in favour of difelikefalin.

An ANCOVA sensitivity analysis was conducted of the change from baseline in 5-D Itch scale at the end of Week 12 for the ITT population, with MI of missing data using control distribution. Compared with the placebo group, the difelikefalin group showed a greater reduction in total 5-D Itch scale score (-4.8 versus -3.8) at the end of Week 12, with a LS mean treatment group difference of -1.0 (95% CI, -1.7 to -0.3).

An ANCOVA sensitivity analysis of the change from baseline in 5-D Itch scale at the end of Week 12 for the ITT Population, with MI of missing data using baseline

distribution was also conducted. Compared with the placebo group, the difelikefalin group showed a greater reduction in total 5-D Itch scale score (-4.3 versus -3.5) at the end of Week 12, with a LS mean treatment group difference of -0.7 (95% CI, -1.4 to -0.0).

The PGIC results were calculated in the same manner as KALM-1, with the subject frequency of different responses for the ITT population, as well as the percentage of subjects who were responders, (i.e., had responses of “Very much improved” or “Much improved”) being calculated. [redacted]

KALM-2 OLE (27)

In the first analysis of treatment effect, all visits in both the double-blind treatment period and open-label treatment period were included; the baseline was the 5-D Itch scale total score collected on Day 1 of the double-blind treatment period prior to randomisation (i.e, double-blind baseline). In the second analysis, only the visits in the open-label treatment period were included; the baseline was the last 5-D Itch scale total score collected in the double-blind treatment period (i.e., open-label baseline). The study was stopped early by the sponsor due to reasons unrelated to safety or lack of drug effect; only limited meaningful conclusions could be drawn from the small number of subjects that completed 52 weeks of treatment (n = 5). Thus, results at Week 36 are discussed (n = 52).

The mean (standard deviation [SD]) baseline 5-D Itch score was [redacted] At the first assessment post-baseline (after the double-blind Week 4, on Day 29), [redacted]

[redacted]

When subjects randomised to placebo during the double-blind treatment period transitioned to open-label treatment (between double-blind Week 12 and open-label Week 4), [redacted] at the end of open-label Week 4 for double-blind placebo (n = 200) and difelikefalin (n = 167) subjects, respectively) (**Error! Reference source not found.**,[redacted]). Over the course of long-term treatment through Week 36, the total number of subjects decreased at each time point, due to discontinuation (with two subjects discontinuing due to lack of therapeutic efficacy). The change from baseline was [redacted] at the end of open-label Week 36 for placebo/difelikefalin [n = 30] and difelikefalin/difelikefalin [n = 22] subjects, respectively) (XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX[redacted]).

[redacted]

[redacted]

[redacted]	[redacted]		[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

The percentage of subjects with an improvement of at least five points from the double-blind baseline [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

DB = double-blind; OL = Open-label; pbo = placebo. Note: difelikefalin is referred to as its previous name ‘CR845’ in this diagram.

Note: The number of subjects with non-missing values at each visit were displayed. Subjects with missing baseline values were excluded. Note: End of Week 52 results were excluded due to the small number of subjects with non-missing data. Note: Baseline was the last assessment prior to the start of double-blind treatment.

At the beginning of the open-label treatment period (end of open-label Week 4), the number of double-blind placebo subjects with an improvement of at least five points from the double-blind baseline [redacted] The percentage of double-blind difelikefalin subjects with an improvement of at least five points from the double-blind baseline [redacted]. Throughout the open-label treatment period, [redacted] treatment sequence groups, respectively.

The mean (SD) open-label baseline 5-D Itch score (i.e., the value from the final 5-D Itch scale total score collected in the double-blind treatment period) was [redacted] for subjects who were randomised to placebo and difelikefalin in the double-blind treatment period, respectively. [redacted]

The percentage of subjects with an improvement of at least five points from the open-label baseline was [redacted]

The same analyses performed for the total 5-D Itch were also performed for the individual 5-D Itch domains of degree, duration, direction, disability, and distribution.

[redacted]

Pooled analysis of KALM-1 and KALM-2

Pooled results on the efficacy from the KALM-1 and KALM-2 trials have been accepted for publication (39). The pooled analysis included 851 randomised patients, with 426 patients receiving difelikefalin and 425 patients receiving placebo. The primary endpoint, proportion of patients achieving ≥3-point improvement in WI-NRS score at Week 12, was achieved in 51.1% of participants in the difelikefalin group and 35.2% of participants in the placebo group. As early as Week 3, rates of complete response in WI-NRS were significantly greater in patients in the pooled population treated with difelikefalin versus placebo; this was maintained through Week 12. In the pooled population, significantly greater proportions of participants in the difelikefalin group achieved clinically meaningful improvements in itch-related QoL versus the placebo group, as measured by ≥15-point improvements in Skindex-10 total scores (55.5% vs 40.5%, respectively, at Week 12; $p < 0.001$) and ≥5-point improvements in 5-D Itch total scores (52.1% vs 42.3%, respectively, at Week 12; $p = 0.01$) over 12 weeks of treatment.

Following the 12-week placebo-controlled trial period, 340 patients from the difelikefalin arm and 372 patients from the placebo arm met the inclusion criteria and entered the OLE. In the OLE, all patients received difelikefalin. For patients who continued difelikefalin treatment, itch improvement (measured as mean 5-D Itch total score) was maintained through the 52-week OLE; additionally, itch improvement was consistently observed to emerge in patients who switched from placebo to difelikefalin during the OLE.

CLIN3105 (28)

Effectiveness conclusions

Treatment with difelikefalin 0.5 mcg/kg resulted in a clinically meaningful reduction in pruritus, as measured by the percentage of subjects with a ≥3-point improvement in WI-NRS score through Week 12. At Week 12, a majority of the subjects reported at least a 3-point (73.7%) or 4-point (59.3%) improvement from baseline in the weekly mean of the daily 24-hour WI-NRS score, which was previously established as a clinically meaningful threshold for this patient population (40). [redacted]

At Week 12, clinically meaningful improvements were also seen in sleep and global health status metrics. Patients completed the Sleep Quality Questionnaire at all three dialysis visits during the run-in period, the final week of the study (Week 12), the first dialysis visit in Week 1, and the first dialysis visit after the last dose of difelikefalin. In total, 66.0% of subjects achieved a 3-point or greater improvement, and 56.7% reported a ≥4-point improvement from baseline in sleep quality as measured by Sleep Quality score[redacted]. Improvements were reported in skin irritation measures of the EQ-PSO: the percentage of subjects reporting no problems in skin irritation increased from 1.4% at baseline to 28.9% in Week 12.

Treatment with difelikefalin also improved itch-related QoL in subjects with CKD-aP undergoing haemodialysis, as measured by the Skindex-10 scale (mean change ± SD = -21.0 ± 15.59, [redacted] and 5-D Itch scale total scores (mean change ± SD = -7.1 ± 4.27,[redacted] The reductions in Skindex-10 scale total score and 5-D Itch total score are both considered to be clinically meaningful (7, 37). In addition, for all three domains of the Skindex-10 scale (disease, mood/emotional distress, and social functioning) and all five domains of the 5-D Itch scale (disability, distribution, duration, degree, and direction), subjects treated with difelikefalin achieved a reduction in score at the end of Week 12.

In conclusion, the effectiveness results in this study, in which all patients knowingly receive active treatment and therefore provide insight into the expected real-world effectiveness, indicate that in subjects undergoing haemodialysis, treatment with difelikefalin reduced CKD-aP and improved itch-related QoL.

B.2.7 Subgroup analysis

All subgroup analyses were pre-planned.

KALM-1

The primary efficacy analysis was conducted separately for interim analysis and post-interim analysis subjects, and for stratification factors. These stratification factors were use of anti-itch medication or not at baseline, and presence or absence of certain medical conditions at baseline.

The difelikefalin group showed a greater percentage of subjects achieving a ≥3-point improvement from baseline in WI-NRS scores at Week 12 regardless of use of anti-itch medication, which was considered to be statistically significant in both cases: 48.8% vs 27.2%, p =0.001 for no use of anti-itch medications at baseline; 53.2% vs 29.4%, p =0.005 for use of anti-itch medication.

A descriptive analysis of the change in WI-NRS from baseline at Week 12 and by study site was also conducted, as well as an analysis of the proportion of subjects achieving a ≥3-point improvement.

Please see Appendix E: Subgroup analysis for results and detailed information for all subgroup analyses.

KALM-1 OLE

No subgroup analyses were conducted.

KALM-2

The primary efficacy endpoint was analysed separately for interim analysis and post-interim analysis subjects, and by stratification factor, study region, and dialysis type. Stratification factors were use or non-use of anti-itch medication at baseline, and presence of absence of specific medical conditions. A descriptive analysis of the change in WI-NRS from baseline at Week 12 and by study site was also conducted,

as well as an analysis of the proportion of subjects achieving a ≥3-point improvement.

The subjects using anti-itch medications at baseline had a greater treatment difference (odds ratio = 2.15; 95% CI, 1.09 to 4.25) favouring difelikefalin than subjects not using anti-itch medications at baseline (odds ratio = 1.36; 95% CI, 0.84 to 2.20).

For completeness, the Week 12 change in WI-NRS score from baseline was summarised using descriptive statistics by study site, along with the counts and proportions (out of the ITT population at that site) of subjects achieving a ≥3-point improvement from baseline by site (for sites that had at least two subjects in each treatment arm with data at Week 12). Please see Appendix E: Subgroup analysis for results and detailed information of all subgroup analyses.

KALM-2 OLE

No subgroup analyses were conducted.

CLIN3105

No subgroup analyses were conducted

B.2.8 Meta-analysis

Not applicable.

B.2.9 Indirect and mixed treatment comparisons

Not applicable.

B.2.10 Adverse reactions

Pooled results on the safety results from the KALM-1 and KALM-2 trials have been accepted for publication (42). Difelikefalin has a good safety profile and is well tolerated by patients. Although more than 60% of patients experienced an adverse event with difelikefalin in KALM-1 and KALM-2, the rate and type of AEs observed with difelikefalin treatment were comparable with those observed with placebo. This has been consistently demonstrated across Phase II and Phase III studies. Furthermore, some of the adverse events reported in the trials, including dizziness

and headache, have been reported to occur in more than half of all patients on HD (43).

Table 33 summarises the occurrence of TEAEs and deaths during the double-blind treatment periods of KALM-1 and KALM-2. In the KALM-1 study, the rate of patients experiencing at least one (TEAE) during the double-blind treatment period was 68.8% in the difelikefalin group, versus 62.2% in the placebo group. The findings of the KALM-2 study were consistent with those of KALM-1: 68.1% of patients receiving difelikefalin and 61.4% of those receiving placebo experienced at least one TEAE.

The rate of serious TEAEs was also comparable between difelikefalin and placebo, with 25.9% in the difelikefalin group and 21.8% in the placebo group experiencing at least one serious TEAE during the double-blind treatment period of KALM-1; the equivalent figures for KALM-2 were 24.7% and 21.6%. The number of deaths was very low, and consistent across the treatment arms of both studies. Both deaths in the difelikefalin group of KALM-1 were attributed to sepsis; the two deaths in the placebo group were due to septic shock. In KALM-1, TEAEs led to discontinuation in 4.8% of patients in the placebo group and 7.9% of patients in the difelikefalin group. A similar pattern was observed in KALM-2, confirming that difelikefalin has a favourable safety profile.

Table 33: Summary of TEAEs and deaths during the double-blind treatment periods of KALM-1 and KALM-2 (double-blind safety populations)

	KALM-1		**KALM-2**	
	Placebo (n=188)	**Difelikefalin (n=189)**	**Placebo (n=236)**	**Difelikefalin (n=235)**
Number of patients with at least 1 TEAE	117 (62.2%)	130 (68.8%)	145 (61.4%)	160 (68.1%)
Number of patients with at least one serious TEAE	41 (21.8%)	49 (25.9%)	51 (21.6%)	58 (24.7%)
Number of deaths	2 (1.1%)	2 (1.1%)	2 (0.8%)	2 (0.9%)
Number of patients with at least 1 TEAE resulting in study drug discontinuation	9 (4.8%)	15 (7.9%)	8 (3.4%)	13 (5.5%)

Abbreviation: TEAE., treatment-emergent adverse event

Safety data from the OLE of KALM-1 and KALM-2 also show comparable TEAE rates across the treatment groups (Table 34).

Table 34: Summary of TEAEs during the open-label treatment period of KALM-1 and KALM-2 (open-label safety population)

	KALM-1 OLE		KALM-2 OLE	
	Placebo/ difelikefalin (n=162)	**Difelikefalin/ difelikefalin (n=162)**	**Placebo/ difelikefalin (n=210)**	**Difelikefalin/ difelikefalin (n=162)**
Number of patients with at least 1 TEAE	132 (81.5%)	125 (82.8%)	117 (61.9%)	256 (64.2%)
Number of patients with at least one serious TEAE	88 (54.3%)	79 (52.3%)	61 (32.3%)	130 (32.6%)
Number of deaths	12 (7.4%)	10 (6.6%)	7 (3.7%)	15 (3.8%)
Number of patients with at least 1 TEAE resulting in study drug discontinuation	15 (9.3%)	10 (6.6%)	8 (4.2%)	20 (5.0%)

Abbreviation: TEAE, treatment-emergent adverse event

Similarly, in the pooled studies, patients reported TEAEs that were mostly mild to moderate in both the placebo-controlled period (difelikefalin: 57.5% [244/424] vs placebo: 52.6% [223/424]) and the OLE period (difelikefalin: 53.6% [427/796]). The incidence rate of common TEAEs and serious TEAEs did not increase with longer-term exposure.

It was demonstrated that 71.2% of patients in the difelikefalin group experienced TEAEs, versus 65.3% in the placebo group (Table 35). The rates of TEAEs leading to study drug discontinuation were low, and comparable between the placebo and difelikefalin groups: 4.0% and 6.8%, respectively. (42). Due to its positive safety profile, difelikefalin is appropriate for the long-term treatment of CKD-aP.

Table 35: Summary of TEAEs according to a pooled analysis of the KALM-1 and KALM-2 safety population

	Pooled analysis		
	Placebo-controlled Weeks 0-12		Placebo-controlled + OLE Weeks 0 up to 64
	Placebo (n=424)	Difelikefalin (n=424)	Difelikefalin (n=796*)
Number of subjects with any TEAE reported	277 (65.3%)	302 (71.2%)	640 (80.4%)
Number of subjects with any non-fatal serious TEAEs reported	96 (22.6%)	107 (25.2%)	354 (44.5%)
Number of subjects with any TEAE leading to death	5 (1.2%)	3 (0.7%)	37 (4.6%)
Number of subjects with any TEAE leading to study drug discontinuation	17 (4%)	29 (6.8%)	72 (9%)

Abbreviation: TEAE, treatment-emergent adverse event
* = Number of patients exposed to difelikefalin in either the placebo-controlled period or the OLE

n's are based on the safety population, defined during the double-blind period as randomised subjects who received at least one dose of double-blind study drug during the placebo-controlled period, and defined during the OLE period as subjects who received at least one dose of study drug during the placebo-controlled or OLE period.

Source: (44)

Table 36 presents the most commonly reported TEAEs in KALM-1 and KALM-2. In KALM-2, nausea and fall were experienced by ≥5% of patients; however, the rate at which these events occurred were comparable between the difelikefalin and placebo groups. The most commonly reported serious AEs were hyperkalaemia (2.1% in both groups), pneumonia (1.6% in the difelikefalin group and 2.7% in the placebo group), sepsis (1.6% in the difelikefalin group and 2.1% in the placebo group), hypotension (1.6% in the difelikefalin group and 1.1% in the placebo group), and chronic obstructive pulmonary disease (1.6% in the difelikefalin group and 0.5% in the placebo group).

The TEAEs observed in KALM-1 and KALM-2 are consistent with those observed in CLIN2101 and other studies in the difelikefalin study programme.

Table 36: TEAEs ≥5% of any treatment group (double-blind treatment period of KALM-1 and KALM-2)

TEAEs at ≥5% frequency	KALM-1		KALM-2	
	Placebo (n=188)	Difelikefalin (n=189)	Placebo (n=236)	Difelikefalin (n=235)
Diarrhoea	7 (3.7%)	18 (9.5%)	13 (5.5%)	19 (8.1%)
Dizziness	2 (1.1%)	13 (6.9%)	12 (5.1%)	13 (5.5%)
Vomiting	6 (3.2%)	10 (5.3%)	14 (5.9%)	15 (6.4%)
Nasopharyngitis*	10 (5.3%)	6 (3.2%)	N/A	N/A
Fall*	5 (2.7%)	5 (2.6%)	12 (5.1%)	16 (6.8%)
Nausea*	9 (4.8%)	6 (3.2%)	10 (4.2%)	15 (6.4%)

* = TEAEs that occurred in ≥5% of patients in only KALM-1 or KALM-2, but results from both studies are reported for consistency.
Abbreviation: TEAE, treatment-emergent adverse event

In the 2-week discontinuation period of KALM-1, subjects were evaluated for TEAEs potentially related to opioid withdrawal. The observed TEAE profile showed no suggestion of drug withdrawal following treatment cessation and no evidence of dependence development. Subjects were also evaluated for potential signs and symptoms of opioid withdrawal using the Short Opiate Withdrawal Scale (ShOWS) and Objective Opiate Withdrawal Scale (OOWS) during the 2-week discontinuation period. The results from both scales indicated no signs of withdrawal in either treatment group. The proportion of patients experiencing at least one TEAE during the double-blind discontinuation period in the difelikefalin group (19.9%) was comparable with that of the placebo group (24.6%), as presented in Table 37. No TEAEs occurred with a frequency of ≥5% during the discontinuation period. Table 37 also presents System Organ Classes with a TEAE frequency of ≥5% during the discontinuation period.

Table 37: TEAEs by System Organ Class (double-blind discontinuation safety population of KALM-1)

	Placebo (n=179)	Difelikefalin (n=176)
Number of patients with ≥1 TEAE	44 (24.6%)	35 (19.9%)
TEAE by System Organ Class (≥5% frequency in either treatment arm)*		
Gastrointestinal disorders	8 (4.5%)	9 (5.1%)
Infections and infestations	10 (5.6%)	2 (1.1%)

* = No specific TEAEs occurred at ≥5% frequency in either treatment arm. Data are shown as the overall rates for each system organ class. Abbreviation: TEAE, treatment-emergent adverse event

CLIN3105

Data from CLIN3105 support the safety and tolerability of difelikefalin reported in KALM-1 and KALM-2. Of 222 patients, 143 (64.4%) reported a total of 414 TEAEs over the course of the study. The most common TEAEs reported (≥4% of all patients) were diarrhoea (5.0%), nausea (4.5%), and hyperkalaemia (4.1%); these events were well tolerated. Overall, 6.3% of patients reported TEAEs that resulted in study drug discontinuation. Of the 143 who reported a TEAE, [redacted] 91 serious TEAEs in 45 patients (20.3%) were reported during the study, none of which were considered related to difelikefalin.

B.2.11 Ongoing studies

There are no ongoing studies that will provide additional evidence in the next 12 months for the indication being appraised.

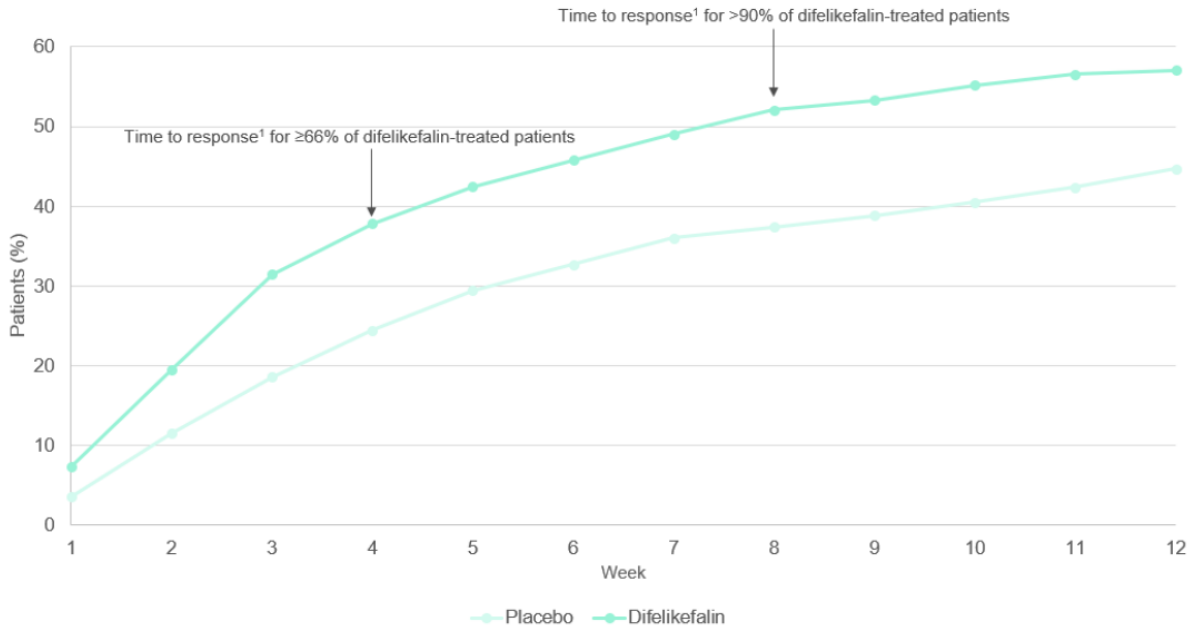
B.2.12 Interpretation of clinical effectiveness and safety evidence

Summary of clinical efficacy

Difelikefalin reduces itch intensity

Difelikefalin clearly displays an ability to reduce itch intensity compared to placebo. Two-thirds of patients who were treated with difelikefalin and reported a response to treatment had reached a ≥3 point reduction on the WI-NRS within 4 weeks, as assessed in a pooled post-hoc analysis of KALM-1 and KALM-2. After four additional weeks of treatment, >90% of these patients had achieved such a reduction (45) (see **Error! Reference source not found.**). 38.7% (95% CI [32.8%, 45.0%]) of patients treated with difelikefalin had a had a ≥4-point improvement in their WI-NRS score (representing a ‘substantial improvement in itching intensity’) from baseline to Week 12; only 23.4% (95% CI [18.7%, 28.8%]) of patients receiving placebo experienced this level of improvement ($p<0.001$). In CLIN3105, an even higher proportion of patients – 59.3% – reported a substantial improvement in itch intensity.

Figure 8 Time to first improvement for maintenance HD patients with moderate-to-severe CKD-aP



1 Time to first improvement in patients who reported ≥3-point improvement in WI-NRS from baseline at any point during the trial

A complete responder was defined as a subject with ≥80% of the non-missing 24-hour WI-NRS scores equal to 0 or 1 on Week 12. In a pooled analysis of KALM-1 and KALM-2, it was shown that there was a significantly better complete responder rate with difelikefalin vs placebo (12.0% vs 6.7%, p=0.006). The odds ratio of this effect was 2.11 (95% CI: 1.32 to 3.39), meaning that patients receiving difelikefalin are more than twice as likely to have very little to no itching at the EOT compared with those receiving placebo.

Difelikefalin maintains its clinical effectiveness

Difelikefalin addresses the chronic symptoms of CKD-aP. Initial improvements in itch intensity with difelikefalin vs placebo are observed in the early weeks of treatment and are followed by sustained and clinically meaningful improvements in itch intensity and QoL throughout the treatment period (up to 64 weeks following patients from KALM-1 double-blind treatment period through to the OLE period).

Difelikefalin improves sleep quality

CLIN3105 demonstrated that difelikefalin improved patients' sleep quality by reducing the impact of their itch on sleep. [redacted] In total, 66.0% of patients achieved a ≥3-point improvement and 56.7% a ≥4-point improvement from baseline in the weekly mean of the Sleep Quality Questionnaire score.

Difelikefalin improves HRQoL

Studies by (6) and (15) have found a statistically significant association between itch intensity and both physical and mental HRQoL. Difelikefalin effectively reduces itch intensity, and as such will have a positive impact on CKD-aP patients HRQoL.

In addition, the impact of difelikefalin on HRQoL was measured using the 5-D Itch scale, which measures the effect of itching on a patient's life across five different dimensions. [redacted]

Summary of clinical safety

Difelikefalin has a good safety profile and is well tolerated by patients. Although more than 60% of patients experienced an adverse event with difelikefalin in KALM-1 and KALM-2, the rate and type of AEs observed with difelikefalin treatment were comparable with those observed with placebo. Pooled analysis of KALM-1 and KALM-2 showed that 71.2% of patients in the difelikefalin group experienced TEAEs, vs 65.3% in the placebo group. The rates of TEAEs leading to study drug discontinuation were low (<10%), and comparable between the placebo and difelikefalin groups (4.0% vs 6.8%) (42).

Data from CLIN3105 support the safety and tolerability of difelikefalin reported in KALM-1 and KALM-2. Of 222 patients, 143 (64.4%) reported a total of 414 TEAEs over the course of the study. The most common TEAEs reported (≥4% of all patients) were diarrhoea (5.0%), nausea (4.5%), and hyperkalaemia (4.1%). These events were well tolerated; overall, 6.3% of patients reported TEAEs that resulted in study drug discontinuation. Of the 143 who reported a TEAE, 68 patients (30.6%) had a maximum severity of mild, 56 (25.2%) had a maximum severity of moderate, and 19 (8.6%) had a maximum severity of severe. TEAEs of special interest (including gait disturbance, falls, dizziness, somnolence, seizures, syncope, mental status changes, mood changes, unusual feeling/sensation, tachycardia, and palpitations) were reported by 10.4% of patients. Three treatment-emergent deaths (1.4% of patients) and 91 serious TEAEs in 45 patients (20.3%) were reported during the study, none of which were considered related to difelikefalin.

Difelikefalin is safe and well tolerated by patients. Unlike centrally acting mu opioid receptor agonists, there is no evidence of physical dependence, abuse, or addiction potential. The frequency of AEs and SAEs with difelikefalin was similar to that of placebo, demonstrating a positive risk-benefit profile.

Strength and limitations of the clinical evidence base

Th efficacy of difelikefalin in treatment of moderate-to-severe pruritus associated with CKD in adult patients receiving in-centre haemodialysis has been demonstrated in two Phase 3, randomised, double-blind, placebo-controlled studies (KALM-1 and KALM-2). These studies provide robust data with a comparator placebo arm to assess the efficacy of difelikefalin, with minimal risk of bias. Adequate concealment of treatment allocation was achieved using an interactive voice/web response system. Successful blinding was also achieved: during the double-blind treatment period patients, investigators, study staff, and the sponsor were blinded to the study drug assignment, only breaking the blind in cases of medical emergency. It should be noted that a placebo effect was seen within both KALM-1 and KALM-2, with 27.6% and 42.2% of placebo patients achieving at least a 3-point improvement from baseline in the WI-NRS by Week 12 in KALM-1 and KALM-2, respectively. However, itch is by nature often subject to a placebo effect, as it is a complex and subjective symptom that can be influenced or exacerbated by both environmental and

psychological factors (46). For example, itch can be highly susceptible to suggestion: studies have shown that verbal suggestion used to influence patient expectation, i.e., telling a patient they are receiving an antipruritic treatment, can lead to an increased placebo response (47-49). Furthermore, the odds ratio for a ≥3-point improvement from baseline with difelikefalin versus placebo was 2.72 (95% CI, 1.72 to 4.30) in KALM-1 and 1.61 (95% CI, 1.08 to 2.41) in KALM-2, both of which are statistically significant, ($p<.001$ and $p=0.20$, respectively), demonstrating a significant treatment effect compared to placebo.

The efficacy of difelikefalin has also been shown in uncontrolled, open-label, single-arm studies (KALM-1 OLE, KALM-2 OLE, and CLIN3105). Non-randomised study designs are associated with an inherent risk of bias, for example selection bias, reporting bias, and incomplete follow-up. A limitation of KALM-2 specifically is that, due to an administrative decision by the sponsor (unrelated to efficacy or safety), KALM-2 was halted early; therefore, 313 subjects (78.4%) could not complete the 52-week open-label treatment period. The benefit of difelikefalin has nevertheless been consistently demonstrated across these studies. In addition, in a CKD-aP modified involving eight nephrologists across the UK, conducted in May 2022, 7 (87.5%) agreed that the patient populations in the KALM trials is broadly generalisable to the UK patient population. Please see Appendix N for full report.

Conclusion

Reducing itch intensity is the cornerstone of CKD-aP treatment because itching drives the wider burden of the condition (e.g., poor QoL, poor sleep, infection, hospitalisations, mortality). Difelikefalin can effectively address the high unmet medical need in CKD-aP patients on HD, with a positive risk-benefit profile. The value of difelikefalin is demonstrated through high-grade clinical evidence from a robust clinical trial programme including two randomised, double-blind, placebo-controlled, Phase III trials conducted in geographically diverse populations. This distinguishes difelikefalin from other treatments used today, which are used off-label and supported by low-grade, weak clinical evidence.

In clinical trials, the majority (>50%) of patients achieved a ≥3-point reduction in itch intensity (WI-NRS) after 12 weeks of difelikefalin treatment, which was independent of concomitant antipruritic medication use. This clinically meaningful change equates

to a reduction in itch intensity of one degree, such as from severe to moderate or from moderate to mild in most patients. Compared with placebo, difelikefalin doubles the chance of patients having very little to no itch at the EOT. [redacted]

Difelikefalin is the only treatment with high-quality, well documented evidence supporting its efficacy and safety and licensed for the treatment of CKD-aP.

B.3 Cost-effectiveness

Summary of cost-effectiveness analysis

- A *de novo* economic model has been developed to assess the cost-effectiveness of difelikefalin compared with established clinical management for treatment of adults with moderate-to-severe CKD-aP.
- The model is a Markov model and comprises 5 core health states as defined by level of itch severity: none, mild, moderate, severe, and very severe.
- The analysis is consistent with the NICE reference case; costs and benefits are discounted at a rate of 3.5.% and a lifetime time horizon is adopted.
- The 5-D Itch scale is a multidimensional questionnaire which assesses itch severity and itch-related quality of life over the previous 2 weeks. The scale has been validated in patients with chronic pruritus, including haemodialysis patients, and has been shown to be sensitive to changes in pruritus over time (37).
- The 5-D Itch scale was completed by trial participants throughout the duration of the double-blind 12-week period, and the open-label extension phase for both the KALM-1 and KALM-2 trials.
- Treatment-specific transition probabilities between CKD-aP severity categories were derived from the pooled KALM-1 and KALM-2 trial data, using 4-weekly (run-in phase) and 52-weekly (long-term) transition count data. Transition matrices are derived from per-cycle probabilities of losing or gaining health states. Each cycle has unique transition probabilities, as the response to treatment is greatest following initiation of treatment and overall response is further stratified by baseline CKD-aP severity.
- In the base case, the mean change in itch score from baseline is assumed to remain unchanged for the established clinical management arm from week 12 to week 64. Furthermore, in line with the clinical opinion that the placebo affect would wane over time, a waning effect is applied in the established clinical management arm equal to a 5% probability for patients to gain a health state (deteriorate) each year following Week 64 (cycle 4 onwards).

- Treatment with difelikefalin with established clinical management compared with established clinical management alone was associated with increased life years (0.06 per person) and increased QALYs ([redacted] per person) at an incremental cost of [redacted] per person at patient access scheme (PAS) price. As a result, difelikefalin with established clinical management is cost-effective compared with established clinical management alone, with an incremental cost-effectiveness ratio (ICER) of £24,293/QALY gained.
- Extensive scenario analyses demonstrate the base case cost-effectiveness results to be robust to variation in model inputs and assumptions, with only 2 of the 10 scenarios did the ICER exceed a willingness to pay (WTP) threshold of £30,000/QALY. Subgroup analysis demonstrates that in patients only with severe or very severe itch at baseline, the cost-effectiveness of difelikefalin improves and falls below a WTP threshold of £20,000/QALY. The ICER in the probabilistic analysis remains cost-effective at PAS price with an ICER of £23,253.

B.3.1 Published cost-effectiveness studies

An SLR was undertaken in April-July 2022 to identify cost-effectiveness studies, treatment pathway guidelines, cost and resource use data, and HRQoL data for CKD-aP, with particular focus given to the UK and Europe. Full details of the SLR search strategy, study selection process, and results are presented in Appendix G. The review identified 16 studies detailing treatment guidelines, 3 studies providing utility evidence, 7 studies providing cost and resource use data, and 1 economic evaluation. Due to the lack of economic evaluation evidence for CKD-aP, disease criteria were extended to include CKD and pruritus independently. The expanded economic review identified 6 NICE health technology assessments (HTA), and 12 studies reporting on cost-effectiveness analyses in CKD and pruritus. The results of the expanded SLR provided insight and guidance on the model development and structure, although only one cost-effectiveness analysis identified in the primary SLR relating to CKD-aP was considered directly relevant.

Soro et al. (2022) present a methodological approach to assess the economic value of difelikefalin for the treatment of CKD-aP. The study outlines a cohort model with 4 health states representing different levels of pruritus and presents the results of a 5-D itch scale to EQ-5D mapping study. A separate model was commissioned by Vifor and has been developed to address the decision problem of the current appraisal. No relevant published cost-effectiveness analyses in CKD-aP for comparator technologies were identified.

B.3.2 Economic analysis

A *de novo* economic model was developed to assess the cost-effectiveness of difelikefalin compared with established clinical management for treatment of adults with moderate-to-severe CKD-aP. The key features of the economic analysis and their justifications are presented in Table 38.

Table 38: Key features of the economic analysis

Factor	Chosen values	Justification
Model structure	Markov model with 5 core health states reflecting CKD-aP severity	Cohort Markov models have been used in previous CKD and pruritus appraisals. The mutually exclusive health states appropriately capture the heterogeneity of HRQoL and healthcare costs incurred in different CKD-aP severity states.
Time horizon	Lifetime	NICE reference case (50); considered to reflect that CKD-aP is chronic and expected to continue for the duration of patients' lifetime.
Comparator	Established clinical management	NICE final scope; considered because there are currently no approved treatments for CKD-aP. and no single treatment is used consistently across England.
Source of utilities	Primary data collection mapping study	As no generic preference-based measures of health were collected in the primary trials, a separate primary data collection study of UK dialysis centres was undertaken to develop a mapping algorithm related to the appropriate outcome measures.
Source of costs	NHS and personal social services (PSS) perspective; sourced from national databases including British National Formulary (BNF), National Cost Collection, and Personal Social Services Research Unit (PSSRU)	NICE reference case (50)
Health effects measure	QALYs	NICE reference case (50)

Factor	Chosen values	Justification
Half-cycle correction	Not applied in 4-week cycles. Applied in yearly cycles.	NICE reference case (50); half-cycle correction not applied in the run-in period due to short cycle length (4-week cycle).
Abbreviations: NHS: National Health Service; NICE: National Institute for Health And Care Excellence; CKD-aP: chronic kidney disease-associated pruritus; QALY: quality-adjusted life year; PSS: personal social services' PSSRU: Personal Social Services Research Unit		

B.3.2.1 Perspective

In accordance with current NICE guidance (50), a cost-utility analysis considering lifetime quality-adjusted life years (QALYs) and costs from a current UK NHS and personal social services perspective was undertaken. Both costs and QALYs were discounted at a rate of 3.5% per year.

B.3.2.2 Intervention and comparator

The proposed intervention is difelikefalin administered by intravenous bolus injection at the end of haemodialysis treatment. Difelikefalin is restricted for in-centre haemodialysis use only. It is proposed that difelikefalin be used as an adjunct to established clinical management where established clinical management is insufficient in reducing pruritus.

There are currently no approved treatments for CKD-aP apart from difelikefalin; treatment instead focuses on symptom management. BAD recommends only capsaicin cream, topical calcipotriol, or oral gabapentin and advises against sedative antihistamines and cetirizine (21). In the modified Delphi panel, it was noted that difelikefalin may be used in adjunct with topical creams and prior to the use of gabapentin (Appendix N: Clinical opinion and consensus report). The KALM trials did not directly include any comparator treatments, although patients using anti-itch medication at baseline were allowed to continue doing so.

B.3.2.3 Patient population

In accordance with the NICE final scope and the licensed indication for difelikefalin, the analysis considers adult patients with moderate-to-severe CKD-aP who are on haemodialysis.

The base case analysis considers the enrolled populations of the KALM-1 and KALM-2 trials. The efficacy results, as presented in B.2.6 Clinical effectiveness results of the relevant studies, demonstrate improved efficacy compared to current established clinical management, with the achievement of the primary endpoint (≥3-point improvement in weekly WI-NRS score) evaluated in subgroups based on baseline characteristics of Phase 3 trials. Subgroup analyses by use of anti-itch medication at baseline are presented in section B.3.11 Subgroup analysis.

The starting cohort age, proportion by sex, weight, and length of time on haemodialysis are used as inputs in the model to account for variations in costs and health outcomes due to demographic factors. The baseline characteristics applied in the model are based on the KALM trial populations. Data reported by the UK Renal Registry (UKRR) has also been included in the model as a secondary option (51). The UKRR collects and reports data annually on approximately 70,000 kidney patients on renal replacement therapy (RRT) in the UK. The baseline characteristics used in the model are summarised in Table 39.

Table 39: Baseline characteristics applied in the model

Characteristic	Pooled KALM trials (42)	UKRR
	Mean (SD)	Median
Starting cohort age (years)	58.3 (12.8)	67.50
Proportion male (%)	58.7	62.10
Weight (kg)	84.4 (21.5)	n/a
Length of time on dialysis (years)	4.78 (4.3) *	3.2 (52)
Note: * estimated from pooled KALM trial patient-level data set		

B.3.2.4 Model structure

A Markov model was constructed to calculate lifetime costs and QALYs for treatment with difelikefalin compared with established clinical management. In a Markov model, a set of mutually exclusive health states are defined which describe what can happen to the population of interest over time. People in the model can only exist in one of these health states at a time. Possible transitions are defined between each of the health states, and the probability of each transition occurring within a defined

period of time (a cycle) is assigned to each possible transition. This approach is deemed appropriate as it is consistent with the 3 appraisals for atopic dermatitis and 2 appraisals for CKD identified in the extended SLR.

The model comprises 5 core health states as defined by level of itch severity: none, mild, moderate, severe, and very severe. For each of the 5 core health states it possible to transition to either transplant or death (absorbing state). Renal transplant is assumed to be a definitive treatment for CKD-aP (21). Consequently, in the model, all patients discontinue CKD-aP treatment on receipt of a transplant. As presented in Table 40, the 5 core health states were defined in line with the outcome measures collected in the KALM-1 and KALM-2 trials. NRS categories were informed by (53) who explored optimal cut-offs for the 5-D Itch scale based on NRS categories in haemodialysis patients.

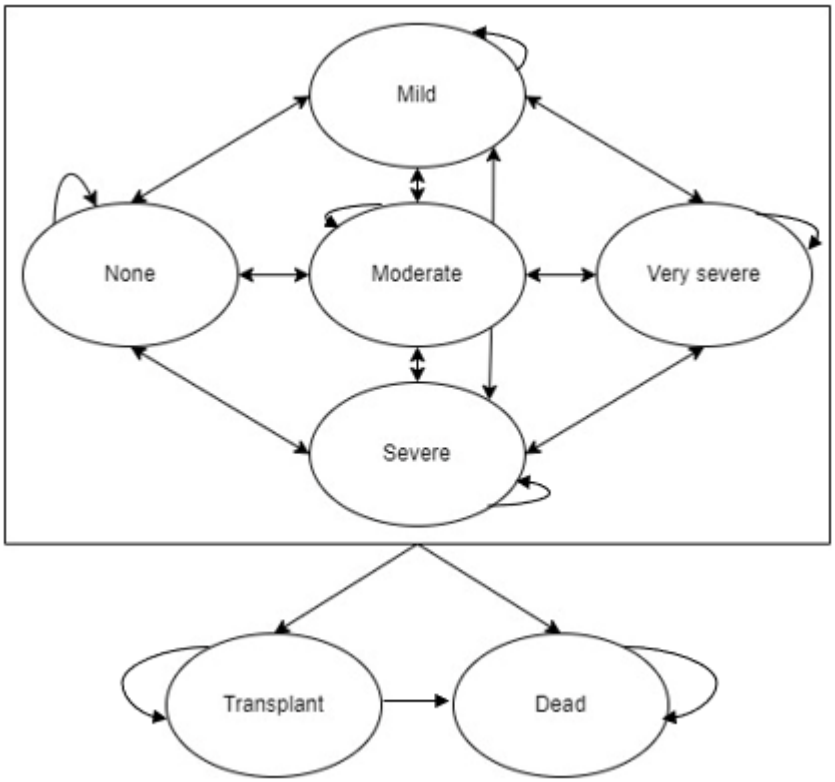
Table 40: Health state definitions by outcome measure scores

Health state	WI-NRS score	5-D Itch Scale (total score)
None	0	5-8
Mild	1-3	9-11
Moderate	4-6	12-17
Severe	7-8	18-21
Very severe	9-10	22-25
Abbreviations: WI-NRS: Worst Itch Numeric Rating Scale		

Error! Reference source not found. illustrates the model structure and the possible transitions between health states at each cycle. The model structure is designed to reflect current UK clinical practice for CKD-aP. It comprises an initial 'run-in' period which reflects short-term treatment decisions and initial response to treatment. This is followed by the long-term course of CKD-aP, whereby patients responding to treatment with difelikefalin remain on treatment for the duration of the model, and non-responders taken off treatment with difelikefalin remain on established clinical management. A 4-week cycle length is used for the first 3 cycles (the 'run-in' period) with a 52-week cycle length used from Cycle 4 onwards, continuing in the model for a lifetime (100 years, lifetime time horizon). A half-cycle correction is applied in the

model after the ‘run-in’ period for all 52-week cycles (Cycle 4 onwards), using the Trapezoidal method (54).

Figure 9: Model schematic



Treatment-dependent rates describe patient transitions between the 5 core health states. Time-dependent rates define how quickly people move from any of the 5 core health states to the transplant state. Time- and state-dependent rates define how quickly people move from any of the 5 core health states to the dead state. Transplant failure is not modelled, as no difference is anticipated between treatment arms following discontinuation of treatment for CKD-aP. This is consistent with previous CKD appraisals, including the cost-effectiveness analysis informing the NICE guidance for RRT and conservative management [NG107] (55).

In each model cycle, people accrue costs and QALY benefits associated with the relevant health state and treatment arm. In the base case, the model estimates total lifetime costs and QALYs for each treatment arm, with the summary measure presented as an ICER.

The model is constructed to allow people to enter the model at any state. In the base case analysis, only patients with moderate-to-severe (including very severe) CKD-aP are considered. A subgroup analysis is presented considering only patients with severe or very severe CKD-aP at baseline. The distribution of patients at model entry is based on the pooled data from the KALM-1 and KALM-2 trials, and by the outcome measure selected. Table 41 presents the possible distribution of patients across the 5 core health states at model entry.

Table 41: Distribution of patients at model entry

Outcome and itch severity at baseline	Proportion of patients in state at model entry (%)				
	None	Mild	Moderate	Severe	Very severe
5-D Itch Scale (total score)					
Moderate and severe	0.00%	0.00%	55.28%	34.17%	10.55%
Severe only	0.00%	0.00%	0.00%	76.40%	23.60%

B.3.3 Clinical parameters and variables

B.3.3.1 Efficacy

The data informing estimates of treatment efficacy in the model has been derived using patient-level data from the pooled KALM-1 and KALM-2 trials.

B.3.3.1.1 Measures of itch severity

As outlined in Table 40, the 5 core model health states reflecting the severity of CKD-aP are defined using the clinical outcome measures used in the KALM-1 and KALM-2 trials. The primary outcome measure used in both trials to assess itch intensity was the Worst Itching Intensity Numerical Rating Scale (WI-NRS), with the primary endpoint being the percentage of patients achieving a ≥3-point improvement at week 12 in weekly mean of daily WI-NRS scores. Intensity of itch is measured using a 0 to 10 numeric rating scale to indicate the intensity of the worst itching over the past 24 hours, where "0" represents "no itching" and "10" represents "worst itching imaginable". The WI-NRS has been widely used for evaluation of chronic itch,

including CKD-aP (7, 16, 56, 57). Anchor- and distribution-based analysis of the Phase 2 study CR845-CLIN2101 dataset supported the idea that a reduction of ≥3-points on the WI-NRS defines a clinically meaningful change threshold in pruritus in patient with CKD-aP undergoing haemodialysis (40). WI-NRS scores were collected throughout the duration of the double-blind 12-week period for both the KALM-1 and KALM-2 trials.

The secondary outcome measure collected in the KALM-1 and KALM-2 trials to assess itch severity and itch-related quality of life was the 5-D Itch scale. The 5-D Itch scale is a multidimensional questionnaire which assesses itch severity and itch-related quality of life over the previous 2 weeks. The questionnaire covers 5 dimensions of itch, including the degree, duration of itch/day, direction (improvement/worsening), disability (impact on activities such as work), and body distribution of itch. The total 5-D Itch scale score ranges from 5 to 25, with higher scores indicating worse responses. The scale has been validated in patients with chronic pruritus, including haemodialysis patients, and has been shown to be sensitive to changes in pruritus over time (37). Analysis of Phase 2 study CR845-CLIN2101 dataset showed that a 5-point reduction in the total 5-D Itch score from baseline represented a clinically meaningful improvement in patients with CKD-aP undergoing haemodialysis. The 5-D Itch scale was completed by trial participants throughout the duration of the double-blind 12-week period, and the open-label extension phase for both the KALM-1 and KALM-2 trials.

As 5-D Itch scale total scores provide estimates of treatment efficacy for up to 64-weeks compared with only 12-weeks using WI-NRS, they were used to inform efficacy estimates within the model base case.

B.3.3.1.2 Run-in phase and stopping rule

As noted in Section B.3.2.4 Model structure, the model comprises an initial ‘run-in’ period to reflect short-term treatment decisions and initial response to treatment. In the modified Delphi panel, clinical experts noted that they would consider stopping treatment if side effects are worse than the itch, or if the treatment is not working (Appendix N: Clinical opinion and consensus report).

As previously noted, the clinically meaningful thresholds for an improvement in itch are:

- a reduction of ≥3-points from baseline in the WI-NRS score
- a ≥5-point reduction from baseline in the total 5-D Itch scale score

A stopping rule has been implemented in the analysis whereby patients on difelikefalin not achieving a clinically significant endpoint will discontinue treatment. In the base case, the stopping rule has been applied at 12-weeks. In the modified Delphi panel, this endpoint was considered suitable as it aligns with the 3-monthly patient review conducted by the consultant nephrologists. Table 42 presents the relevant proportional split applied in the model at Week 12 in the base case analysis. The proportion of patients who remain on treatment following the application of the stopping rule was estimated as the count of patients in the health state at Week 12 who achieved a clinically meaningful itch score improvement divided by the total count of patients in the health state at Week 12. A scenario analysis is presented where the stopping rule is instead applied at Week 8.

Table 42: Stopping rule split applied in model base case

Difelikefalin arm	None	Mild	Moderate	Severe	Very severe	Total
Count at baseline; all patients	0	0	224	129	40	**393**
Count at Week 12; all patients	75	126	161	20	11	**393**
Count at Week 12; patients achieving clinically meaningful threshold	69	88	54	4	0	**215**
Proportion who remain on treatment after Week 12	92.00%	69.84%	33.54%	20.00%	0.00%	**54.71%**

B.3.3.1.3 Long-term extrapolation – difelikefalin treatment arm

Long-term data for the difelikefalin treatment arm in the model is informed by data collected during the OLE phase of the KALM-1 and KALM-2 trial. Data informing efficacy estimates has been derived from both patients who received difelikefalin and

patients who received placebo during the double-blind treatment period. **Error! Reference source not found.** presents results for the mean improvement in 5-D Itch scale total score from baseline across the double-blind treatment period and OLE phase, with and without the stopping rule applied.

[redacted]

As no data was collected beyond the 52-week OLE phase, in the base case, efficacy remains unchanged after Week 64 (Cycle 4). No treatment waning is anticipated; however, this has been explored in scenario analyses.

B.3.3.1.4 Long-term extrapolation – established clinical management arm

For both the KALM-1 and KALM-2 trials, all study participants in the OLE phase received difelikefalin. As no data informing the long-term efficacy for patients receiving placebo within the trials were available, 3 possible extrapolation methods were considered in the analysis: mean difference (MD), ratio of means (RoM), and no change in efficacy.

The MD (or 'difference in means') approach is a standard statistic that measures the absolute difference between the mean value in two groups in a clinical trial. For this analysis, the average of the mean difference between the placebo arm and the difelikefalin arm at each observation has been used to estimate a mean change in 5-D Itch scale total score from baseline to Week 64. This mean change value is then added to the baseline score of patients treated with placebo to simulate a 5-D Itch scale total score for Week 64.

The RoM method is an alternative to the MD approach, in which the average of the ratio of means between the placebo arm and the difelikefalin arm at each observation is multiplied by the baseline score of patients treated with placebo. This simulates a 5-D Itch scale total score for Week 64.

The company sought clinical opinion on the natural progression of CKD-aP and the potential trend in the mean change in itch score that could be expected in the extrapolation period for patients receiving placebo in the KALM trials (Appendix N: Clinical opinion and consensus report). It was noted that the placebo effect would wane over time in line with the natural progression of the disease, which is likely to get worse over time. Soro et al., 2022 graphically present data on changes in pruritus severity from the SHAREHD stepped wedge cluster randomised controlled trial, in which data were collected on 17 POS-S renal symptoms (including pruritus). Over the 18 months, a general trend was observed, with the prevalence of moderate pruritus remaining stable, mild/none increasing, and severe/overwhelming decreasing. Given that no robust quantitative data on the natural progression of CKD-aP were available, the third extrapolation approach considers that efficacy (mean change in itch from baseline) remains unchanged from Week 12 onwards.

Error! Reference source not found. plots the mean improvement in 5-D Itch scale total score from baseline across the double-blind treatment period and OLE phase for the placebo arm of the pooled KALM-1 and KALM-2 trial data, for all 3 possible extrapolation approaches. In the base case, the mean change in itch score from baseline is assumed to remain unchanged for the established clinical management arm from Week 12 to Week 64. Furthermore, in line with the clinical opinion that the placebo affect would wane over time, a waning effect is applied in the established clinical management arm equal to a 5% probability for patients to gain a health state (deteriorate) each year following Week 64 (Cycle 4 onwards).

[redacted]

B.3.3.1.5 Transition matrices

Treatment-specific transition probabilities between CKD-aP severity categories were derived from the pooled KALM-1 and KALM-2 trial data, using 4-weekly (run-in phase) and 52-weekly (long-term) transition count data.

Multiple imputation was used to fill in missing data values in the patient-level data set for the total 5-D Itch scale scores. This was carried out in R. Multiple imputation is based on the assumption that the data is missing completely at random, which was verified via a Missing Completely at Random test in the Misty package (58). The MICE package was used to perform multiple imputation with the Predictive Mean Matching approach; the number of imputations and maximum iteration were set to 5 and 40, respectively (59). The missing values were estimated based on treatment group, baseline itch score, and patient characteristics including age band, sex, diabetes status, length of ESRD, length of haemodialysis, length of CKD-aP, and use of anti-itch medication at baseline.

Table 43 and Table 44 summarise the number of observations included at baseline and the number of observations included within each cycle when using observed data, with and without missing data imputation. The number of observations in the difelikefalin model arm for cycles 1 to 3 (393 observations) reflects the population that received treatment with difelikefalin in the DB treatment period of the KALM-1 and KALM-2 trials, had a baseline observation, and had at least moderate CKD-aP

at baseline. The number of observations in the difelikefalin model arm for Cycle 4 onwards (279 observations) reflects the population included in previous cycles that achieved a clinically meaningful treatment response at Week 12 and that entered the OLE period of the KALM-1 and KALM-2 trials, plus the population that received placebo in the DB period who were eligible to enter the OLE period and achieved a clinically meaningful treatment response at Week 12 of the OLE period.

Table 43: Number of data observations included in analysis from KALM-1 and KALM-2

5-D Itch Scale total scores	Difelikefalin	Placebo
Total number of patient observations for KALM-1 and KALM-2	424	424
Patients with a missing baseline score	-7	0
Patients with 'None' and 'Mild' CKD-aP at baseline	-24	-21
Number of observations at baseline used in analysis	**393**	**403**

Table 44: Number of data observations included in the analysis at each model cycle from KALM-1 and KALM-2

5-D Itch Scale total scores	Difelikefalin		Placebo	
Cycle	Observed only	Missing data imputation	Observed only	Missing data imputation
Baseline count	393	393	403	403
Cycle 1 (baseline to Week 4)	356	393	371	403
Cycle 2 (Week 4 to Week 8)	333	393	357	403
Cycle 3 (Week 8 to Week 12)	330	393	359	403
Cycle 4 (Week 12 to Week 64)	74	279	N/A	N/A

Creating a matrix which calculates the probability of moving from any one state to each of the other states can result in small observation numbers estimating a single probability value, which may lead to unrealistic outcomes. Furthermore, because an extrapolation of the trial data was required to estimate the long-term efficacy for patients receiving placebo (Cycle 4 onwards), using estimates of a mean change in itch score from baseline would result in all placebo patients remaining as mild or moderate CKD-aP. To circumvent this issue, it is assumed that the probability of improving or deteriorating CKD-aP in each cycle is equal regardless of current health state. For example, the probability of remaining in the severe health state is the same as the probability of remaining in the moderate health state. As such, transition matrices are derived from per-cycle probabilities of losing or gaining health states. Each cycle has unique transition probabilities, as the response to treatment is greatest following initiation of treatment and overall response is further stratified by baseline CKD-aP severity. (**Error! Reference source not found.**).

To take account of the fact that patients do not experience drastic changes to their itch state, and additionally to reduce reliance on extreme values, probabilities were generated for a maximum improvement or deterioration by 3 health states. This assumption was examined: it was observed that only 3 patients in the difelikefalin treatment arm and 1 patient in the placebo arm improved by 4 health states across

the duration of both the KALM-1 and KALM-2 trials. Furthermore, no patients deteriorated by 4 health states in either treatment arm.

In the base case analysis, transition probability matrices were estimated from a simulated data set using the mean change from baseline in itch scores by CKD-aP severity at baseline: moderate, severe or very severe. In addition, the analysis used count data to inform the distribution of patients across 5-D Itch scale total scores at baseline.

Table 45 presents the mean change values used in the modelled base case. Mean change values for the placebo arm of the trial data assume no change in score following Week 12.

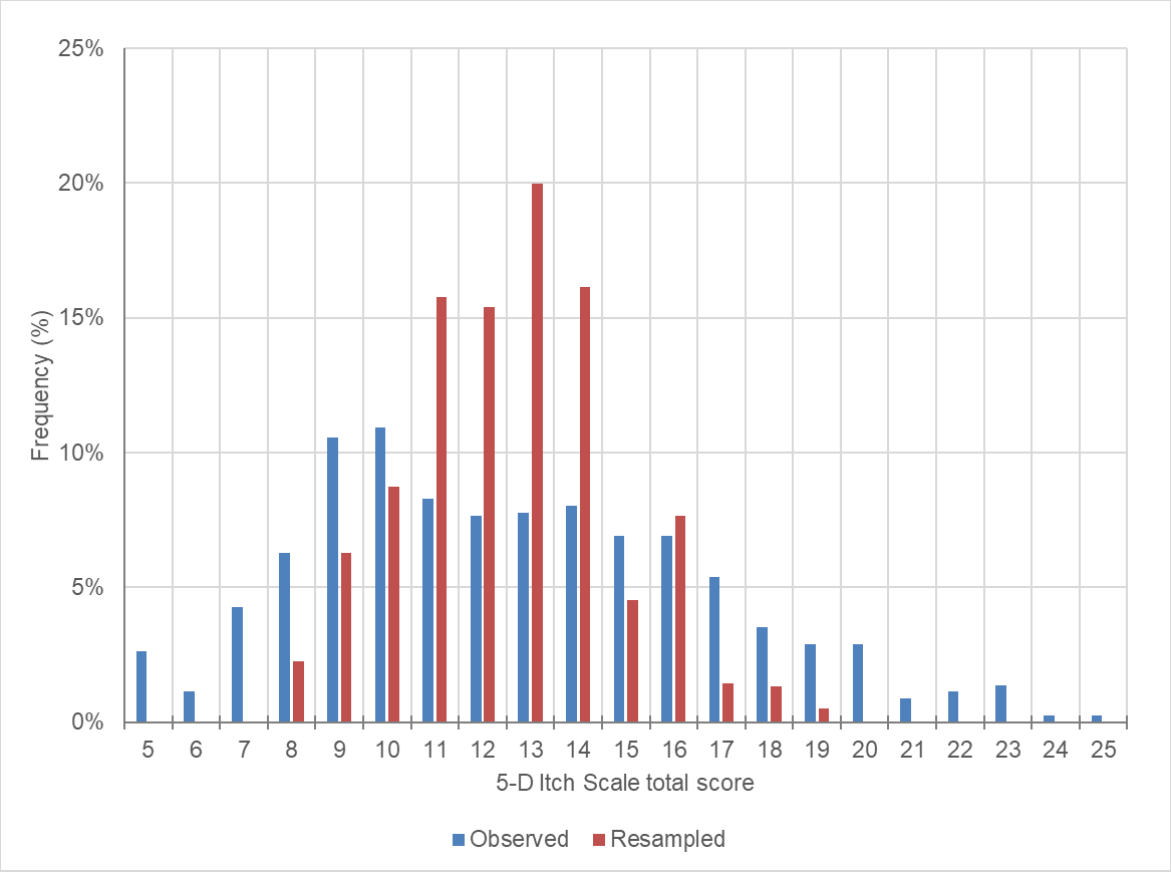
Table 45: Mean change in 5-D Itch scale total score from baseline

Mean change from baseline (5-D Itch scale total scores)	Week 4	Week 8	Week 12	Week 64
Difelikefalin treatment arm				
Moderate (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Very severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Established clinical management arm				
Moderate (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Very severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SE; standard error
Note: This table corresponds with the curves presented in **Error! Reference source not found.**

A histogram plot for the frequency of observations for each 5-D Itch scale total score at Week 12 using the observational data set and the simulated data set is shown in **Error! Reference source not found.**. Using the simulated data set in the base case approach was the preferred option, as it offers a better reflection of the underlying trend in the data and a more appropriate quantification of the uncertainty in the mean change in itch scores through probabilistic analysis.

Figure 10: Frequency of 5-D Itch scale total score observations for the observed data set and simulated data set at Week 12 for all patients



The base case transition probability matrices for the difelikefalin treatment arm and the established clinical management arm are presented in Table 46 and Table 47, respectively.

Table 46: Transition probabilities - difelikefalin arm base case

		After				
		None	**Mild**	**Moderate**	**Severe**	**Very severe**
Cycle 1 (baseline to Week 4)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	66.71%	33.29%	0.00%	0.00%	0.00%
	Moderate	6.91%	59.80%	33.29%	0.00%	0.00%
	Severe	0.00%	6.91%	59.80%	33.29%	0.00%
	Very severe	0.00%	0.00%	6.91%	59.80%	33.29%
Cycle 2 (Week 4 to Week 8)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	18.47%	81.53%	0.00%	0.00%	0.00%
	Moderate	0.00%	18.47%	81.53%	0.00%	0.00%
	Severe	0.00%	0.00%	18.47%	81.53%	0.00%
	Very severe	0.00%	0.00%	0.00%	18.47%	81.53%
Cycle 3 (Week 8 to Week 12)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	11.81%	88.19%	0.00%	0.00%	0.00%
	Moderate	0.00%	11.81%	88.19%	0.00%	0.00%
	Severe	0.00%	0.00%	11.81%	88.19%	0.00%
	Very severe	0.00%	0.00%	0.00%	11.81%	88.19%
Cycle 4 (Week 12 to Week 64)						

		After				
		None	**Mild**	**Moderate**	**Severe**	**Very severe**
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	91.83%	8.17%	0.00%	0.00%	0.00%
	Moderate	8.54%	83.29%	8.17%	0.00%	0.00%
	Severe	0.00%	8.54%	83.29%	8.17%	0.00%
	Very severe	0.00%	0.00%	8.54%	83.29%	8.17%
Cycle 5 (Week 64 onwards)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	0.00%	100.00%	0.00%	0.00%	0.00%
	Moderate	0.00%	0.00%	100.00%	0.00%	0.00%
	Severe	0.00%	0.00%	0.00%	100.00%	0.00%
	Very severe	0.00%	0.00%	0.00%	0.00%	100.00%

Table 47: Transition probabilities - established clinical management arm base case

		After				
		None	**Mild**	**Moderate**	**Severe**	**Very severe**
Cycle 1 (baseline to Week 4)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	49.25%	50.75%	0.00%	0.00%	0.00%
	Moderate	5.03%	44.22%	50.75%	0.00%	0.00%
	Severe	0.00%	5.03%	44.22%	50.75%	0.00%

		After				
		None	**Mild**	**Moderate**	**Severe**	**Very severe**
	Very severe	0.00%	0.00%	5.03%	44.22%	50.75%
Cycle 2 (Week 4 to Week 8)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	9.92%	90.08%	0.00%	0.00%	0.00%
	Moderate	0.00%	9.92%	90.08%	0.00%	0.00%
	Severe	0.00%	0.00%	9.92%	90.08%	0.00%
	Very severe	0.00%	0.00%	0.00%	9.92%	90.08%
Cycle 3 (Week 8 to Week 12)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	9.42%	90.58%	0.00%	0.00%	0.00%
	Moderate	0.00%	9.42%	90.58%	0.00%	0.00%
	Severe	0.00%	0.00%	9.42%	90.58%	0.00%
	Very severe	0.00%	0.00%	0.00%	9.42%	90.58%
Cycle 4 (Week 12 to Week 64)						
Before	None	100.00%	0.00%	0.00%	0.00%	0.00%
	Mild	0.00%	100.00%	0.00%	0.00%	0.00%
	Moderate	0.00%	0.00%	100.00%	0.00%	0.00%
	Severe	0.00%	0.00%	0.00%	100.00%	0.00%
	Very severe	0.00%	0.00%	0.00%	0.00%	100.00%
Cycle 5 (Week 64 onwards)						
Before	None	95.00%	5.00%	0.00%	0.00%	0.00%

		After				
		None	**Mild**	**Moderate**	**Severe**	**Very severe**
	Mild	0.00%	95.00%	5.00%	0.00%	0.00%
	Moderate	0.00%	0.00%	95.00%	5.00%	0.00%
	Severe	0.00%	0.00%	0.00%	95.00%	5.00%
	Very severe	0.00%	0.00%	0.00%	0.00%	100.00%

B.3.3.2 Mortality and transplant rates

Fishbane et al., (2022) report outcomes for the all-difelikefalin-exposure cohort, including participants who received one or more doses of IV difelikefalin during the DB or OLE period of the KALM-1 and KALM-2 trials, as well as participants from the two additional open-label Phase 3 supportive studies (CLIN3101 and CLIN3105). They report a total of 56 deaths among the 1,306 participants (incidence rate of 69.0/1,000 PY) in the all-difelikefalin exposure cohort. No renal transplants were observed.

In the cost-effectiveness analysis of haemodiafiltration versus high flux HD as presented in NICE guideline for RRT and conservative management (NG107), a time-dependent annual transplant probability was applied to those who were alive on HD and a time-dependent annual mortality probability was applied to those who were alive on HD and had not had a transplant (55). Annual probabilities of death and transplant were estimated for people on HD from 1 to 10 years after initiating dialysis. The annual probability of death after Year 10 was assumed to be the same as in Year 10.

The probabilities used in the NG107 cost-effectiveness analysis were informed by a novel analysis of data from the UKRR using data on a UK adult incident cohort starting RRT on HD between January 2005 and December 2014. The analysis conducted by the UKRR has not been updated since, but a review of annual mortality rates in patients on HD from 2015 to 2019 indicates that this data may still be generalisable to the population included in this analysis.

In the base case, both mortality and transplant rates were modelled using the methods adopted in the cost-effectiveness analysis of HDF versus high flux HD, as presented in NICE guideline for RRT and conservative management (NG107), given the improved generalisability to the UK ESRD population. The resulting probabilities are summarised in Table 48.

Table 48: Probability of death and transplant each year post initiation of HD used in model

Year	Probability of death	Probability of transplant
1	0.187	0.039
2	0.140	0.050

Year	Probability of death	Probability of transplant
3	0.144	0.058
4	0.156	0.060
5	0.166	0.058
6	0.170	0.051
7	0.188	0.049
8	0.201	0.030
9	0.187	0.030
10	0.200	0.017
11+	0.200	0.000

The post-transplant population is assumed to follow age-adjusted all-cause mortality using rates obtained from UK life tables. A weighted age-dependent mortality probability was calculated using the proportion of male patients in the model.

The Dialysis Outcomes and Practice Patterns Study (DOPPS) is an international prospective cohort study of adult patients treated with in-centre HD in 21 countries. In a recent review of DOPPS data, Sukul et al., 2021, report the adjusted all-cause mortality hazard ratio, in comparison with patients who did not report pruritus, for patients extremely, very much, and moderately bothered by pruritus to be 1.24 (95% CI, 1.08-1.41), 1.02 (95% CI, 0.91-1.14), and 1.11 (95% CI, 1.00-1.22), respectively. All models adjusted for potential confounders, including 15 comorbid conditions; results show that CKD-aP is an independent predictor of patient mortality.

In the base case analysis, an increased mortality risk for the very severe, severe, and moderate CKD-aP population is applied using a hazard ratio of 1.24, 1.02, and 1.11, which is based on the findings reported by Sukul et al., (2021). The time-dependent probabilities of death are appropriately converted to rates, adjusted using the hazard ratio, and converted back to a probability for use in the model.

B.3.3.3 Adverse events

The AEs considered in the model are based on the commonly reported treatment-emergent AEs (TEAE) from the all-difelikefalin-exposure cohort reported in Fishbane et al., (2022). The most commonly reported TEAEs (≥2%) occurring in participants in the difelikefalin group and with ≥1% higher incidence than placebo were diarrhoea, dizziness, nausea, gait disturbance including falls, hyperkalaemia, headache,

somnolence, and mental status changes. The incidence and annual probability of AEs used in the model are summarised in Table 49.

Table 49: Incidence and annual probability of AEs used in model

Adverse event	Incidence rate (per 1,000 PY) from Fishbane et al., 2022		Annual probability used in model	
	All-difelikefalin-exposure	Placebo	Difelikefalin	Established clinical management
Diarrhoea	266.2	267.2	0.234	0.234
Dizziness	151.6	188.0	0.141	0.171
Nausea	225.6	207.8	0.202	0.188
Gait disturbance (falls)	267.5	237.5	0.235	0.211
Hyperkalaemia	157.8	158.3	0.146	0.146
Headache	106.0	118.7	0.101	0.112
Somnolence	39.4	95.9	0.039	0.091
Abbreviations: PY: person years Annual probability calculated as = 1-EXP(-r/t)				

As the rate of disease progression in patients receiving no treatment compared with patients receiving established clinical management is unknown, treatment discontinuation as a result of TEAEs has not been included in the model. This limitation is considered to be conservative against the cost-effectiveness of difelikefalin as the incidence rate of TEAEs leading to discontinuation was lower for all-difelikefalin-exposure cohort compared with placebo (196.0/1,000 PY for the all-difelikefalin-exposure cohort and 395.8/1,000 PY for the placebo cohort) (42) respectively.

Update following NICE clarification questions: Adverse events were restricted to be applied only within the first 3 model cycles (12-weeks).

B.3.4 Measurement and valuation of health effects

B.3.4.1 Health-related quality of life data from clinical trials

No generic preference-based measures of health were collected in the KALM-1 or KALM-2 trials. As noted in Section B.3.3.1.1, itch-related quality of life was included

as a secondary measure using the 5-D Itch scale, a multidimensional questionnaire which assesses itch-related quality of life over the previous 2 weeks.

Using NRS categories for the 5-D Itch scale (53), the 5-D Itch scale total scores collected in the KALM-1 and KALM-2 trials have been used to inform health state definitions and transition probabilities within the model.

B.3.4.2 Mapping

As no generic preference-based measures of health were collected in the KALM-1 or KALM-2 trials, a separate primary data collection study across UK dialysis centres was undertaken to develop a mapping algorithm relating the WI-NRS and 5-D Itch Scale to the EQ-5D-3L (Appendix J1.2 Mapping study).

Full details of the mapping study can be found in Appendix J. In summary, primary data collection was undertaken between November 2020 and June 2021 across 5 sites in England on adult patients (18+) who had been receiving haemodialysis for at least 3 months. The data collected was used to estimate EQ-5D-3L mapping functions from 5-D Itch scale scores, WI-NRS, and 5-D Itch scale scores and WI-NRS combined. All mapping functions included age, sex, diabetes status, and length of time on dialysis as additional conditioning variables. Despite limitations with missing observations, the 5-D Itch scale score to EQ-5D-3L mapping algorithm was considered the most appropriate option, given the paucity of published data in CKD-aP.

Table 50 provides the results of the EQ-5D-3L predictions based on the KALM-1 and KALM-2 data and using the 5-D Itch scale mapping algorithm. In the mapping study, the severe and very severe (unbearable) populations were merged, given the small numbers of observations in each group. In the base case analysis, the utility scores for the severe and very severe populations are set to be equal.

Table 50: CKD-aP severity utility scores used in the model

CKD-aP severity	Mean utility	95% CI
Not present	[redacted]	[redacted]
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]

CKD-aP severity	Mean utility	95% CI
Severe/unbearable	[redacted]	[redacted]
Abbreviations: CKD-aP: chronic kidney disease-associated pruritus; CI: confidence interval		

B.3.4.3 Health-related quality of life studies

An SLR was undertaken in April-July 2022 to identify HRQoL data for CKD-aP, with a particular focus on the UK and Europe. Full details of the SLR search strategy, study selection process, and results are presented in Appendix G. The review identified 3 studies providing utility evidence.

The study by (60) presents the results of the mapping study discussed in Section B.3.4.2, as well as utility values estimated from the SHAREHD database. Utility values from the SHAREHD database are presented in scenario analysis.

The study by (61) assessed different approaches to mapping individual questions from itch-related outcome data collected in the Phase 2 difelikefalin trial (CLIN2101) to estimate EQ-5D-5L scores.

The study by (62) investigated the effectiveness of zolpidem and acupressure therapy on food acupoints in improving the sleep quality and overall quality of life of haemodialysis patients with CKD-aP in Pakistan. The study categorised participants into “no problems” and “problems” and observed a numerical improvement in the mean EQ-5D index score in the control group, from 0.49 (±0.30) at baseline to 0.53 (±0.30) at Week 8 (p=0.187).

B.3.4.4 Adverse reactions

The SLR for utilities did not identify any HRQoL values for adverse events in patients with CKD-aP.

Utility values used to reflect the CKD-aP health states in the base case are informed by the mapping study. Adverse events were not reported; however, it is assumed that utility scores reported include any disutility associated with AEs. Furthermore, the incremental incidence of adverse events reported in Fishbane et al., (2022) for the results of the pooled KALM-1 and KALM-2 trials were small and in general lower in those patients treated with DFK, suggesting that observed AEs are likely to be a feature of underlying disease. As such, any utility decrements associated with their

incidence will be implicitly captured in health state utility values. To avoid the risk of double counting QALY loss, adverse event QALY loss is set to 0 for all adverse events in the base case analysis.

B.3.4.5 Health-related quality of life data used in the cost-effectiveness analysis

The SLR for utilities did not identify values of HRQoL for transplant. Instead, NICE HTAs identified in the expanded SLR were reviewed. Health state utility values for transplant were informed by Lee et al., 2005, which was identified in NICE TA775.

Table 51 provides a summary of the utility values used in the cost-effectiveness analysis for difelikefalin compared with established clinical management for treatment of adults with moderate-to-severe CKD-aP.

Table 51: Summary of utility values for cost-effectiveness analysis

	Mean	SE	Source	Reference in submission
Health state utility values				
None	[redacted]	[redacted]	Mapping study (Appendix J)	B.3.4.2 Mapping
Mild	[redacted]	[redacted]		B.3.4.2 Mapping
Moderate	[redacted]	[redacted]		B.3.4.2 Mapping
Severe	[redacted]	[redacted]		B.3.4.2 Mapping
Very severe	[redacted]	[redacted]		B.3.4.2 Mapping
Transplant	0.71	0.04	TA775 referencing (63)	B.3.4.5 Health-related quality of life data used in the cost-effectiveness analysis
Adverse event QALY loss				
Diarrhoea	0.00	0.00	Assumed zero to avoid double counting	B.3.4.4 Adverse reactions
Dizziness	0.00	0.00		B.3.4.4 Adverse reactions
Nausea	0.00	0.00		B.3.4.4 Adverse reactions
Gait disturbance (falls)	0.00	0.00		B.3.4.4 Adverse reactions
Hyperkalaemia	0.00	0.00		B.3.4.4 Adverse reactions
Headache	0.00	0.00		B.3.4.4 Adverse reactions
Somnolence	0.00	0.00		B.3.4.4 Adverse reactions

Abbreviations: QALY: quality adjusted life year; SE: standard error

B.3.5 Cost and healthcare resource use identification, measurement and valuation

B.3.5.1 Intervention and comparators' treatment costs

Difelikefalin is administered by intravenous bolus injection at the end of haemodialysis treatment. The recommended dose of difelikefalin is 0.5 micrograms/kg dry body weight. The total dose volume (mL) required from the vial should be calculated as follows: 0.01 × dry body weight (kg), rounded to the nearest tenth (0.1 mL) (64).

£[redacted] per vial, a [redacted] reduction.

Table 52 provides the injection volumes and required number of vials per weight band as per the SmPC. Because vials are recommended for single use only, the number of vials required has been rounded to account for unused fill volume. The list price of difelikefalin is £35 per vial and the PAS price of difelikefalin is £[redacted] per vial, a [redacted] reduction.

Table 52: Difelikefalin dose, injection volumes and required number of vials

Weight range (kg) *<lower bound>*	Injection volume (mL)	Vials required
40	0.40	1
45	0.50	1
55	0.60	1
65	0.70	1
75	0.80	1
85	0.90	1
95	1.00	1
105	1.10	2
115	1.20	2
125	1.30	2
Notes: Number of vials required is rounded to account for wastage **Abbreviations:** Kg: kilograms; mL: millilitre		

The UKRR reports that in 2019, 5.5% of patients were undertaking in-centre haemodialysis (ICHD) less than 3 times per week, 92.7% exactly 3 times per week, and 1.8% more than 3 times per week. In the model, a weighted frequency of 2.96

dialyses sessions per week was used by assuming 2 and 4 sessions per week for those under and over 3 ICHD sessions per week, respectively. Using a frequency of 2.96 sessions per week, the annual cost of difelikefalin is estimated at £5,392.66 at list price, and £[redacted] at PAS price.

As noted in Section B.3.2.2, there are currently no approved treatments for CKD-aP and treatment instead focuses on symptom management. BAD recommends only capsaicin cream, topical calcipotriol, or oral gabapentin, and advises against sedative antihistamines and cetirizine (21). The average annual cost for established clinical management applied in the model was estimated based on data on background CKD-aP treatment collected in the mapping study. This has been detailed in Section B.3.4.2. Table 53 provides a summary of the data collected and used in the model to inform costs for established clinical management. Unit costs, dose and pack size for the established clinical management treatments were sourced from the BNF. In the mapping study, the severe and very severe (unbearable) populations were merged, given the small numbers of observations in each group. In the base case analysis, the resource use for the severe and very severe populations are set to be equal.

Table 53: Established clinical management resource use and treatment costs

	None	Mild	Moderate	Severe	Very severe	Dose (per pack	Pack size	Pack cost	Weekly cost	Source
Receiving anti-pruritic medication	**40.20%**	**38.50%**	**36.60%**	**55.60%**	**55.60%**					
Topical corticosteroids	2.40%	3.40%	1.60%	7.40%	7.40%	1.50	15	£1.26	£0.88	BNF; Hyrdocortisone; Mild inflammatory skin disorders; 1% cream AAH Pharmaceuticals
Oral corticosteroids	15.20%	9.40%	8.10%	16.00%	16.00%	1.00	30	£0.86	£0.20	BNF; Loratadine; Symptomatic relief of allergy such as hay fever, chronic idiopathic urticaria; 10mg
Antihistamines	4.90%	7.70%	13.80%	24.70%	24.70%	1.50	84	£2.36	£0.30	BNF; Hyrdroxyzxine hydrochloride; Elderly dose; 10mg tablets AAH Pharmaceuticals
Gabapentin/pregabalin	11.00%	7.70%	6.50%	17.30%	17.30%	3.00	100	£2.74	£0.58	BNF; Gabapentin; Peripheral neuropathic pain; 300mg capsule; Alliance Healthcare (Distribution) Ltd
Monteleukast	1.80%	2.60%	0.80%	1.20%	1.20%	1.00	28	£1.37	£0.34	BNF; Monteleukast; Prophylaxis of asthma; 10mg tablet; A A H Pharmaceuticals Ltd NHS Indicative price
Antidepressants	12.50%	17.90%	1.63%	21.00%	21.00%	3.50	30	£13.66	£11.16	BNF; Doxepin; Pruritus in eczema; Xepin 5% cream Cambridge Healthcare Supplies Ltd
Anxiolytic/sedatives	4.30%	4.30%	1.60%	4.90%	4.90%	-	-	-	-	No appropriate cost could be identified.

Abbreviations: BNF: British National Formulary

In the modified Delphi panel, it was noted that treatment for patients with CKD-aP is additive, and that difelikefalin may be used as an adjunct with established clinical management where established clinical management is insufficient in reducing pruritus (Appendix N: Clinical opinion and consensus report). Therefore, total treatment costs for the difelikefalin arm include the weighted treatment cost for established clinical management. Table 54 provides a summary of the total treatment costs. Using the resource use data collected in the mapping study (Table 53), it was noted that the total weighted treatment cost for moderate CKD-aP was lower than that for 'mild' and 'none' severity. This was driven by the greatly reduced proportion of people with moderate CKD-aP using antidepressants. Given that established clinical management is additive, in the base case analysis, the weighted total treatment costs for moderate CKD-aP were set to be equal to the weighted total treatment costs for mild CKD-aP.

Table 54: Summary of total weighted treatment costs by health state

CKD-aP severity	Established clinical management arm	Difelikefalin arm	
		List price	PAS price
None	£31.98	£5,424.64	[redacted]
Mild	£42.48	£5,435.14	[redacted]
Moderate	£42.48*	£5,435.14	[redacted]
Severe	£75.65	£5,468.31	[redacted]
Very severe	£75.65	£5,468.31	[redacted]

Notes: * Established clinical management costs for moderate CKD-aP were adjusted in model to equal costs for mild CKD-aP; the very severe health state is assumed equivalent to the severe health state
Abbreviations: CKD-aP: chronic kidney disease associate pruritus; PAS: patient access scheme

B.3.5.2 Health state management costs

An SLR was undertaken in April-July 2022 to identify cost and healthcare resources, using data for CKD-aP which focused primarily on the UK and Europe. Full details of the SLR search strategy, study selection process, and results are presented in Appendix G. Among the 7 included studies assessing cost and resource use data in CKD-aP patients, 2 were conducted in the US (6, 65) and 1 study each was

conducted globally (11), in Italy (66), India (67), Saudi Arabia (68), and Taiwan (69). For brevity, only the results of the global resource use study are presented here.

As noted in Section B.3.3.2, Sukul et al., (2021) analysed the data from the international perspective cohort study, DOPPS. The analysis included 23,264 haemodialysis patients who responded to a survey question asking about the extent the patient was bothered by itchy skin during the past 4 weeks. The proportions of patients not at all, somewhat, moderately, very much, and extremely bothered by pruritus in the UK were 29%, 24%, 21%, 15%, and 12% respectively. Sukul et al., used Cox regression for time-to-event outcomes and modified Poisson regression for binary outcomes to estimate the risk of mortality and hospitalisation in people with CKD-aP. All models adjusted for potential confounders, including 15 comorbid conditions. Table 55 shows the hazard ratios for adjusted all-cause hospitalisation compared with patients who reported being not at all bothered by itchy skin. The adjusted annual rate of hospitalisation in patients who reported being not at all bothered by itchy skin is 0.895. It has been assumed that the verbal rating scale included in the DOPPS questionnaire corresponds to the verbal rating scale used in Lai et al., (2017) and the KALM trials.

Table 55: Hazard ratios for all-cause hospitalisation

CKD-aP severity	Hazard ratio	95% CI
None	[redacted]	
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]
Severe	[redacted]	[redacted]
Very severe	[redacted]	[redacted]

Abbreviations: CKD-aP: chronic kidney disease-associated pruritus; CI: confidence interval

The unit cost per hospitalisation was estimated to be £3,004.43, a figure determined using the National Cost Collection (2020/2021) weighted average of codes for CKD with and without interventions for all CC scores (LA08G, LA08H, LA08J, LA08K, LA08L, LA08M, LA08N, and LA08P).

In the modified Delphi panel, it was noted that patients would have a 3-monthly patient review conducted by the consultant nephrologists. A specialist visit was costed at £242.48 per visit using the National Cost Collection (2020/2021) cost for a Nephrologist consultant led, non-admitted face-to-face follow-up attendance (WF01A).

The cost of a haemodialysis session was estimated to be £169.34 using the National Cost Collection (2020/2021) weighted average of codes LD05A and LD06A. In the base case, dialysis costs were not included given the adjustment in risk of mortality for the very severe CKD-aP population and the resulting indirect increase in survival for the difelikefalin treatment arm. This approach was deemed appropriate with reference to NICE guidance (Section 4.4.16) which states that where a technology increases survival in people for whom the NHS is currently providing expensive care, background care costs may be removed.

The cost of transplant was estimated to be £20,901.72 using the National Cost Collection (2020/2021) weighted average of codes LA01A, LA02A, and LA03A, plus the weighted average of codes LA11Z, and LA12A, plus the weighted average of codes LA13A, and LA14Z. The post-transplant cost was estimated to be £5,913.50 per year, informed by the NHS Blood and Transplant fact sheet 7 (2009). The post-transplant costs were inflated to 2021 prices using PSSRU inflation indices (70).

B.3.5.3 Adverse reaction unit costs and resource use

The AEs included in the model have been detailed in Section B.3.4.4. For the all-difelikefalin-exposure and placebo cohort of the difelikefalin trials, AEs were mild or moderate in severity (≥65% of any of the events) in the majority of patients (42). Given that no relevant or appropriate costs for AEs were identified in either the SLR or the adapted SLR, in the base case analysis, AEs were costed as a single GP appointment (£33.19; PSSRU 2021).

B.3.6 Severity

The technology is not expected to meet the criteria for a severity weight.

B.3.7 Summary of base case analysis inputs and assumptions

B.3.7.1 Summary of base case analysis inputs

A summary of the base case cost-effectiveness analysis inputs is provided in Table 56.

Table 56: Summary of variables applied in the economic model

Variable	Value	SE	Distribution	Section in submission
General settings				
Time horizon	Lifetime			B.3.2.1 Perspective
Discount rate (costs and outcomes)	3.5%			
Baseline demographics				
Mean age	58.30 years			B.3.2.3 Patient population
Mean weight	84.40kg	4.31	Normal	
Time spent on dialysis	4.78 years			
Sex (% male)	59.58%			
Proportion in state at model entry				
None	0.00%			B.3.2.4 Model structure
Mild	0.00%			
Moderate	55.28%			
Severe	34.17%			
Very severe	10.55%			
Adverse event (annual probability) – difelikefalin				
Diarrhoea	0.234			B.3.3.3 Adverse events
Dizziness	0.141			
Nausea	0.202			
Gait disturbance (falls)	0.235			
Hyperkalaemia	0.146			
Headache	0.101			
Somnolence	0.039			
Adverse event (annual probability) – established clinical management				

Variable	Value	SE	Distribution	Section in submission
Diarrhoea	0.234			B.3.3.3 Adverse events
Dizziness	0.171			
Nausea	0.188			
Gait disturbance (falls)	0.211			
Hyperkalaemia	0.146			
Headache	0.112			
Somnolence	0.091			
Adverse event – cost per event	£33.19	£1.96	Gamma	B.3.5.3 Adverse reaction unit costs and resource use
Adverse event – QALY loss per event	0.00			B.3.4.5 Health-related quality of life data used in the cost-effectiveness analysis
Treatment costs – difelikefalin (list price; PAS price)				
None	£5,424.64; [redacted]			B.3.5.1 Intervention and comparators' treatment costs
Mild	£5,435.14; [redacted]			
Moderate	£5,435.14; [redacted]			
Severe	£5,468.31; [redacted]			
Very severe	£5,468.31; [redacted]			
Treatment costs – established clinical management				
None	£31.98	+-10%	Gamma	B.3.5.1 Intervention and comparators' treatment costs
Mild	£42.48			
Moderate	£42.48			
Severe	£75.65			

Variable	Value	SE	Distribution	Section in submission
Very severe	£75.65			
Management costs				
None	£3,659.29			B.3.5.2 Health state management costs
Mild	£3,686.19			
Moderate	£3,820.66			
Severe	£4,008.91			
Very severe	£4,224.06			
Transplant	£20,901.72	+-10%	Gamma	
Post-transplant	£5,913.50			
Hospitalization hazard ratios				
Mild	1.01	0.03	Normal	B.3.5.2 Health state management costs
Moderate	1.06	0.03		
Severe	1.13	0.04		
Very severe	1.21	0.05		
Utilities				
None	0.617	0.03	Beta	B.3.4.5 Health-related quality of life data used in the cost-effectiveness analysis
Mild	0.579	0.02		
Moderate	0.514	0.02		
Severe	0.429	0.03		
Very severe	0.429	0.03		
Transplant	0.712	+-10%		
Transplant by time on haemodialysis (annual probability)				
Year 1	3.90%	0.20%	Beta	B.3.3.2 Mortality and transplant rates
Year 2	5.00%	0.26%		
Year 3	5.80%	0.30%		
Year 4	6.00%	0.31%		
Year 5	5.80%	0.30%		
Year 6	5.10%	0.26%		
Year 7	4.90%	0.25%		
Year 8	3.00%	0.15%		

Variable	Value	SE	Distribution	Section in submission
Year 9	3.00%	0.15%		
Year 10	1.70%	0.09%		
Year 11+	0.00%	0.00%		
Mortality by time on haemodialysis (annual probability)				
Year 1	18.70%	0.95%	Beta	B.3.3.2 Mortality and transplant rates
Year 2	14.00%	0.71%		
Year 3	14.40%	0.73%		
Year 4	15.60%	0.80%		
Year 5	16.60%	0.85%		
Year 6	17.00%	0.87%		
Year 7	18.80%	0.96%		
Year 8	20.10%	1.03%		
Year 9	18.70%	0.95%		
Year 10	20.00%	1.02%		
Year 11+	20.00%	1.02%		
Risk of mortality by health state (hazard ratio)				
None	1.00			B.3.3.2 Mortality and transplant rates
Mild	1.00			
Moderate	1.12	0.06	Normal	
Severe	1.01	0.06		
Very severe	1.24	0.09		
Efficacy – 5-D Itch scale total score mean change from baseline				
DFK Moderate - Week 4	-2.90	0.24	Normal	B.3.3.1 Efficacy
DFK Moderate - Week 8	-3.54	0.24		
DFK Moderate - Week 12	-3.92	0.24		
DFK Moderate - Week 64	-6.96	0.24		
DFK Severe - Week 4	-5.13	0.30		
DFK Severe - Week 8	-5.90	0.31		
DFK Severe - Week 12	-6.66	0.35		
DFK Severe - Week 64	-10.75	0.35		

Variable	Value	SE	Distribution	Section in submission
DFK Very severe - Week 4	-6.05	0.63		
DFK Very severe - Week 8	-6.78	0.77		
DFK Very severe - Week 12	-7.93	0.70		
DFK Very severe - Week 64	-12.49	0.70		
ECM Moderate - Week 4	-1.28	0.19		
ECM Moderate - Week 8	-2.16	0.21		
ECM Moderate - Week 12	-2.65	0.23		
ECM Moderate - Week 64	-3.39	0.26		
ECM Severe - Week 4	-3.81	0.30		
ECM Severe - Week 8	-4.93	0.35		
ECM Severe - Week 12	-5.09	0.37		
ECM Severe - Week 64	-7.45	0.36		
ECM Very severe - Week 4	-5.09	0.63		
ECM Very severe - Week 8	-5.91	0.75		
ECM Very severe - Week 12	-6.14	0.81		
ECM Very severe - Week 64	-9.35	0.69		
Distribution of count of 5-D Itch scale total score at baseline				
5	0.0%	0.00%		
6	0.0%	0.00%		
7	0.0%	0.00%		
8	0.0%	0.00%		
9	0.0%	0.00%		
10	0.0%	0.00%		
11	0.0%	0.00%	Beta	B.3.3.1 Efficacy
12	4.5%	0.74%		
13	8.0%	0.96%		
14	9.4%	1.04%		
15	11.3%	1.12%		
16	10.9%	1.11%		
17	11.1%	1.11%		

Variable	Value	SE	Distribution	Section in submission
18	10.8%	1.10%		
19	8.5%	0.99%		
20	7.2%	0.91%		
21	7.7%	0.94%		
22	5.0%	0.77%		
23	1.9%	0.48%		
24	2.6%	0.57%		
25	1.0%	0.35%		
Note: All annual probabilities were converted to monthly probabilities for the first 4 cycles in the model **Abbreviations:** ECM: established clinical management; DFK: difelikefalin				

B.3.7.2 Assumptions

A list of the assumptions made in the base case analysis and their justifications is provided in Table 57. Where appropriate, the exploration of the potential impact of these assumptions in a scenario analysis is noted.

Table 57: Summary of assumptions used in the analysis

Model input	Description of assumption	Justification
Model structure	Renal transplant is assumed to be a definitive treatment for CKD-aP.	CKD-aP is associated with renal failure. Following transplant, patients will have regained kidney function.
Discontinuation	It is assumed that people may discontinue treatment with difelikefalin following a 'run-in' period to determine a clinical response to treatment. People not achieving a clinically meaningful response will discontinue treatment and progress through the rest of the model at the same rate as the ECM arm.	As patients in the KALM trials were not discontinued in this way, it is not possible to exclusively measure efficacy and model outcomes for this patient group.
Treatment waning	In the base case, a treatment waning was modelled for the ECM arm.	The company sought clinical opinion on the natural progression of CKD-aP and the potential trend in the mean change in itch score that could be expected in the extrapolation period for patients receiving placebo and DFK in the

Model input	Description of assumption	Justification
		KALM trials. It was noted that the placebo effect would wane over time in line with the natural progression of the disease (which is likely to get worse over time), and that data for the DFK arm would suggest no treatment waning effect. In the base case, to reflect clinical opinion, treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year. The waning effect has been applied to both treatment arms in scenario analysis.
Efficacy	It is assumed that the probability of improving or deteriorating CKD-aP in each cycle is equal no matter the current health state.	Creating a matrix which calculates the probability of moving from any one state to each of the other states can result in small observation numbers estimating a single probability value, which may lead to unrealistic outcomes. Furthermore, because an extrapolation of the trial data was required to estimate the long-term efficacy for patients receiving placebo (Cycle 4 onwards), using estimates of a mean change in itch score from baseline would result in all placebo patients remaining as mild or moderate CKD-aP. As such, transition matrices are derived from per-cycle probabilities of losing or gaining health states.
Efficacy	Transition probabilities were generated for a maximum improvement or deterioration by 3 health states.	This assumption was examined, and it was observed that only 3 patients in the difelikefalin treatment arm and 1 patient in the placebo arm improved by 4 health states across the duration of both the KALM-1 and KALM-2 trials. Furthermore, no patients deteriorated by 4 health states in either treatment arm.
Utilities	It is assumed that utility scores reported and estimated from the mapping study include any disutility associated with AEs supported in the model.	The incremental incidence of adverse events reported in Fishbane et al., (2022) for the results of the pooled KALM-1 and KALM-2 trials were small and in general lower in those patients treated with DFK, suggesting that observed AEs are likely to be a

Model input	Description of assumption	Justification
		feature of underlying disease. As such, any utility decrements associated with their incidence will be implicitly captured in health state utility values.
Treatment costs	The weighted total treatment costs for moderate CKD-aP were set to be equal to the weighted total treatment costs for mild CKD-aP.	Using the resource use data collected in the mapping study (Table 53), it was noted that the total weighted treatment cost for moderate CKD-aP was lower than that for 'mild' and 'none' severity. This was noticed as being driven by the greatly reduced proportion of people with moderate CKD-aP using antidepressants. To align with clinical guidance that treatment is additive, in the base case, treatment costs for moderate CKD-aP were set to be equal to treatment costs for mild CKD-aP.

B.3.8 Base case results

The deterministic base case cost-effectiveness analysis results of difelikefalin compared with established clinical management for treatment of adults with moderate-to-severe CKD-aP over a lifetime time horizon are summarised in Table 58 (List price; £35.00 per vial) and Table 59 (PAS price; [redacted]).

Treatment with difelikefalin with established clinical management compared with established clinical management alone was associated with increased life years (0.06 per person) and increased QALYs ([redacted] per person) at an incremental cost of £8,453 per person at list price and [redacted] per person at PAS price.

At PAS price, treatment with difelikefalin with established clinical management is cost-effective at a willingness-to-pay (WTP) threshold of £30,000/QALY.

The incremental QALYs were driven by an increase in the number of people in less severe CKD-aP states. The Markov traces displaying the distribution of patients across health states are presented in Appendix J.

Table 58: Base case deterministic results (list price)

	Difelikefalin + ECM	ECM	Incremental	ICER
Total costs (£)	£32,097	£23,644	£8,453	£61,157
Total life years (LY)	4.65	4.59	0.06	
Total QALYs	[redacted]	2.75	[redacted]	
Abbreviations: ECM: established clinical management; LY: Life years; QALY: quality-adjusted life year; ICER: incremental cost-effectiveness ratio				

Table 59: Base case deterministic results (PAS price)

	Difelikefalin + ECM	ECM	Incremental	ICER
Total costs (£)	[redacted]	£23,644	[redacted]	£23,277
Total LY	4.65	4.59	0.06	
Total QALYs	[redacted]	2.75	[redacted]	
Abbreviations: ECM: established clinical management; LY: Life years; QALY: quality-adjusted life year; ICER: incremental cost-effectiveness ratio				

Table 60 provides a summary of the disaggregated costs, and Table 61 provides a summary of the QALYs and life years associated with difelikefalin with established clinical management and established clinical management alone.

Table 60: Base case disaggregated costs

	Difelikefalin + ECM	ECM	Incremental
Treatment (list price)	£8,545	£136	£8,408
Treatment (PAS price)	[redacted]	£136	[redacted]
Adverse events	£17	£37	-£20
Management			
None	£3,717	£679	£3,038
Mild	£2,256	£2,664	-£408
Moderate	£2,384	£3,710	-£1,325

	Difelikefalin + ECM	ECM	Incremental
Severe	£1,385	£2,344	-£959
Very severe	£447	£777	-£330
Transplant	£13,345	£13,297	£48
Abbreviations: ECM: established clinical management			

Table 61: Base case disaggregated QALYs and life years

	Difelikefalin + ECM	ECM	Incremental
Life years (total)	**4.65**	**4.59**	**0.06**
None	1.02	0.19	0.83
Mild	0.61	0.72	-0.11
Moderate	0.62	0.97	-0.35
Severe	0.35	0.58	-0.24
Very severe	0.11	0.18	-0.08
Transplant	1.95	1.94	0.01
QALYs (total)	[redacted]	**2.75**	[redacted]
None	[redacted]	0.11	[redacted]
Mild	[redacted]	0.42	[redacted]
Moderate	[redacted]	0.50	[redacted]
Severe	[redacted]	0.25	[redacted]
Very severe	[redacted]	0.08	[redacted]
Transplant	[redacted]	1.38	[redacted]
Adverse events	[redacted]	0.00	[redacted]
Abbreviations: ECM: established clinical management; QALY: quality-adjusted life year			

B.3.9 Exploring uncertainty

As the base case ICER of difelikefalin compared with established clinical management for treatment of adults with moderate-to-severe CKD-aP well exceeded a WTP of £30,000/QALY at list price, only results at PAS price are presented in the sensitivity and subgroup analyses.

B.3.9.1 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was performed to explore the effect of uncertainty associated with key model inputs. PSA results for 1,000 iterations are presented in Table 62. The mean incremental costs and QALYs of difelikefalin with established clinical management compared with established clinical management alone were calculated to estimate the probabilistic ICER.

The probabilistic results were comparable with the deterministic results. The incremental per patients QALYs and costs in the probabilistic analysis results were [redacted] and [redacted] respectively, compared to [redacted] and [redacted] in the deterministic analysis results. The PSA scatter plot and cost-effectiveness acceptability curve are shown in **Error! Reference source not found.** and **Error! Reference source not found.**.

The ICER in the probabilistic analysis remained cost-effective at PAS price with an ICER of £23,253. The probability of cost-effectiveness is 84% at a WTP threshold of £30,000/QALY and 24% at a WTP threshold of £20,000/QALY.

Table 62: Probabilistic results (PAS price)

	Difelikefalin + ECM	ECM	Incremental	ICER
Total costs (£)	[redacted]	£23,684	[redacted]	£23,253
Total LY	-	-	-	
Total QALYs	[redacted]	2.77	[redacted]	
Abbreviations: ECM: established clinical management; LY: Life years; QALY: quality-adjusted life year; ICER: incremental cost-effectiveness ratio				

[redacted]

[redacted]

[redacted]

B.3.9.2 Deterministic sensitivity analysis

Deterministic sensitivity analyses (DSA) were performed to explore the effect of uncertainty associated with varying individual model inputs. The inputs with an impact on the ICER of ≥£1,000/QALY are presented in descending order as a tornado plot in **Error! Reference source not found.**.

The cost-effectiveness of difelikefalin with established clinical management is most sensitive to changes in the health state utility scores. The largest driver of change is the utility score for the ‘none’ CKD-aP severity health state. In only 1 of the varied parameter values with impact on the ICER of ≥£1,000/QALY did the ICER exceed a WTP threshold of £30,000/QALY.

[redacted]

B.3.9.3. Scenario analysis

A range of scenario analyses were conducted to test the robustness of the model to alternative model inputs and assumptions. The details of the undertaken analyses and the results of the scenario analyses, presented as the ICER of difelikefalin with established clinical management compared with established clinical management alone, are shown in Table 63.

Table 63: Summary of scenario analysis with ICER results at PAS price

Scenario	Description	ICER
Base case deterministic results		**£23,277**
1.	Established clinical management extrapolation As noted in Section B.3.3.1.4, 3 possible extrapolation methods were considered in the analysis: mean difference (MD), ratio of means (RoM), and no change in efficacy. The results of the MD and ROM extrapolations are presented here.	
1.a.	Mean difference long-term extrapolation for ECM arm.	£30,054
1.b	Ratio of means long-term extrapolation for ECM arm	£25,409
2.	Stopping rule at Week 8 In the base case analysis, a stopping rule is implemented at Week 12 (end of Cycle 3) to reflect short-term treatment decisions and initial response to treatment. Here, patients who do not achieve a clinically significant response on difelikefalin discontinue treatment to remain only on established clinical management. A scenario is presented where the stopping rule is applied in Week 8 (end of Cycle 2).	
2.a.	Stopping rule applied in Week 8.	£23,077
3.	KALM-1 and KALM-2 separately In the base case, data from the pooled KALM-1 and KALM-2 trials were presented. Two scenarios are presented using patient-level data taken from each trial independently.	
3.a.	KALM-1 trial data only	£25,817
3.b.	KALM-2 trial data only	£19,805
4.	Observed data In the base case analysis, transition probability matrices were estimated using a simulated data set using the mean change from baseline in itch scores by CKD-aP severity at baseline. Using the simulated data set in the base case approach was preferred as it offers a better reflection of the underlying trend in the data and a more appropriate quantification of the uncertainty in the mean change in itch scores through probabilistic analysis. A scenario analysis is presented using the observed data directly.	
4.a	Observed data set	£37,913
5.	Efficacy plateau In the base case analysis, the treatment effect of difelikefalin was assumed to remain unchanged from Week 64 (end of Cycle 4) onwards. A scenario is presented whereby the mean change from Week 12 to Week 64 (Cycle 4) for the difelikefalin arm is continued for an additional year (Cycle 5) before being set to plateau for the duration of the model (remain unchanged).	

Scenario	Description	ICER
5.a	Efficacy plateau after Year 2	£21,475
6.	Treatment waning The company sought clinical opinion on the natural progression of CKD-aP and the potential trend in the mean change in itch score which could expected in the extrapolation period for patients receiving placebo in the KALM trials. It was noted that the placebo effect would wane over time in line with the natural progression of the disease (which is likely to get worse over time). In the base case analysis, a waning effect was applied only to the established clinical management arm. Two scenarios are presented for which treatment waning is applied to both treatment arms.	
6.a.	In Cycle 5, assume a probability of 5% for patients to gain a health state (deteriorate) each cycle.	£25,915
6.b.	In Cycle 5, assume a probability of 10% for patients to gain a health state (deteriorate) each cycle.	£26,016
7.	SHAREHD utility scores In deterministic analysis, it was noted that health state utility scores were a driver of the ICER for difelikefalin. A scenario is presented using utility scores for the 5 CKD-aP health states derived from the SHAREHD study and presented in Soro et al., 2022.	
7.a.	SHAREHD utility scores	£21,584

B.3.11 Subgroup analysis

Subgroup analysis was conducted to explore the effects of variations in the target population: specifically, use of anti-itch medication at baseline and baseline itch severity. The results of the subgroup analyses, presented as the ICER of difelikefalin with established clinical management compared with established clinical management alone, are shown below.

Table 64 presents cost-effectiveness analysis results of difelikefalin compared with established clinical management in patients **only receiving anti-itch medication at baseline.** In this scenario, efficacy is updated to reflect the subgroup and established clinical management treatment costs are updated to reflect 100% use of treatment costs.

Table 64: Subgroup analysis A: Only receiving anti-itch medication at baseline

	Difelikefalin + ECM	ECM	Incremental	ICER
Total costs (£)	[redacted]	£23,814	[redacted]	£23,993
Total LY	4.66	4.59	0.06	
Total QALYs	[redacted]	2.74	[redacted]	
Abbreviations: ECM: established clinical management; LY: Life years; QALY: quality-adjusted life year; ICER: incremental cost-effectiveness ratio				

Table 65 presents cost-effectiveness analysis results of difelikefalin compared with established clinical management in patients **not receiving anti-itch medication at baseline.** In this scenario, efficacy is updated to reflect the subgroup, and established clinical management treatment costs are updated to reflect 0% use of treatment costs.

Table 65: Subgroup analysis B: Not receiving anti-itch medication at baseline

	Difelikefalin + ECM	ECM	Incremental	ICER
Total costs (£)	[redacted]	£23,511	[redacted]	£25,922
Total LY	4.65	4.59	0.05	
Total QALYs	[redacted]	2.75	[redacted]	
Abbreviations: ECM: established clinical management; LY: Life years; QALY: quality-adjusted life year; ICER: incremental cost-effectiveness ratio				

Table 66 presents cost-effectiveness analysis results of difelikefalin compared with established clinical management in patients **only with severe or very severe itch at baseline.** In this scenario, efficacy and model entry proportions are updated to reflect the subgroup.

Table 66: Subgroup Analysis C: Severe and Very severe itch at baseline

	Difelikefalin + ECM	ECM	Incremental	ICER
Total costs (£)	[redacted]	£23,486	[redacted]	£18,642
Total LY	4.66	4.54	0.12	

	Difelikefalin + ECM	ECM	Incremental	ICER
Total QALYs	[redacted]	2.67	[redacted]	
Abbreviations: ECM: established clinical management; LY: Life years; QALY: quality-adjusted life year; ICER: incremental cost-effectiveness ratio				

B.3.12 Benefits not captured in the QALY calculation

- Sukul et al., 2021 report that patients with extremely itchy skin were 50% more likely to withdraw from dialysis than patients not at all bothered by itchy skin. It has not been possible to capture the impact of withdrawal from dialysis for patients.
- Research indicates that EQ-5D is not sensitive to changes in some psychological disorders and conditions affecting sensory functions. Improvements in, for example, sleep and skin irritation are not adequately captured by changes to EQ-5D index scores. Therefore, the use of difelikefalin may result in significant health-related benefits that are unlikely to be included in the QALY calculation.

B.3.13 Validation

Internal quality assurance measures were undertaken throughout the model development. The model was validated through the use of extreme values and formula auditing to ensure the consistency of model estimates.

The model structure and inputs were critiqued and validated by an external and clinician and health economics consultant. Where appropriate, any errors were amended. Overall, the validation identified no issues with the structural or computational accuracy of the model.

B.3.14 Interpretation and conclusions of economic evidence

The cost-effectiveness of difelikefalin with established clinical management compared with established clinical management alone for treatment of adults with moderate-to-severe CKD-aP has been evaluated in line with the NICE final scope.

The treatment effect of difelikefalin was derived from the pooled analysis of the KALM-1 and KALM-2 trials, in which significantly greater proportions of participants in the difelikefalin group achieved clinically meaningful improvements in itch severity and itch-related quality of life versus the placebo group (5-D Itch scale: : [redacted] respectively, at Week 12; P = 0.01) over 12 weeks of treatment. For patients who continued difelikefalin treatment in the OLE extension period, itch improvement (measured as mean 5-D Itch scale total score) was maintained and emerged in patients who switched from placebo to difelikefalin during the OLE. The health state utility values were derived from a separate primary data collection study across UK dialysis centres, undertaken to develop a mapping algorithm relating the WI-NRS and 5-D Itch Scale to the EQ-5D-3L. Additionally, scenario analysis was conducted, and results are presented using health state utility values derived from the SHAREHD trial conducted in 12 renal centres in England. Costs were identified from UK sources, including NHS reference costs and literature. Where required, additional model inputs were sourced from published NICE technical appraisals in CKD and atopic dermatitis, including the cost-effectiveness analysis of HDF versus high flux HD as presented in NICE guideline for RRT and conservative management (NG107).

Extensive scenario analyses demonstrate the base case cost-effectiveness results to be robust to variation in model inputs and assumptions, with only 2 of the 10 scenarios did the ICER exceed a WTP threshold of £30,000/QALY. Deterministic sensitivity analysis demonstrates the results to be sensitive to changes in the utility scores for the 5 CKD-aP severity states. When using alternative utility scores, as derived from the SHAREHD analysis, the cost-effectiveness of difelikefalin improves. Subgroup analysis demonstrates that in patients only with severe or very severe itch at baseline, the cost-effectiveness of difelikefalin improves and falls below a WTP threshold of £20,000/QALY.

In summary, the results of this analysis demonstrate that difelikefalin represents a cost-effective use of NHS resources in adults with moderate-to-severe CKD-aP where established clinical management is insufficient in reducing pruritus, with an ICER of £23,277/QALY gained.

B.4 References

1. Kidney Research UK. Kidney Health Inequalities in the UK - An agenda for change 2018.
2. Crews DC, Bello AK, Saadi G. Burden, access, and disparities in kidney disease. Journal of nephrology. 2019;32(1).
3. Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al. Global Prevalence of Chronic Kidney Disease - A Systematic Review and Meta-Analysis. PloS one. 2016;11(7).
4. WHO. Leading causes of death - Cause-specific mortality, 2000–2019 2020 [Available from: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death.
5. Miskawaan Health. Understanding Chronic Kidney Disease - What Is Chronic Kidney Disease? 2022 [Available from: https://www.miskawaanhealth.com/understanding-chronic-kidney-disease/
6. Ramakrishnan K, Bond TC, Claxton A, Sood VC, Kootsikas M, Agnese W, et al. Clinical characteristics and outcomes of end-stage renal disease patients with self-reported pruritus symptoms. International journal of nephrology and renovascular disease. 2013;7.
7. Mathur VS, Lindberg J, Germain M, Block G, Tumlin J, Smith M, et al. A longitudinal study of uremic pruritus in hemodialysis patients. Clinical journal of the American Society of Nephrology : CJASN. 2010;5(8).
8. Mettang T, Kremer AE. Uremic pruritus. Kidney international. 2015;87(4).
9. Sukul N, Speyer E, Tu C, Bieber BA, Li Y, Lopes AA, et al. Pruritus and Patient Reported Outcomes in Non-Dialysis CKD. Clinical journal of the American Society of Nephrology : CJASN. 2019;14(5).
10. Rayner HC, Larkina M, Wang M, Graham-Brown M, van der Veer SN, Ecder T, et al. International Comparisons of Prevalence, Awareness, and Treatment of Pruritus in People on Hemodialysis. Clinical journal of the American Society of Nephrology : CJASN. 2017;12(12).
11. Sukul N, Karaboyas A, Csomor PA, Schaufler T, Wen W, Menzaghi F, et al. Self-reported Pruritus and Clinical, Dialysis-Related, and Patient-Reported Outcomes in Hemodialysis Patients. Kidney medicine. 2021;3(1).
12. Locatelli F, Legat FJ. Chronic Kidney Disease-Associated Pruritus: From Epidemiology to Treatment. European Medical Journal. 2021;9(4):2-10.
13. Nguyen NTQ, Cockwell P, Maxwell AP, Griffin M, O'Brien T, O'Neill C. Chronic kidney disease, health-related quality of life and their associated economic burden among a nationally representative sample of community dwelling adults in England. PloS one. 2018;13(11).
14. Maung SC, El Sara A, Chapman C, Cohen D, Cukor D. Sleep disorders and chronic kidney disease. World journal of nephrology. 2016;5(3).
15. van der Willik EM, Lengton R, Hemmelder MH, Hoogeveen EK, Bart HAJ, van Ittersum FJ, et al. Itching in dialysis patients: impact on health-related quality of life and interactions with sleep problems and psychological symptoms - results from the RENINE/PROMs registry. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2022.
16. Pisoni RL, Wikström B, Elder SJ, Akizawa T, Asano Y, Keen ML, et al. Pruritus in haemodialysis patients: International results from the Dialysis Outcomes

and Practice Patterns Study (DOPPS). Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2006;21(12).
17. Palmer S, Vecchio M, Craig JC, Tonelli M, Johnson DW, Nicolucci A, et al. Prevalence of depression in chronic kidney disease: systematic review and meta-analysis of observational studies. Kidney international. 2013;84(1).
18. Balaskas EV, Grapsa I. Uremic pruritus is a poor prognostic factor of outcome. Peritoneal dialysis international : journal of the International Society for Peritoneal Dialysis. 1995;15(2).
19. Mettang T, Pauli-Magnus C, Alscher A. Uraemic pruritus--new perspectives and insights from recent trials. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2002;17(9).
20. Kimmel M, Alscher DM, Dunst R, Braun N, Machleidt C, Kiefer T, et al. The role of micro-inflammation in the pathogenesis of uraemic pruritus in haemodialysis patients. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2006;21(3).
21. Millington GWM, Mustapa MF, Levell NJ, Collins A, Lovell CR, Leslie TA, et al. British Association of Dermatologists' guidelines for the investigation and management of generalized pruritus in adults without an underlying dermatosis, 2018. The British journal of dermatology. 2018;178(1).
22. Committee for Medicinal Products for Human Use. Kapruvia - Assessment Report 2022.
23. Committee for Medicinal Products for Human Use. Kapruvia - Summary of opinion 2022.
24. Cara Therapeutics Inc. CR845-CLIN3102 Clinical Study Report – Double-blind Phase (AIC). 2020 26/11/2020.
25. Cara Therapeutics Inc. CR845-CLIN3103 Clinical Study Report – Double-blind Phase (AIC). 2020 6/10/2020.
26. Cara Therapeutics Inc. CR845-CLIN3102 Clinical Study Report – Open-label Extension Phase (AIC). 2020 10/12/2020.
27. Cara Therapeutics Inc. CR845-CLIN3103 Clinical Study Report – Open-label Extension Phase (AIC). 2020 12/11/2020.
28. Cara Therapeutics Inc. CR845-CLIN3105 Clinical Study Report (AIC). 2020 29/09/20.
29. Cara Therapeutics Inc. CR845-CLIN3102 Protocol (AIC). 2017 21/12/17.
30. Cara Therapeutics Inc. CR845-CLIN3103 Protocol (AIC). 2019 10/09/2019
31. Cara Therapeutics Inc. CR845-CLIN3105 Protocol (AIC). 2019 12/02/2019.
32. Cara Therapeutics Inc. CR845-CLIN3102 Statistical Analaysis Plan - Double-Blind Phase (AIC). 2019 07/05/2019.
33. Cara Therapeutics Inc. CR845-CLIN3102 Statistical Analysis Plan - Open-Label Extension Phase (AIC). 2020 09/03/2020.
34. Cara Therapeutics Inc. CR845-CLIN3103 Statistical Analysis Plan - Double-Blind Phase (AIC). 2020 09/03/2020.
35. Cara Therapeutics Inc. CR845-CLIN3103 Statistical Analysis Plan - Open Label Long Term Extension Phase (AIC). 2020.
36. Cara Therapeutics Inc. CR845-CLIN3105 Statistical Analysis Plan (AIC). 2020 09/03/2020.
37. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. The British journal of dermatology. 2010;162(3).

38. Dworkin RH, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005;113(1-2).
39. Topf J, Wooldridge T, McCafferty K, Schömig M, Csiky B, Zwiech R, et al. Efficacy of Difelikefalin for the Treatment of Moderate-to-Severe Pruritus in Hemodialysis Patients: Pooled Analysis of KALM-1 and KALM-2 Phase 3 Studies. Kidney medicine. 2022.
40. Vernon M, Ständer S, Munera C, Spencer RH, Menzaghi F. Clinically meaningful change in itch intensity scores: An evaluation in patients with chronic kidney disease-associated pruritus. Journal of the American Academy of Dermatology. 2021;84(4).
41. Data on file. CLIN3105 Post-hoc Analysis (CIC). 2022.
42. Fishbane S, Wen W, Munera C, Lin R, Bagal S, McCafferty K, et al. Safety and Tolerability of Difelikefalin for the Treatment of Moderate-to-Severe Pruritus in Hemodialysis Patients: Pooled Analysis From the Phase 3 Clinical Trial Program. Kidney Medicine. 2022.
43. Caplin B, Kumar S, Davenport A. Patients' perspective of haemodialysis-associated symptoms. Nephrol Dial Transplant. 2011;26(8):2656-63.
44. Fishbane S, Wen W, Munera C, Menzaghi F, McCafferty K. Long-term Safety and Efficacy of Difelikefalin in Subjects With Chronic Kidney Disease–Associated Pruritus: Analysis From KALM-1 and KALM-2. SKIN The Journal of Cutaneous Medicine. 2021;5.
45. Csiky B, Walpen S, Schaufler T, Morin I, Wen W, Menzaghi F. MO467: Time to Improvement of Itch Intensity in Patients With Chronic Kidney Disease-Associated Pruritus Treated With Difelikefalin. Nephrology Dialysis Transplantation. 2022;37(Supplement_3).
46. Krajnik M, Zylicz Z. Understanding pruritus in systemic disease. Journal of pain and symptom management. 2001;21(2).
47. Meeuwis SH, van Middendorp H, van Laarhoven AIM, van Leijenhorst C, Pacheco-Lopez G, Lavrijsen APM, et al. Placebo and nocebo effects for itch and itch-related immune outcomes: A systematic review of animal and human studies. Neuroscience and biobehavioral reviews. 2020;113.
48. van Laarhoven AIM, Vogelaar ML, Wilder-Smith OH, van Riel P, van de Kerkhof PCM, Kraaimaat FW, et al. Induction of nocebo and placebo effects on itch and pain by verbal suggestions. Pain. 2011;152(7).
49. Bartels DJ, van Laarhoven AI, Haverkamp EA, Wilder-Smith OH, Donders AR, van Middendorp H, et al. Role of conditioning and verbal suggestion in placebo and nocebo effects on itch. PloS one. 2014;9(3).
50. NICE. NICE health technology evaluations: the manual 2022 [Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation.
51. UK Renal Registry. UK Renal Registry, 23rd Annual Report. Data to 31/12/2019. 2021.
52. The Renal Association. Clinical Practice Guideline - Haemodialysis. 2019.
53. Lai JW, Chen HC, Chou CY, Yen HR, Li TC, Sun MF, et al. Transformation of 5-D itch scale and numerical rating scale in chronic hemodialysis patients. BMC nephrology. 2017;18(1).
54. Pérez-Martín J, Bermejo I, Díez F, J. Evaluation of Markov Models with Discontinuities. Medical decision making : an international journal of the Society for Medical Decision Making. 2019;39(4).

55. NICE. RRT and conservative management - Cost-effectiveness analysis: HDF versus high flux HD. 2018.
56. Kumagai H, Ebata T, Takamori K, Muramatsu T, Nakamoto H, Suzuki H. Effect of a novel kappa-receptor agonist, nalfurafine hydrochloride, on severe itch in 337 haemodialysis patients: a Phase III, randomized, double-blind, placebo-controlled study. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association. 2010;25(4).
57. Ständer S, Augustin M, Reich A, Blome C, Ebata T, Phan NQ, et al. Pruritus assessment in clinical trials: consensus recommendations from the International Forum for the Study of Itch (IFSI) Special Interest Group Scoring Itch in Clinical Trials. Acta dermato-venereologica. 2013;93(5).
58. Yanagida T. Package 'misty' - Miscellaneous Functions 2022 2022-06-08.
59. van Buuren S. Package 'mice' - Multivariate Imputation by Chained Equations. 2021 24/11/2021.
60. Soro M, Thokala P, Fotheringham J. POSC312 A Methodological Approach to Assess the Economic Value of Difelikefalin to Treat Chronic Kidney Disease Associated Pruritus (CKD-AP) - Value in Health. Value in Health. 2022;25.
61. Schaufler T, C. M, Ramirez de Arellano Serna A, F. M. PUK28 EMULATING EQ-5D DATA FROM AVAILABLE PROS TO OPTIMIZE CLINICAL TRIAL PLANNING IN CHRONIC KIDNEY DISEASE-ASSOCIATED PRURITUS. Value in Health. 2019;22.
62. Rehman IU, Ahmed R, Rahman AU, Wu DBC, Munib S, Shah Y, et al. Effectiveness and safety profiling of zolpidem and acupressure in CKD associated pruritus: An interventional study. Medicine. 2021;100(21).
63. Lee AJ, Morgan CL, Conway P, Currie C, J. . Characterisation and comparison of health-related quality of life for patients with renal failure. Current medical research and opinion. 2005;21(11).
64. EMA. Kapruvia - Summary of Product Characteristics 2022.
65. Ramakrishnan K, Sood V, Sibbel SP. Economic Impact of Pruritus Among End-Stage Renal Disease Patients Receiving Hemodialysis | Ramakrishnan, K; Sood, V; Sibbel, S.P | download. Value in Health. 2013;16(7).
66. Calabria S, Ronconi G, Piccinni C, Dondi L, Cinconze E, Pedrini A, et al. A Retrospective Observational Analysis of Italian Reimbursed Healthcare for Patients on Hemodialysis and Treated for Chronic Kidney Disease-Associated Pruritus. . Pre Print not yet peer reviewed. 2022.
67. Suja A, Anju R, Anju V, Neethu J, Peeyush P, Saraswathy R. Economic evaluation of end stage renal disease patients undergoing hemodialysis. Journal of pharmacy & bioallied sciences. 2012;4(2).
68. Khan TM, Al-Haider I, Syed Sulaiman SA, Hassali MS. Linguistic validation of the 5D itching scale to Arabic in patients with end-stage kidney disease. Journal of renal care. 2013;39(4).
69. Ting SW, Fan PC, Lin YS, Lin MS, Lee CC, Kuo G, et al. Uremic pruritus and long-term morbidities in the dialysis population. PloS one. 2020;15(10).
70. Jones K, Burns A. Unit Costs of Health and Social Care 2021. Personal Social Services Research Unit Kent, UK, ; 2021.
71. Ruikar V. Interactive Voice/Web Response System in clinical research. Perspectives in clinical research. 2016;7(1).

72. Suresh K. An overview of randomization techniques: An unbiased assessment of outcome in clinical research. Journal of human reproductive sciences. 2011;4(1).

B.5 Appendices

Appendix C: Summary of product characteristics (SmPC) and UK public assessment report

C1.1 SmPC

Please see document ‘Appendix C1 – SmPC’

C1.2 Public assessment report

No UK public assessment report is available – please refer to document ‘Appendix C2 – EPAR’

Appendix D: Identification, selection and synthesis of clinical evidence

D1.1 Identification and selection of relevant studies

Please see document 'Appendix D, G, H, I – SLR Results'

D1.2 Participant flow in the relevant randomised control trials

KALM-1

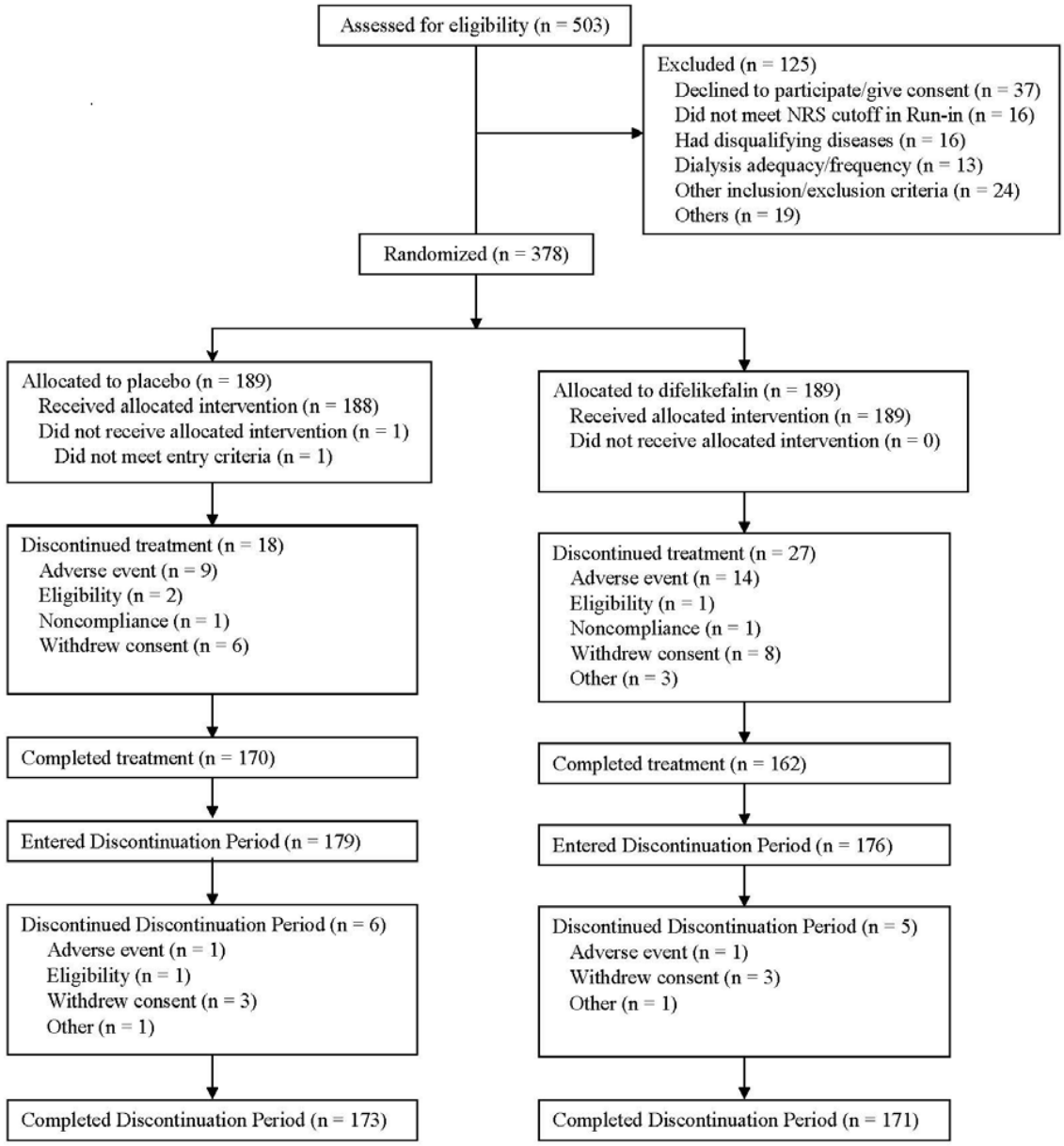
Subject disposition

A total of 503 subjects enrolled in the study, of whom 125 failed screening, 378 underwent randomisation, and 377 received at least one treatment with difelikefalin (189 subjects) or placebo (188 subjects) during the double-blind treatment period (Figure 11). One subject, Subject 108013, was randomised but did not receive any study drug. No unblinding was performed for medical management of study subjects during the double-blind phase of the study. A total of 332 subjects (88.1% of the 377 subjects randomised and exposed to study drug) completed the double-blind treatment period. 45 subjects (11.9%) discontinued early from the double-blind treatment period, with the discontinuation rate being greater in the difelikefalin group than in the placebo group (14.3% versus 9.6%). The most common (≥2% of all subjects) reasons for early discontinuation from the double-blind treatment period were AEs (6.1%) and withdrawal of subject consent (3.7%). The most frequently reported TEAEs leading to study drug discontinuation were septic shock (3 subjects [0.8%]), dizziness (3 subjects [0.8%]), pneumonia, sepsis, and mental status changes (2 subjects [0.5%] each) (Table 27). Other reasons for early discontinuation from the double-blind treatment period were eligibility criteria (0.8%), "other" (0.8%), and subject noncompliance (0.5%). Reasons categorised as "other" or "withdrawal of consent" were reviewed to confirm that study drug was not discontinued for a different reason (none of them were). With respect to treatment group differences in reasons for early discontinuation, only the percentage of subjects who discontinued because of an AE was greater (by factor of 1.5 or more) in the difelikefalin group than in the placebo group (7.4% versus 4.8%). Per the study protocol, subjects who discontinued early from the double-blind treatment period were still to complete the

double-blind discontinuation period, if possible. A total of 355 subjects (94.2% of the 377 subjects randomised and exposed to study drug) entered the double-blind discontinuation period, and 344 of these subjects completed this period.

The percentages of subjects entering and completing the discontinuation period were similar between the difelikefalin and placebo groups (93.1% and 95.2% , for entry, respectively; 97.2% and 96.6% for completion, respectively). Of the 11 subjects (3.1% of 355 subjects entering the period) who prematurely discontinued from the discontinuation period, the most common reason was withdrawal of subject consent (1.7%), followed by AE (0.6%), “other” (0.6%), and eligibility (0.3%). The difelikefalin and placebo groups were comparable with respect to the frequency of early withdrawal from the discontinuation period and the reasons for early withdrawal.

Figure 11 Study disposition flow diagram, KALM-1



Abbreviations: NRS = numerical rating scale

Protocol deviations

111 subjects (29.4%) reported at least one major protocol deviation, with the percentages being similar between the difelikefalin and placebo groups (29.6% and 29.3%, respectively). The most frequently (≥2% of all subjects) identified categories of major deviations were informed consent (5.3%), investigational product accountability management (4.5%), delegation of authority (4.5%), ≥25% WI-NRS scores missing (4.5%), tests/assessments/procedure (3.7%), subject not dosed in

either Week 11 or Week 12 (3.7%), and received <80% of planned study doses (3.2%). Subjects who had deviations categorised as “other” represented 5.6% of the double-blind safety population; examples of these deviations included dispensing of investigational product based on incorrect stratification for medical history, incorrect recording of use of anti-itch medication at the time of randomisation, and stratification of subject by medical history in error (Listing 16.2.2). Overall, the analysis showed no notable patterns with respect to major protocol deviations. The composition of the per protocol population, which was based primarily on excluding subjects with major protocol deviations, is described in Section 11.1. Minor protocol deviations were identified in most study subjects, i.e., 341 of the 377 subjects (90.5%) in the double-blind safety population, and the incidence was similar between the difelikefalin and placebo groups (89.4% and 91.5%, respectively).

KALM-1 OLE

Subject disposition

[redacted]

[redacted]

[redacted]

[redacted]

Protocol deviations

[redacted]

[redacted]

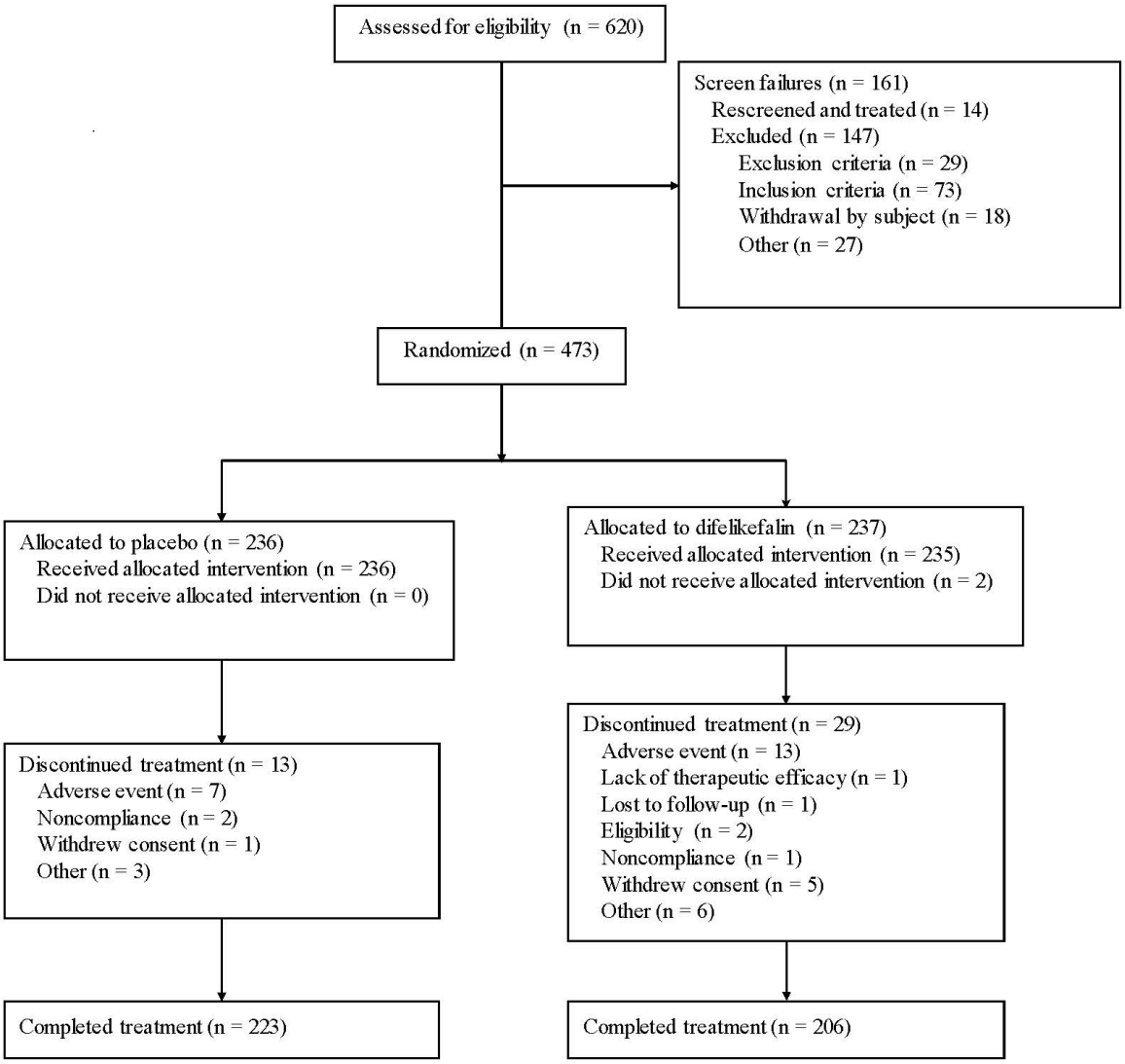
KALM-2

Subject disposition

A total of 620 subjects enrolled in the study, of whom 161 failed screening, 473 underwent randomisation, and 471 received at least one treatment with difelikefalin (235 subjects) or placebo (236 subjects) during the double-blind treatment period (Figure 12). Two subjects, Subject 036008005 and Subject 840002006, were randomised to difelikefalin but did not receive any study drug. Unblinding was performed for one subject (036003001) who experienced the SAE of small bowel obstruction, which was considered related to study treatment.

A total of 429 subjects (91.1% of the 471 subjects randomised and exposed to study drug) completed the double-blind treatment period. 42 subjects (8.9%) discontinued early from the double-blind treatment period, with the percentage being greater in the difelikefalin group than in the placebo group (12.3% versus 5.5%). The most common (≥2% of all subjects) reason for early discontinuation from the double-blind treatment period was an AE (4.2%). The most frequently reported TEAEs leading to study drug discontinuation in the difelikefalin group were anxiety (0.9%) and insomnia (0.9%). Other reasons for early discontinuation from the double-blind treatment period were “other” (1.9%), subject withdrawal of consent (1.3%), subject noncompliance (0.6%), and eligibility criteria (0.4%). Reasons categorised as “other” or “withdrawal of consent” were reviewed to confirm that study drug was not discontinued for a different reason (none of them were). With respect to treatment group differences in reasons for early discontinuation, the percentage of subjects who discontinued was greater in the difelikefalin group than in the placebo group for the following reasons: AE (5.5% versus 3.0%), subject withdrew consent (2.1% versus 0.4%), and “other” (2.6% versus 1.3%).

Figure 12 Study disposition flow diagram, KALM-2



Protocol deviations

161 subjects (34.2%) reported at least one major protocol deviation, with this percentage being comparable between the difelikefalin and placebo groups (37.9% and 30.5%, respectively). The most frequently (≥2% of all subjects) identified categories of major deviations were dosing noncompliance (17.4%), procedure not performed (15.3%), and procedure performed out of window (6.2%). Subjects who had deviations categorised as “other” represented 14.0% of the double-blind safety population; examples of these deviations included error in stratification of subject by medical history, incorrect recording of use of anti-itch medication at the time of randomisation, AE/SAE not reported within 24 hours, incomplete laboratory sample processing, inadequate informed consent administration, incorrect handling of IP,

and insufficient subject training on questionnaires and other documentation. Overall, the analysis showed no notable patterns with respect to major protocol deviations.

178 subjects (37.8%) reported at least one minor protocol deviation, and the incidence was comparable between the difelikefalin and placebo groups (37.0% and 38.6%, respectively).

KALM-2 OLE

Subject disposition

[redacted]

[redacted]

[redacted]

[redacted]

Protocol deviations

[redacted]

CLIN3105

Subject disposition

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Protocol deviations

[redacted]

D1.3 Critical appraisal for each study

Below are the complete quality assessments for each trial included in this submission:

Table 67 KALM-1 complete quality assessment

Trial number (acronym)	Response	How is the question addressed?
Was randomisation carried out appropriately?	Yes	Before the start of the study a computer-generated randomisation schedule was prepared using Interactive Web Response Systems (IWRS) and Interactive Voice Response Systems (IVRS) (71). A stratified randomisation method was used, and patients were randomised to a ratio of 1:1 to receive either difelikefalin or placebo. Using a stratified randomisation method addresses the need to control and balance any potential influence that covariates may have on the clinical outcomes (72). Patients were stratified according to their use or non-use of concomitant medications to treat their itch during the week prior to randomisation (run-in period), as well as the presence or absence of specific medical conditions: history of fall or fracture, confusional state or mental status change or altered mental status, disorientation, and gait disturbance or movement disorder.
Was the concealment of treatment allocation adequate?	Yes	An interactive voice/ web response was used to determine treatment assignment (71). The labelling of the study drug (either difelikefalin 0.5 mcg/kg or placebo IV solution) was also blinded.

Trial number (acronym)	Response	How is the question addressed?
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	The groups were similar with regard to demographic and baseline disease characteristics.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	During the double-blind treatment period, patients, investigators, study staff, and the sponsor were blinded to the study drug assignment. For medically urgent or emergent situations that necessitate knowledge of study drug assignment for patient management, the blind may be broken via the IVRS/IWRS. Whenever possible, the Medical Monitor was to be contacted prior to breaking the blind. Blinding of patients, investigators, study staff and the sponsor was not changed during the 'unblinded interim analysis'. This analysis was conducted by the IDMC. Members of the IDMC did not participate in the Data Safety Monitoring Board (DSMB) and were not members of the study team.
Were there any unexpected imbalances in drop-outs between groups?	No	Discontinuations of study medication were low and well-balanced between treatment arms.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Based on the clinical study report, all outcomes outlined in the methodology are reported in detail.
Did the analysis include an ITT analysis? If so, was this appropriate	Yes	Analysis was performed on an ITT population (defined as group of patients who are randomised to a treatment group) consisting of 378 subjects (189 difelikefalin and 189 placebo). Subjects in the ITT population were analysed according to their

Trial number (acronym)	Response	How is the question addressed?
and were appropriate methods used to account for missing data?		randomised treatment, regardless of the actual treatment received. All efficacy analysis was conducted on the ITT population.

Table 68 KALM-1 OLE complete quality assessment

Question	Response	How is the question addressed?
Was the cohort recruited in an acceptable way?	Yes	To be eligible for the OLE phase of the study, a subject had to have received at least 30 doses of the study drug (either placebo or active) during the 12-week double-blind treatment period and had to continue to meet the other eligibility criteria listed in Appendix O.
Was the exposure accurately measured to minimise bias?	Yes	The instruments used to assess the impact of interventions on CKD-aP have been carefully developed and validated in populations relevant to difelikefalin. As well as assessing clinical changes in itch intensity, the instruments are designed to establish the patient's perception of their itch and its impact on their QoL. As such, the findings of the studies described in this submission demonstrate both the clinical benefit of difelikefalin and its humanistic benefit (i.e., the noticeable improvements experienced by patients across all aspects of their lives). Subjects were also analysed according to their previous treatment group in the double-blind treatment period of KALM-1.

Question	Response	How is the question addressed?
Was the outcome accurately measured to minimise bias?	Yes	The maintenance of the effect of difelikefalin on itch was measured by a patient-reported outcome (PRO), the 5-D Itch scale, with which data had already been recorded during the double-blind treatment period. The 5-D Itch scale is a brief multidimensional questionnaire designed to be useful as an outcome measure in clinical studies. The scale has been validated in patients with chronic pruritus, including patients undergoing haemodialysis, and has been shown to be sensitive to changes in pruritus over time (Elman et al., 2010). The 5-D Itch scale is appropriate for assessing itch and its impact in the subject population for this investigational product indication.
Have the authors identified all important confounding factors?	Yes	-
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes	All subjects participating in the OLE phase of the study received difelikefalin 0.5 mcg/kg. However, many of the analyses for the OLE phase are based on the subject's treatment assignment during the double-blind phase, in which subjects were randomised to treatment with either difelikefalin 0.5 mcg/kg or placebo. Randomisation during the double-blind phase was stratified based on use or non-use of concomitant medications to treat their itch during the pre-randomisation week (the run-in period), and the presence or absence of specific medical conditions.

Question	Response	How is the question addressed?
		All prior and concomitant medications used during the trial were recorded. The administration of erythropoiesis-stimulating agents and IV iron to the subject was recorded, as per the schedule of assessments. Subject compliance with study drug was documented as part of standard procedures at the dialysis units where the study drug was administered. Please see Table 14 for details on how missing data was accounted for in analyses.
Was the follow-up of patients complete?	Yes	In total, 127 (78.4%) in the Placebo/DFK group completed the follow-up visit, and 117 (77.5%) of the DFK/ DFK group completed the follow-up visit.
How precise (for example, in terms of confidence interval and p values) are the results?	Yes	For each time point, standard descriptive statistics were used to report observed scores and the changes from baseline, along with the LS means, standard errors, 95% CIs, and differences from baseline within each treatment sequence (reported with LS means, standard errors, and 95% CIs). See results section for KALM-1 OLE.

Table 69 KALM-2 complete quality assessment

Question	Response	How is the question addressed?
Was randomisation carried out appropriately?	Yes	Before the start of the study a computer-generated randomisation schedule was prepared using IWRS and IVRS (71). A stratified randomisation method was used, and patients were randomised to a ratio of 1:1 to receive either difelikefalin or placebo. Using a stratified randomisation method addresses the need to control and balance any potential influence that covariates may have on the clinical outcomes (72). Patients were stratified according to their use or non-use of concomitant medications to treat their itch during the week prior to randomisation (the run-in period) as well as the presence or absence of specific medical conditions: history of fall or fracture, confusional state or mental status, change or altered mental status, disorientation, gait disturbance, and movement disorder.
Was the concealment of treatment allocation adequate?	Yes	An interactive voice/ web response was used to determine treatment assignment (71). The labelling of the study drug (either difelikefalin 0.5 mcg/kg or placebo IV solution) was also blinded.
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	The groups were similar with regard to demographic and baseline disease characteristics.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	During the double-blind treatment period, patients, investigators, study staff, and the sponsor were blinded to the study drug assignment. For medically urgent or emergent situations that necessitate knowledge of study drug assignment for patient management, the blind may be broken via the IVRS/IWRS. Whenever

Question	Response	How is the question addressed?
		possible, the Medical Monitor was to be contacted prior to breaking the blind. Blinding of patients, investigators, study staff and the sponsor was not changed during the 'unblinded interim analysis'. This analysis was conducted by the IDMC. Members of the IDMC did not participate in the DSMB, and were not members of the study team.
Were there any unexpected imbalances in drop-outs between groups?	No	Discontinuations of study medication were low and well-balanced between treatment arms.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Based on the clinical study report, all outcomes outlined in the methodology are reported in detail.
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Analysis was performed on an ITT population (defined as group of patients who are randomised to a treatment group) consisting of 473 subjects (237 difelikefalin and 236 placebo). Subjects in the ITT population were analysed according to their randomised treatment, regardless of the actual treatment received. All efficacy analysis was conducted on the ITT population. Please see Table 14 for details on how missing data was accounted for in analyses.

Table 70 KALM-2 OLE complete quality assessment

Question	Response	How is the question addressed?
Was the cohort recruited in an acceptable way?	Yes	To be eligible for the OLEpPhase of the study, a subject had to have received at least 30 doses of the study drug (either placebo or active) during the 12-week double-blind treatment period, and had to continue to meet the other eligibility criteria listed in Appendix O.
Was the exposure accurately measured to minimise bias?	Yes	The instruments used to assess the impact of interventions on CKD-aP have been carefully developed and validated in populations relevant to difelikefalin. As well as assessing clinical changes in itch intensity, the instruments are designed to establish the patient's perception of their itch and its impact on their QoL. As such, the findings of the studies described in this submission demonstrate both the clinical benefit of difelikefalin and its humanistic benefit (i.e., the noticeable improvements experienced by patients across all aspects of their lives). Subjects were also analysed according to their previous treatment group in the double-blind treatment period of KALM-2.
Was the outcome accurately measured to minimise bias?	Yes	The maintenance of the effect of difelikefalin on itch was measured by a PRO, the 5-D Itch scale, with which data had already been during the double-blind treatment period. The 5-D Itch scale is a multidimensional questionnaire designed to be useful as an outcome measure in clinical studies. The scale has been validated in patients with chronic pruritus, including patients undergoing haemodialysis, and has been shown to be sensitive to changes in pruritus over

Question	Response	How is the question addressed?
		time (Elman et al., 2010). The 5-D Itch scale is appropriate for assessing itch and its impact in the subject population for this investigational product indication.
Have the authors identified all important confounding factors?	Yes	-
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes	All subjects participating in the OLE phase of the study received difelikefalin 0.5 mcg/kg. However, many of the analyses for the OLE phase are based on the subject's treatment assignment during the double-blind phase, in which subjects were randomised to treatment with either difelikefalin 0.5 mcg/kg or placebo. Randomisation during the double-blind phase was stratified based on use or non-use of concomitant medications to treat their itch during the pre-randomisation week (the run-in period), and the presence or absence of specific medical conditions. All prior and concomitant medications used during the trial were recorded. The use of anti-itch medications during the study was recorded on an ongoing basis, starting at screening. Subject compliance with study drug was documented as part of standard procedures at the dialysis units where the study drug was administered.
Was the follow-up of patients complete?	No	Due to an administrative decision by the sponsor (unrelated to treatment efficacy or safety), KALM-2 was halted early; as a result, 313 subjects (78.4%) could not complete the 52-week open-label treatment period.

Question	Response	How is the question addressed?
How precise (for example, in terms of confidence interval and p values) are the results?	Yes	For each time point, standard descriptive statistics were used to report observed scores and the changes from baseline, along with the LS means, standard errors, 95% CIs, and differences from baseline within each treatment sequence.

Table 71 CLIN3105 complete quality assessment

Question	Response	How is the question addressed?
Was the cohort recruited in an acceptable way?	Yes	Along with meeting other eligibility criteria listed in Appendix O, patients with ESRD receiving haemodialysis 3 times a week with moderate-to-severe pruritus were considered for participation in this study. The screening period occurred within 28 days prior to treatment to assess eligibility. It consisted of a screening visit and a run-in period. During the screening period, patients signed the ICF and then were evaluated for eligibility by assessment of inclusion/exclusion criteria. Eligible patients were then moved to the run-in period. The purpose of the run-in period was to confirm that each patient had moderate-to-severe pruritus as measured by the patient-reported Worst Itching Intensity NRS (i.e., weekly average worst itching score >5), and to establish a baseline itch intensity.
Was the exposure accurately measured to minimise bias?	Yes	The instruments used to assess the impact of interventions on CKD-aP have been carefully developed and validated in populations relevant to difelikefalin. As well as assessing clinical changes in itch intensity, the instruments are designed to establish the patient's perception of their itch and its impact on their QoL. As

Question	Response	How is the question addressed?
		such, the findings of the studies described in this submission demonstrate both the clinical benefit of difelikefalin and its humanistic benefit (i.e., the noticeable improvements experienced by patients across all aspects of their lives).
Was the outcome accurately measured to minimise bias?	Yes	Patients were trained on completion of the Worst Itching Intensity NRS, Sleep Quality, and EQ-5D-5L-P questionnaires prior to the first visit of the run-in period, and were trained on the other itch-related PRO measures at any time prior to dosing on Day 1 of the treatment period. The MOS Sleep Questionnaire was developed to assess sleep problems, and has been validated for use in measuring sleep disturbances in CKD-aP patients on HD (7). Because of the subjective nature of itch, it is generally accepted that at least two different measurements should be used to assess pruritus intensity in clinical studies (53). This is the case for the difelikefalin studies reported in this dossier. All instruments used to assess the impact of interventions on CKD-aP have been carefully developed and validated in populations relevant to difelikefalin.
Have the authors identified all important confounding factors?	Yes	-
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes	All prior and concomitant medications used during the trial were recorded. Use of antipruritic medications during the study was recorded on an ongoing basis, starting at screening.

Question	Response	How is the question addressed?
		Subject compliance with study drug was documented as part of standard procedures at the haemodialysis units where study drug was administered. Missed haemodialysis visits were documented on an ongoing basis during the treatment period. The eligibility of subjects was assessed during the run-in period. All eligible subjects received the same treatment.
Was the follow-up of patients complete?	Yes	-
How precise (for example, in terms of confidence interval and p values) are the results?	Yes	For each time point, standard descriptive statistics were used to report observed scores and the changes from baseline, along with the LS means, standard errors, 95% CIs, and differences from baseline within each treatment sequence.

Appendix E: Subgroup analysis

The following sections present in depth results of any subgroup analyses performed in both KALM-1 and KALM-2. An overview is provided in Table 72 below

Table 72 Summary of Week 12 ≥3-Point Worst Itching Intensity NRS improvement by stratification factors in KALM-1 and KALM-2

KALM-1			KALM-2		
	Placebo	**Difelikefalin**		**Placebo**	**Difelikefalin**
Anti-itch medication at baseline = no					
n	111	117	n	151	150
Observed Week 12 ≥3-point NRS improvement [1] — (%)					
Yes	31 (31.3%)	53 (53.5%)	Yes	51 (38.3%)	60 (48.0%)
No	68 (68.7%)	46 (46.5%)	No	82 (61.7%)	65 (52.0%)
Missing	12	18	Missing	18	25
LS means estimate of percent with improvement [2]					
Percent (95% CI)	27.2% (17.8%, 39.2%)	48.8% (36.4%, 61.4%)	Percent (95% CI)	36.0% (23.8%, 50.3%)	43.3% (30.1%, 57.5%)
Odds ratio (95% CI)		2.55 (1.44, 4.53)	Odds ratio (95% CI)		1.36 (0.84, 2.20)
P-value		.001	P-value		0.213
Anti-itch medication at baseline = yes					
n	78	72	n	85	87
Observed Week 12 ≥3-point NRS improvement [1] - n(%)					
Yes	20 (30.3%)	29 (50.0%)	Yes	26 (35.1%)	35 (53.0%)
No	46 (69.7%)	29 (50.0%)	No	48 (64.9%)	31 (47.0%)

Missing	12	14	Missing	11	21
LS means estimate of percent with improvement [2]					
Percent (95% CI)	29.4% (18.8%, 42.8%)	53.2% (39.2%, 66.6%)	Percent (95% CI)	45.9% (32.4%, 60.1%)	64.6% (49.3%, 77.4%)
Odds ratio (95% CI)		2.73 (1.35, 5.51)	Odds ratio (95% CI)		2.15 (1.09, 4.25)
P-value		.005	P-value		0.028
Medical conditions at baseline = no					
n	161	164	n	199	195
Observed Week 12 ≥3-point NRS improvement [1] - n(%)					
Yes	46 (32.5%)	67 (50.4%)	Yes	66 (37.7%)	79 (49.4%)
No	95 (67.4%)	66 (49.6%)	No	109 (62.3%)	81 (50.6%)
Missing	20	31	Missing	24	35
LS means estimate of percent with improvement [2]					
Percent (95% CI)	31.8% (24.6%, 40.0%)	50.8% (42.6%, 58.9%)	Percent (95% CI)	41.2% (32.6%, 50.4%)	51.6% (42.2%, 60.8%)
Odds ratio (95% CI)		2.21 (1.38, 3.55)	Odds ratio (95% CI)		1.52 (0.99, 2.32)
P-value		.001	P-value		0.054
Medical conditions at baseline = yes					
n	28	25	n	37	42
Observed Week 12 ≥3-point NRS improvement [1] - n(%)			Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	5 (20.8%)	15 (62.5%)	Yes	11 (34.4%)	16 (51.6%)

No	19 (79.2%)	9 (37.5%)	No	21 (65.6%)	15 (48.4%)
Missing	4	1	Missing	5	11
LS means estimate of percent with improvement [2]					
Percent (95% CI)	17.6% (7.4%, 36.3%)	63.7% (43.2%, 80.2%)	Percent (95% CI)	42.6% (26.6%, 60.3%)	55.9% (37.8%, 72.6%)
Odds ratio (95% CI)		8.20 (2.24, 29.99)	Odds ratio (95% CI)		1.71 (0.67, 4.36)
P-value		.001	P-value		0.259

KALM-1

When the primary analysis was conducted separately for interim analysis subjects and post-interim analysis subjects, the results were consistent with the combined analysis presented in Table 20 of this submission. In interim analysis subjects, the odds ratio for achieving a ≥3-point improvement from baseline in the WI-NRS at Week 12 with difelikefalin versus placebo was 3.31 (95% CI, 1.67 to 6.57; P <.001); in post-interim subjects, the odds ratio was 2.20 (95% CI, 1.21 to 3.99; P =.009).

The primary efficacy analysis was conducted separately for interim analysis and post-interim analysis subjects, and for stratification factors. These stratification factors whether a patient used anti-itch medication at baseline, and the presence or absence of certain medical conditions at baseline. These specific medical conditions include:

- History of fall or fracture (related to fall)
- Confusional state or mental status change or altered mental status or disorientation
- Gait disturbance or movement disorder

Table 73 below summarises the proportion of ITT subjects with a ≥3-point improvement in the mean 24-hour WI-NRS scores from baseline at Week 12 by stratification factor (use of anti-itch medications at baseline and presence of certain

medical conditions at baseline). The difelikefalin group showed a statistically significant greater percentage of subjects achieving a ≥3-point improvement from baseline in WI-NRS scores at Week 12 regardless of stratification factor: p =.001 for no use of anti-itch medications at baseline; p =.005 for use of anti-itch medications at baseline; p =.001 for no medical conditions at baseline; and p =.001 for medical conditions at baseline.

Table 73 Summary of Week 12 ≥3-Point Worst Itching Intensity NRS improvement by stratification factors in KALM-1 study

	Placebo	Difelikefalin
Anti-itch medication at baseline = no		
n	111	117
Observed Week 12 ≥3-point NRS improvement [1] — (%)		
Yes	31 (31.3%)	53 (53.5%)
No	68 (68.7%)	46 (46.5%)
Missing	12	18
LS means estimate of percent with improvement [2]		
Percent (95% CI)	27.2% (17.8%, 39.2%)	48.8% (36.4%, 61.4%)
Odds ratio (95% CI)		2.55 (1.44, 4.53)
P-value		.001
Anti-itch medication at baseline = yes		
n	78	72
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	20 (30.3%)	29 (50.0%)
No	46 (69.7%)	29 (50.0%)
Missing	12	14
LS means estimate of percent with improvement [2]		
Percent (95% CI)	29.4% (18.8%, 42.8%)	53.2% (39.2%, 66.6%)

	Placebo	Difelikefalin
Odds ratio (95% CI)		2.73 (1.35, 5.51)
P-value		.005
Medical conditions at baseline = no		
n	161	164
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	46 (32.5%)	67 (50.4%)
No	95 (67.4%)	66 (49.6%)
Missing	20	31
LS means estimate of percent with improvement [2]		
Percent (95% CI)	31.8% (24.6%, 40.0%)	50.8% (42.6%, 58.9%)
Odds ratio (95% CI)		2.21 (1.38, 3.55)
P-value		.001
Medical conditions at baseline = yes		
n	28	25
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	5 (20.8%)	15 (62.5%)
No	19 (79.2%)	9 (37.5%)
Missing	4	1
LS means estimate of percent with improvement [2]		
Percent (95% CI)	17.6% (7.4%, 36.3%)	63.7% (43.2%, 80.2%)
Odds ratio (95% CI)		8.20 (2.24, 29.99)
P-value		.001

Based on results from Phase 2 studies, it was anticipated that there would be a limited number of subjects from most study sites. Therefore, the randomisation was centralised and not stratified by centre. This approach helped achieve a balance between the difelikefalin and placebo treatment groups with respect to stratification factors and demographic and baseline characteristics across study sites, but not necessarily within study sites. In accordance with Section 3.2 of ICH E-9 (Statistical Principles for Clinical Trials), the study site was not included as a variable in the statistical models used to analyse the efficacy endpoints. Nevertheless, for completeness, the Week 12 change in WI-NRS score from baseline was summarised using descriptive statistics by study site, along with the counts and proportions (out of the ITT population at that site) of subjects achieving a ≥3-point improvement from baseline by site (for sites that had at least two subjects in each treatment arm with data at Week 12). For the majority of sites (22 of the 29 sites providing data), the mean change (reduction) from baseline in WI-NRS score at Week 12 was greater in the difelikefalin group than in the placebo group; for 13 of the sites, the mean change was at least twice that in the placebo group. Similarly, for the majority of sites (20 of the 29 sites providing data), a greater percentage of subjects in the difelikefalin group than in the placebo group achieved a ≥3 point improvement in WI-NRS score at Week 12; for 13 of the sites, the percentage was at least twice that in the placebo group.

It should be noted that the above treatment group comparisons were limited by the small number of subjects at some study sites.

KALM-2

Analyses of the primary efficacy endpoint were conducted separately for interim analysis subjects and post-interim analysis subjects, and are consistent with the combined analysis presented above. For interim analysis subjects, the odds ratio for achieving a ≥3-point improvement from baseline in the WI-NRS at Week 12 with difelikefalin versus placebo was 1.88 (95% CI, 0.97 to 3.65); in post-interim subjects, the odds ratio was 1.42 (95% CI, 0.88 to 2.30).

Table 74 summarises the proportion of ITT subjects achieving a ≥3-point improvement in the mean 24-hour WI-NRS scores from baseline at Week 12 by

stratification variables (use of anti-itch medications at baseline and presence of certain medical conditions at baseline).

Results were similar regardless of the presence of certain medical conditions at baseline (odds ratio of 1.71 and 1.52 for “yes” and “no”, respectively). The subjects using anti-itch medications at baseline had a greater treatment difference (odds ratio = 2.15; 95% CI, 1.09 to 4.25) favouring difelikefalin than subjects not using anti-itch medications at baseline (odds ratio = 1.36; 95% CI, 0.84 to 2.20).

Table 74 Summary of Week 12 ≥3-Point Worst Itching Intensity NRS improvement by stratification factors in KALM-2 study

	Placebo	Difelikefalin
Anti-itch medication at baseline = no		
n	151	150
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	51 (38.3%)	60 (48.0%)
No	82 (61.7%)	65 (52.0%)
Missing	18	25
LS means estimate of percent with improvement [2]		
Percent (95% CI)	36.0% (23.8%, 50.3%)	43.3% (30.1%, 57.5%)
Odds ratio (95% CI)		1.36 (0.84, 2.20)
P-value		0.213
Anti-itch medication at baseline = yes		
n	85	87
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	26 (35.1%)	35 (53.0%)
No	48 (64.9%)	31 (47.0%)
Missing	11	21

	Placebo	Difelikefalin
LS means estimate of percent with improvement [2]		
Percent (95% CI)	45.9% (32.4%, 60.1%)	64.6% (49.3%, 77.4%)
Odds ratio (95% CI)		2.15 (1.09, 4.25)
P-value		0.028
Medical conditions at baseline = no		
n	199	195
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	66 (37.7%)	79 (49.4%)
No	109 (62.3%)	81 (50.6%)
Missing	24	35
LS means estimate of percent with improvement [2]		
Percent (95% CI)	41.2% (32.6%, 50.4%)	51.6% (42.2%, 60.8%)
Odds ratio (95% CI)		1.52 (0.99, 2.32)
P-value		0.054
Medical conditions at baseline = yes		
n	37	42
Observed Week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	11 (34.4%)	16 (51.6%)
No	21 (65.6%)	15 (48.4%)
Missing	5	11
LS means estimate of percent with improvement [2]		
Percent (95% CI)	42.6% (26.6%, 60.3%)	55.9% (37.8%, 72.6%)

	Placebo	Difelikefalin
Odds ratio (95% CI)		1.71 (0.67, 4.36)
P-value		0.259

The proportion of ITT subjects achieving a ≥3- and ≥4-point improvement in mean 24-hour WI-NRS from baseline at Week 12 by region (USA, Asia, Eastern Europe, or Western Europe) was also assessed. The proportion of 3-point and 4-point responders was larger in the difelikefalin group compared to placebo across all regions. Treatment differences between difelikefalin and placebo were similar in the US and Western Europe (odds ratios of 1.25 and 1.30, respectively, for ≥3-point improvement and odds ratios of 1.48 and 1.21, respectively, for ≥4-point improvement). The difference between difelikefalin and placebo was generally larger in Eastern Europe than in other regions (odds ratio of 3.06; 95% CI 1.38 to 6.80 for ≥3-point improvement and odds ratio of 2.80; 95% CI 1.21 to 6.46 for ≥4-point improvement).

The number of subjects randomised to Asian countries was small (n = 20). The point estimates for the odds ratio in Asia varied depending on the endpoint (odds ratio of 1.90; 95% CI 0.21 to 17.07 for ≥3-point improvement and odds ratio of 5.42; 95% CI 0.13 to 226.01 for ≥4-point improvement).

Finally, the proportion of ITT subjects achieving a ≥3- and ≥4-point improvement in mean 24-hour WI-NRS from baseline at Week 12 was analysed by dialysis type (haemodialysis or haemodiafiltration). Results were numerically similar regardless of dialysis type for ≥3-point improvement (odds ratio of 1.55 and 1.82 for haemodialysis and haemodiafiltration, respectively) and for ≥4-point improvement (odds ratio of 1.70 and 1.72 for haemodialysis and haemodiafiltration, respectively).

Based on results from Phase 2 studies, it was anticipated that there would be a limited number of subjects from most study sites. Therefore, the randomisation was centralised and not stratified by centre. This approach helped achieve a balance between the difelikefalin and placebo treatment groups with respect to stratification factors, and demographic and baseline characteristics across study sites, but not

necessarily within study sites. In accordance with Section 3.2 of ICH E-9 (Statistical Principles for Clinical Trials), study site was not included as a variable in the statistical models used to analyse the efficacy endpoints. Nevertheless, for completeness, the Week 12 change in WI-NRS score from baseline was summarised using descriptive statistics by study site, along with counts and proportions (out of the ITT population at that site) of subjects achieving a ≥3-point improvement from baseline by site (for sites that had at least two subjects in each treatment arm with data at Week 12).

It should be noted that the treatment group comparisons within each centre were limited by the small number of subjects at some study sites.

All other analyses of efficacy were based on the data from all study sites.

Appendix F: Adverse reactions

No additional adverse reactions to report.

Appendix G: Published cost-effectiveness studies

Please see document ‘Appendix D, G, H, I – SLR Results’

Appendix H: Health-related quality of life studies

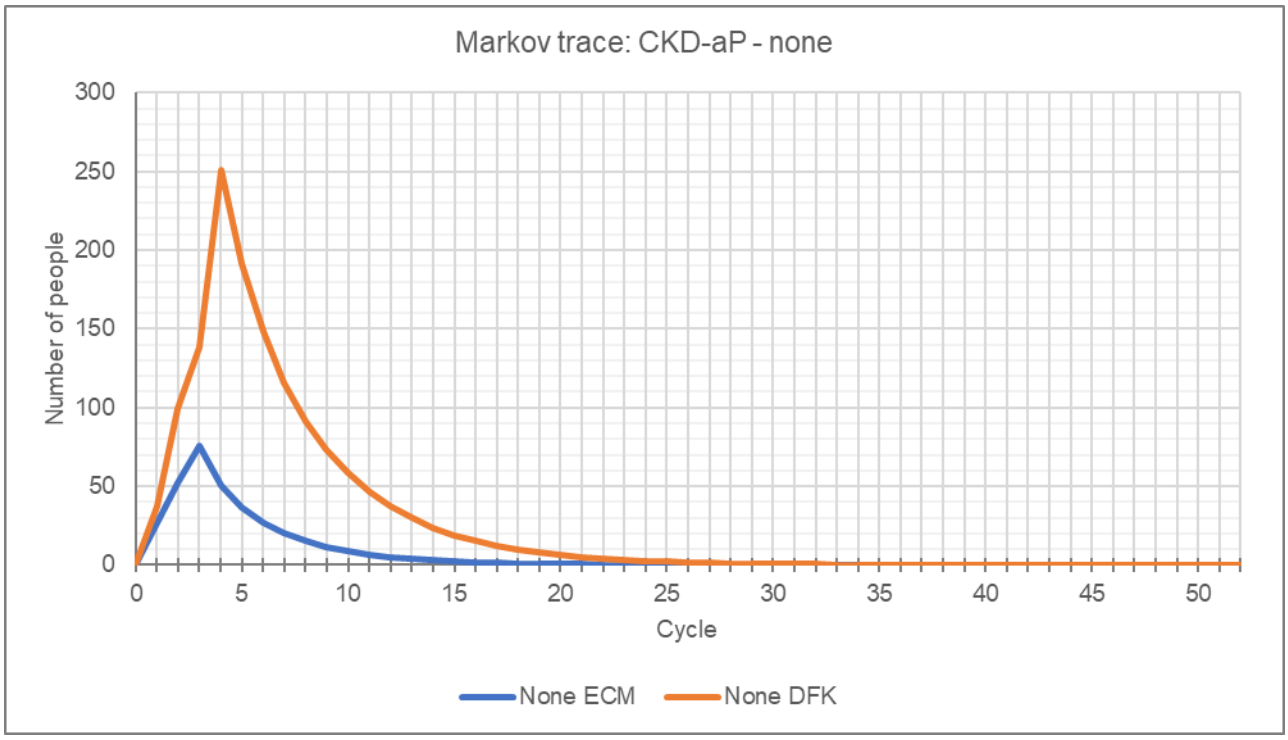
Please see document ‘Appendix D, G, H, I – SLR Results’

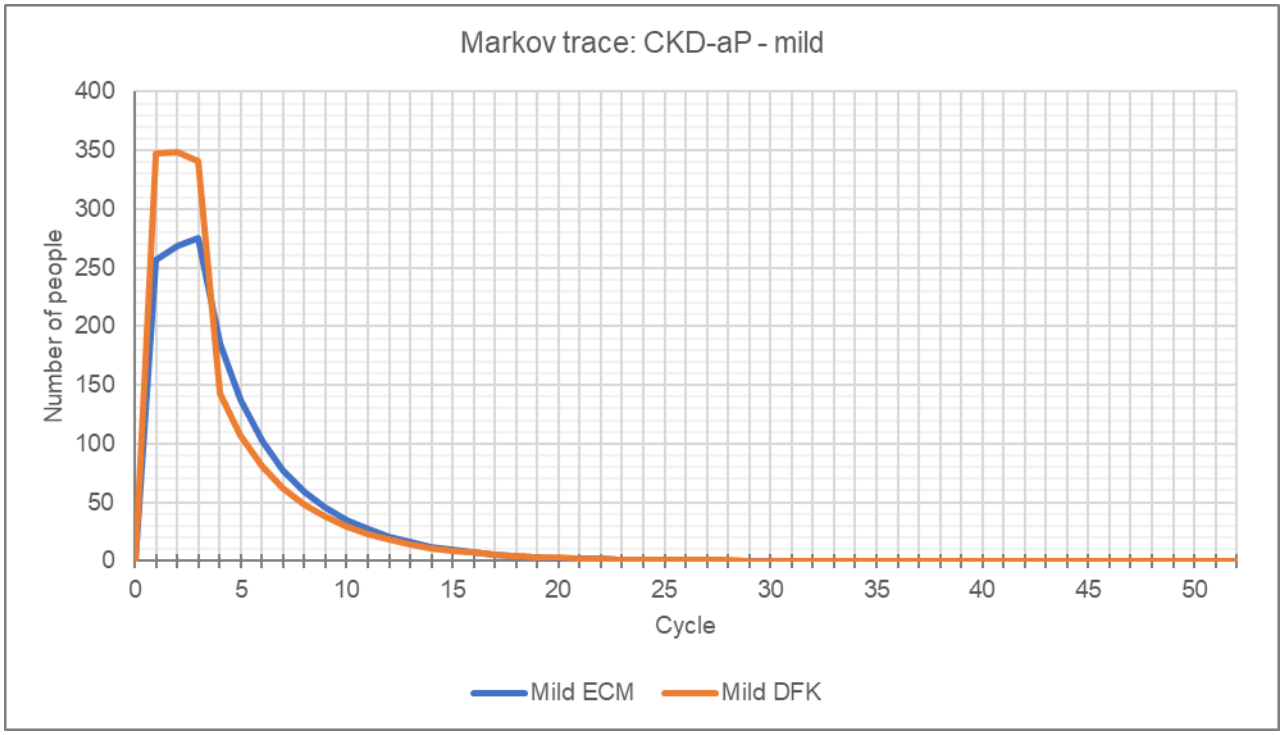
Appendix I: Cost and healthcare resource identification, measurement and valuation

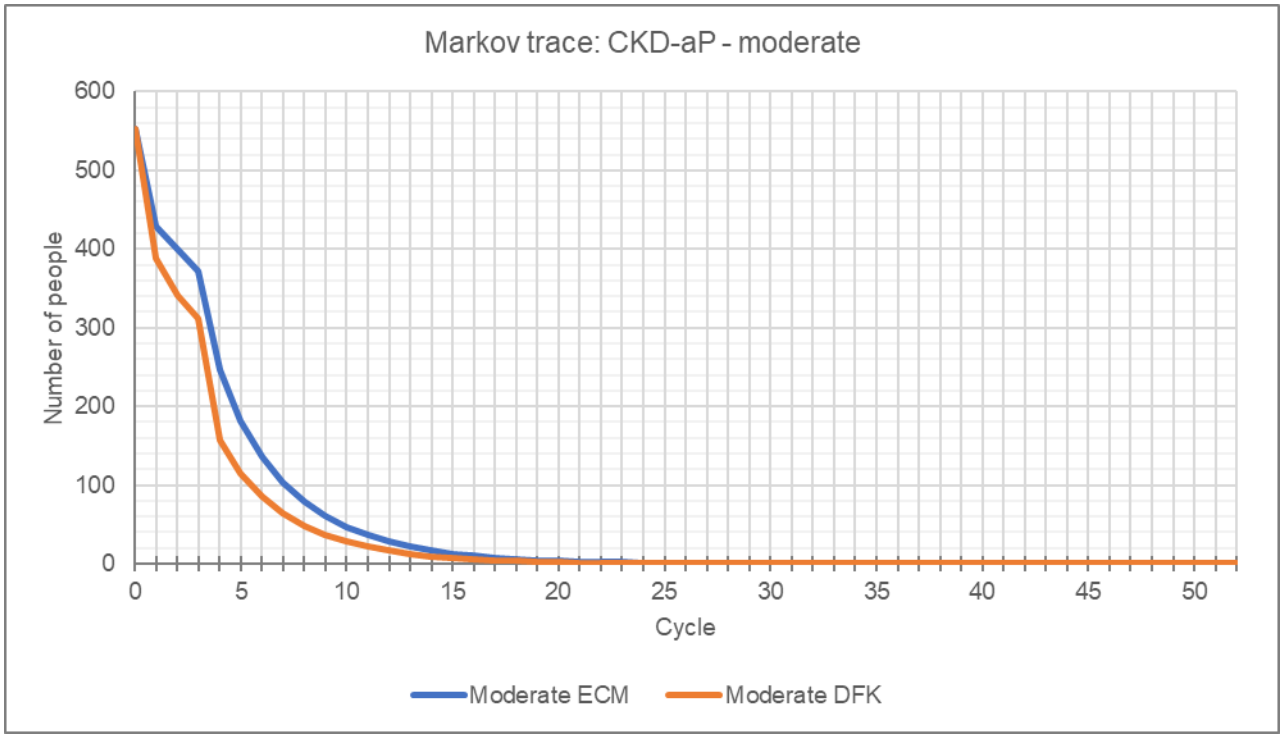
Please see document ‘Appendix D, G, H, I – SLR Results’

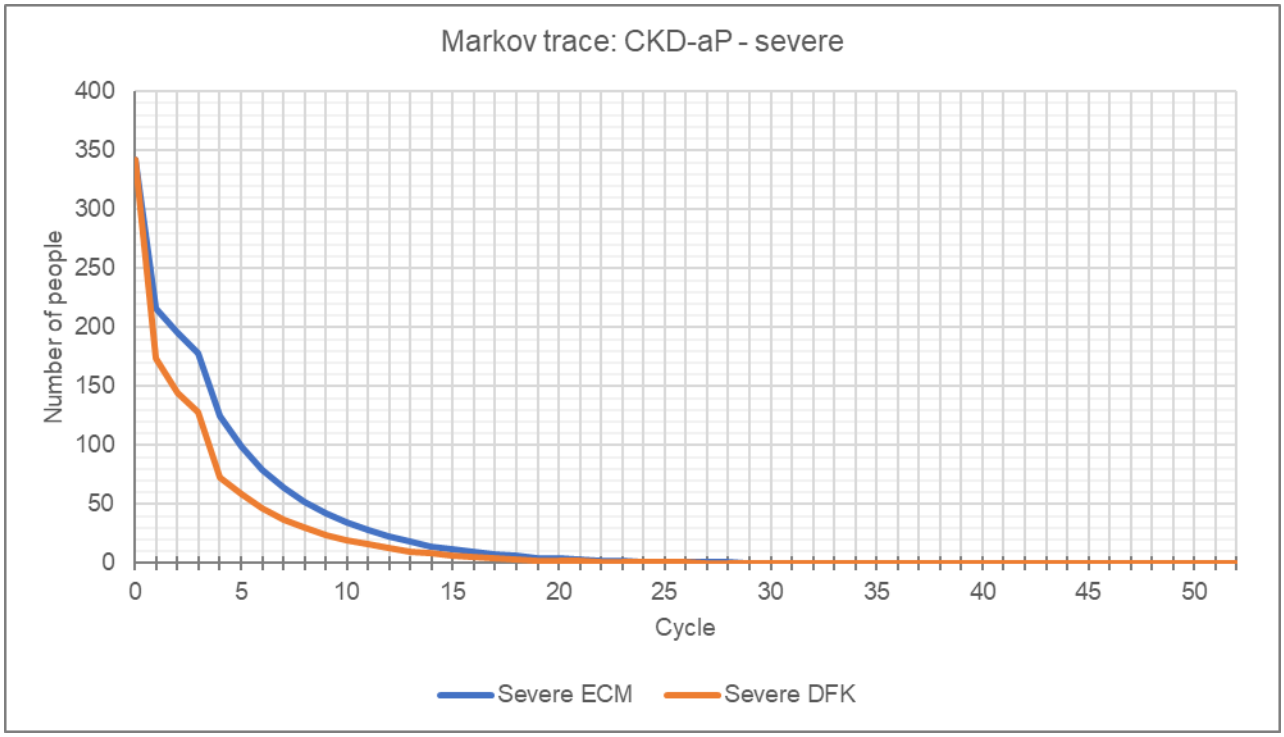
Appendix J: Clinical outcomes and disaggregated results from the model

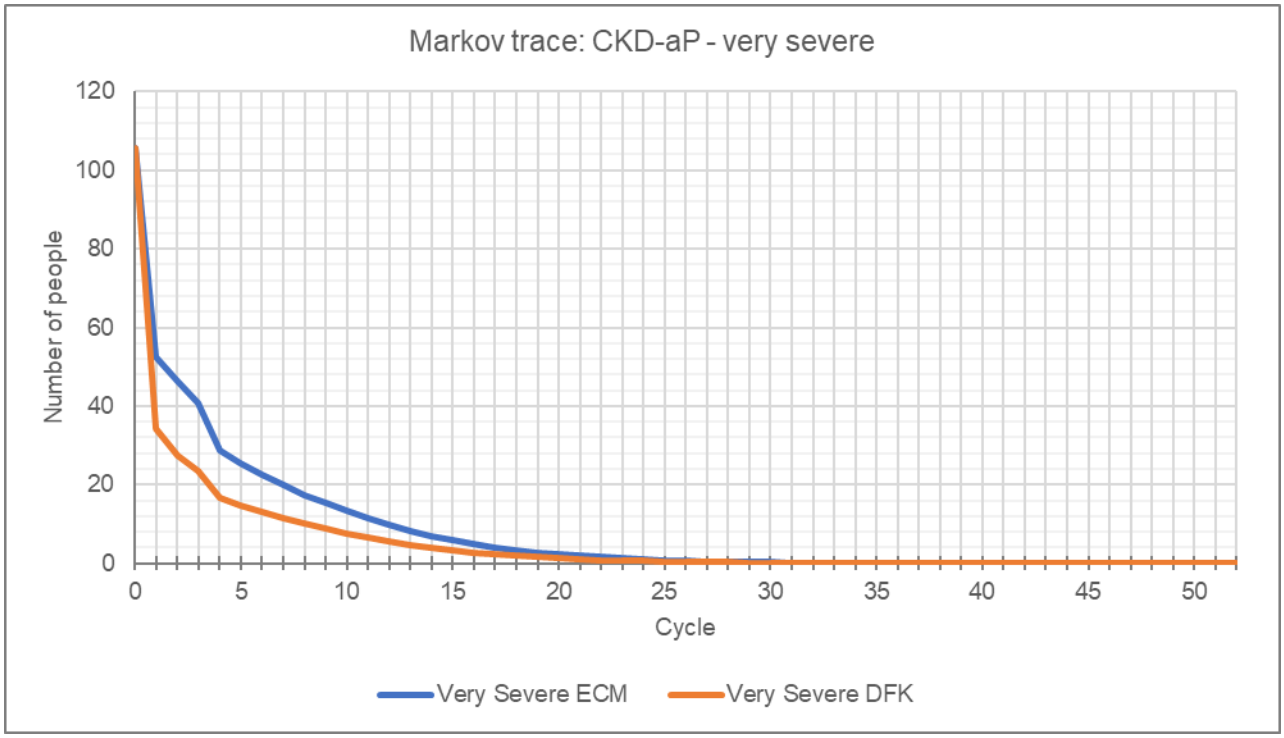
J1.1 Markov traces











J1.2 Mapping study

Please see attached: CKD-ap mapping paper_v3 22.12.21

Note: study is academic in confidence

Appendix K: Price details of treatments included in the submission

Relevant information provided in main body of submission.

Appendix L: Checklist of confidential information

Please see document ‘Difelikefalin confidentiality checklist’

Appendix M: All outcomes’ measures

KALM-1

Primary efficacy:

- Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period.

Secondary efficacy:

- Change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the 5-D Itch scale
- Change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the total Skindex-10 scale score
- Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period

Other efficacy:

- Proportion of subjects achieving a >0-, ≥1-, ≥2-, ≥3-, ≥4-, ≥5-, and ≥6-point improvement from baseline with respect to weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period

- Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period by stratification variables
- Change from baseline in the weekly mean of the 24-hour WI-NRS score at each week of the double-blind treatment period
- Proportion of subjects with “Very Much Improved” or “Much Improved” on the PGIC at Week 12 of the double-blind treatment period
- Proportion of subjects who were complete responders based on the WI-NRS
- Change from baseline in total Skindex-10 scale score at each week
- Change from baseline in each of the three Skindex-10 scale domain scores at each week
- Proportion of subjects with a ≥15-point improvement from baseline in total Skindex-10 scale score at each week
- Change from baseline in the total 5-D Itch scale score at each week
- Change from baseline in each of the five 5-D Itch scale domain scores at each week
- Proportion of subjects achieving a ≥5-point improvement from baseline in total 5-D Itch scale at each week

Safety:

The following assessments were used to evaluate the safety of difelikefalin in subjects

undergoing haemodialysis and experiencing moderate-to-severe pruritus:

- AEs
- Clinical laboratory test results
- Vital sign measurements
- 12-lead ECG results

KALM-1 OLE

Efficacy:

The maintenance of the effect of difelikefalin on itch was measured by a PRO, the 5-D Itch scale, with which data had already been recorded during the double-blind treatment period. The 5-D Itch scale was completed by subjects periodically during the OLE phase and was used to evaluate the effect of difelikefalin, focusing on the change in total score and change by domain score from baseline.

Safety:

The following assessments were used to evaluate the safety of difelikefalin in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus:

- AEs
- Clinical laboratory test results
- Vital sign measurements
- 12-lead ECG results

Adverse events:

An AE was defined as any untoward medical occurrence in a patient or clinical investigational subject administered a medicinal product. An AE did not necessarily have to have a causal relationship with the treatment.

An SAE was defined as any untoward medical occurrence that:

- Resulted in death;
- Was life-threatening; NOTE: The term “life-threatening” in the definition of “serious” referred to an event in which the subject was at risk of death at the time of the event; it did not refer to an event that hypothetically might have caused death if it had been more severe.
- Required inpatient hospitalisation or prolongation of existing hospitalisation;
- Resulted in persistent or significant disability/incapacity;
- Was a congenital anomaly/birth defect; or

- Was an important medical event that may have jeopardised the subject or required intervention to prevent one of the other outcomes listed above

In addition, AEs of special interest also were monitored through standard AE reporting. Specific custom Medical Dictionary for Regulatory Activities (MedDRA) Preferred Terms (PTs) were sued to define these events.

Clinical laboratory test results:

Blood samples for clinical laboratory tests, including haematology, serum chemistry, and serum pregnancy, were taken prior to haemodialysis and analysed by a central laboratory.

Vital sign measurements:

Vital sign measurements included body temperature, heart rate (in sitting or semi-recumbent position), and systolic and diastolic blood pressure (in sitting or semi-recumbent position).

12-lead ECG results:

A 12-lead ECG was obtained prior to the start of the first haemodialysis of Week 53 (or at early termination/EOT, as applicable). The ECG was read locally by the investigator or qualified designee (endorsed by the investigator). Clinically significant abnormalities or worsening of findings after the first dose of the study drug were reported as TEAEs.

KALM-2

Primary efficacy:

- Subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period.

Secondary efficacy

- Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period

- Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 8 of the double-blind treatment period
- Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 4 of the double-blind treatment period
- Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 8 of the double-blind treatment period
- Proportion of subjects achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 4 of the double-blind treatment period
- Change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the total Skindex-10 scale score
- Change from baseline in itch-related QoL at the end of Week 12 of the double-blind treatment period, as assessed by the 5-D Itch scale score

Other efficacy endpoints:

- Proportion of subjects achieving a >0-, ≥1-, ≥2-, ≥3-, ≥4-, ≥5-, and ≥6-point improvement from baseline with respect to weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period
- Proportion of subjects achieving a ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 of the double-blind treatment period by stratification variables
- Proportion of subjects achieving a ≥3 and ≥4-point improvement from baseline in WI-NRS at Week 12 of the double-blind treatment period by region and dialysis type (each individually)
- Change from baseline in the weekly mean of the 24-hour WI-NRS score at each week of the double-blind treatment period

- Proportion of subjects with “Very much improved” or “Much improved” on the PGIC at Week 12 of the double-blind treatment period
- Proportion of subjects who were complete responders based on the WI-NRS
- Change from baseline in total Skindex-10 scale score at each week
- Change from baseline in each of the three Skindex-10 scale domain scores at each week
- Proportion of subjects with a ≥15-point improvement from baseline in total Skindex-10 scale score at each week
- Change from baseline in the total 5-D Itch scale score at each week
- Change from baseline in each of the five 5-D Itch scale domain scores at each week
- Proportion of subjects achieving a ≥5-point improvement from baseline in total 5-D Itch scale at each week

Safety endpoints:

The following assessments were used to evaluate the safety of difelikefalin in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus:

- AEs
- Clinical laboratory test results
- Vital sign measurements
- 12-lead ECG results

KALM-2 OLE

See description under ‘KALM-1 OLE’ in Appendix M.

CLIN3105

Efficacy:

No primary effectiveness endpoint was defined for this study. The following effectiveness endpoints were used:

Worst Itching Intensity NRS:

- Change from baseline in the weekly mean of the 24-hour WI-NRS score to Week 12.
- Percentage of subjects achieving >0, ≥1, ≥2, ≥3, ≥4, ≥5, and ≥6-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at Week 12 (with ≥3 and ≥4-point improvement also reported by region).

Sleep Quality Questionnaire:

- Change from baseline in the weekly mean of the 24-hour Sleep Quality Score to Week 12.
- Percentage of subjects achieving >0, ≥1, ≥2, ≥3, ≥4, ≥5, and ≥6-point improvement from baseline with respect to the weekly mean of the 24-hour Sleep Quality Score at Week 12 (with ≥3 and ≥4-point improvement also reported by region).

5-D Itch scale and Skindex-10 scale:

- Change from baseline in itch-related QoL to Week 12 as assessed by the 5-D Itch scale total score and each 5-D Itch scale domain score.
- Change from baseline in itch-related QoL to Week 12 as assessed by the Skindex-10 scale total score and each Skindex-10 subdomain score.

EQ-5D-5L:

- Percentage of subjects with reported problems by level (1 to 5) and EQ-5D-5L dimension at baseline and Week 12.
- Percentage of subject with no problems (i.e., with a level 1 response) by EQ-5D-5L dimension at baseline and Week 12.
- Descriptive statistics (mean, standard deviation [SD], median, and range) for observed Overall Self-Rated Health Status EQ VAS at baseline, Week 12, and the change from baseline at Week 12.
- Each subject's health state (listing) expressed using the 5 EQ-5D-5L dimensions.

- Descriptive statistics for QALY weights, using the US value set, for baseline, Week 12, and the change from baseline at Week 12.

EQ-PSO:

- Percentage of subjects with reported problems by level (1 to 5) and EQ-PSO dimension at baseline and Week 12.
- Percentage of subjects with no itching/no problems (i.e., with a level 1 response) by EQ-PSO dimension at baseline and Week 12.

Safety:

The following assessments were used to evaluate the safety of difelikefalin in subjects undergoing haemodialysis and experiencing moderate-to-severe pruritus:

- AEs
- Vital signs
- Electrocardiograms
- Clinical laboratory values

Appendix N: Clinical opinion and consensus report

Please see document 'Appendix N - CKD-aP Clinician opinion and consensus report (May 2022)'

Appendix O: Full lists of inclusion and exclusion criteria

KALM-1

Inclusion criteria:

To be eligible for inclusion in the double-blind phase of the study, a subject had to meet the following criteria:

1. Was willing and able to provide written informed consent prior to participating in the study

2. Was able to communicate clearly with the investigator and staff, able to understand the study procedures, and able and willing to comply with the study requirements

3. Was 18 years of age or older

4. Had ESRD and had been undergoing haemodialysis 3 times per week for at least 3 months prior to the start of screening

Note 1: Subjects who required an occasional additional haemodialysis treatment to manage fluid overload could be enrolled as long as it was anticipated that no more than one such treatment would be required in any given week.

Note 2: Subjects undergoing in-home dialysis could participate as long as they had switched to in-centre haemodialysis at least 2 weeks prior to screening, and planned to remain on in-centre haemodialysis for the duration of the study.

5. If female, was not pregnant or nursing during any period of the study

6. If female:

a. Was surgically sterile; or

b. Had been amenorrhoeic for at least 1 year and was over the age of 55 years; or

c. Had a negative serum pregnancy test at screening and agreed to use acceptable contraceptive measures (e.g., hormonal contraceptives, barrier with spermicide, intrauterine device, vasectomised partner, or abstinence) from the time of informed consent until 7 days after the last dose of study drug.

7. If male, had agreed not to donate sperm after the first dose of study drug until 7 days after the last dose of study drug, and had agreed to use a condom with spermicide or abstain from heterosexual intercourse during the study until 7 days after the last dose of study drug

Note: No restrictions were required for a vasectomised male, provided his vasectomy was performed ≥4 months prior to screening.

8. Had a prescription dry body weight between 40.0 and 135.0 kg, inclusive

9. Had at least two single-pool measurements ≥1.2 for [dialyser clearance of urea × dialysis time] / volume of distribution of urea (or Kt/V) or at least two urea reduction

ratio measurements ≥65%, or one single-pool Kt/V measurement ≥1.2 and 1 urea reduction ratio measurement ≥65% on different dialysis days during the 3-month period prior to screening

10. Prior to randomisation:

a. Had completed at least 4 WI-NRS worksheets from the start of the 7-day run-in period, up to and including the pre-randomisation assessment on Day 1

b. Had a mean baseline WI-NRS score >4, defined as the average of all non-missing scores reported from the start of the 7-day run-in period, up to and including the pre-randomisation assessment on Day 1

Exclusion criteria:

A subject was excluded from the double-blind phase of the study if any of the following criteria were met:

1. Had known noncompliance with dialysis treatment that, in the opinion of the investigator, would have impeded completion or validity of the study

2. Was scheduled to receive a kidney transplant during the study

3. Had a known history of allergic reaction to opiates, such as hives

Note: Side effects related to the use of opioids, such as constipation or nausea, did not exclude subjects from the study.

4. Had a concomitant disease or a history of any medical condition that, in the opinion of the investigator, could have posed undue risk to the subject, could have impeded completion of the study procedures, or would have compromised the validity of the study measurements. These conditions included but were not limited to:

a. Known or suspected history of alcohol, narcotic, or other drug abuse, or substance dependence within 12 months prior to screening

b. Significant systolic or diastolic heart failure (eg, New York Heart Association Class IV congestive heart failure)

c. Severe mental illness or cognitive impairment (eg, dementia)

d. Any other relevant acute or chronic medical or neuropsychiatric condition within the 3 months prior to screening (eg, diagnosis of encephalopathy, coma, delirium)

5. Had received new or changed treatment for itch, including antihistamines and corticosteroids (oral, IV, or topical), within 14 days prior to screening

6. Had received new or changed prescription for opioids, gabapentin, or pregabalin within 14 days prior to screening

7. Had received another investigational drug within 30 days prior to the start of screening or was planning to participate in another clinical study while enrolled in this study

8. In the opinion of the investigator, had pruritus attributed to a cause other than ESRD) or its complications (eg, subjects with concomitant pruritic dermatological disease or cholestatic liver disease)

Note: Subjects whose pruritus was attributed to ESRD complications, such as hyperparathyroidism, hyperphosphatemia, anaemia, or the dialysis procedure or prescription, could be enrolled.

9. Had localised itch restricted to the palms of the hands

10. Had pruritus only during the haemodialysis session (by subject report)

11. Was receiving ongoing ultraviolet B treatment and anticipated receiving such treatment during the study

12. Had participated in a previous clinical study with difelikefalin

KALM-1 OLE

Inclusion criteria

To be eligible for inclusion into the OLE phase of the study, each subject had to fulfil the additional following criteria at the time of entry into the OLE phase:

1. Had received at least 30 doses of the planned 36 doses of study drug during the double-blind phase of the study

2. Had a prescription dry body weight ≥40 kg

3. Continued to meet inclusion criteria 1 through 7 of the double-blind phase

Exclusion criteria

A subject was excluded from the OLE phase of the study if any of the additional following criteria were met at the time of entry into the OLE phase:

1. Completed the double-blind phase of this study but exhibited AEs during the course of the treatment period that may have precluded continued exposure to the study drug

2. Was noncompliant with protocol procedures during the double-blind phase of the study, which was indicative of an inability to follow protocol procedures

3. Had developed a concomitant disease or any medical condition that, in the opinion of the investigator, could have posed undue risk to the subject, impeded completion of the study procedures, or compromised the validity of the study measurements

KALM-2

Inclusion criteria:

To be eligible for inclusion in the double-blind phase of the study, a subject had to meet the following criteria:

1. Was willing and able to provide written informed consent prior to participating in the study

2. Was able to communicate clearly with the investigator and staff, able to understand the study procedures, and able and willing to comply with the study requirements, including providing written responses to questionnaires

3. Was between 18 and 85 years of age, inclusive

Note: Subjects in Korea had to be between 19 and 85 years of age, inclusive.

4. Had ESRD and had been undergoing haemodialysis 3 times per week for at least 3 months prior to the start of screening

Note 1: Subjects who required an occasional additional haemodialysis treatment to manage fluid overload could be enrolled as long as it was anticipated that no more than one such treatment would be required in any given week.

Note 2: Subjects undergoing in-home dialysis could participate as long as they had switched to in-centre haemodialysis at least 2 weeks prior to screening, and planned to remain on in-centre haemodialysis for the duration of the study.

Note 3: Subjects receiving alternate dialysis modalities, such as nocturnal dialysis, were not eligible.

5. If female, was not pregnant or nursing during any period of the study

6. If female:

a. Was surgically sterile; or

b. Had been amenorrheic for at least 1 year and was over the age of 55 years; or

c. Had a negative serum pregnancy test at screening and agreed to use acceptable contraceptive measures (eg, hormonal contraceptives, barrier with spermicide, intrauterine device, vasectomised partner, or abstinence) from the time of informed consent until 7 days after the last dose of study drug.

7. If male, had agreed not to donate sperm after the first dose of study drug until 7 days after the last dose of study drug, and had agreed to use a condom with spermicide or abstain from heterosexual intercourse during the study until 7 days after the last dose of study drug

Note: No restrictions were required for a vasectomised male provided his vasectomy was performed ≥4 months prior to screening.

8. Had a prescription dry body weight between 40.0 and 135.0 kg, inclusive

9. Had at least two single-pool measurements ≥1.2 for [dialyser clearance of urea × dialysis time] / volume of distribution of urea (or Kt/V),

Or at least two urea reduction ratio measurements ≥65%,

Or one single-pool Kt/V measurement ≥1.2 and 1 urea reduction ratio measurement ≥65% on different dialysis days during the 3-month period prior to screening

10. Prior to randomisation:

a. Had completed at least 4 WI-NRS worksheets from the start of the 7-day run-in period, up to and including the pre-randomisation assessment on Day 1

b. Had a mean baseline WI-NRS score ≥5, defined as the average of all non-missing scores reported from the start of the 7-day run-in period up to and including the pre-randomisation assessment on Day 1

Exclusion criteria:

A subject was excluded from the double-blind phase of the study if any of the following criteria were met:

1. Had known noncompliance with dialysis treatment that, in the opinion of the investigator, would have impeded completion or validity of the study

2. Was scheduled to receive a kidney transplant during the study

3. Had a known history of allergic reaction to opiates, such as hives

Note: Side effects related to the use of opioids, such as constipation or nausea, did not exclude subjects from the study.

4. Had a concomitant disease or a history of any medical condition that, in the opinion of the investigator, could have posed undue risk to the subject, could have impeded completion of the study procedures, or would have compromised the validity of the study measurements; these conditions included but were not limited to:

a. Known or suspected history of alcohol, narcotic, or other drug abuse, or substance dependence within 12 months prior to screening

b. Significant systolic or diastolic heart failure (e.g., New York Heart Association Class IV congestive heart failure)

c. Severe mental illness or cognitive impairment (e.g., dementia)

d. Any other relevant acute or chronic medical or neuropsychiatric condition within 3 months prior to screening (e.g., diagnosis of encephalopathy, coma, delirium)

5. Had received new or changed treatment for itch, including antihistamines and corticosteroids (oral, IV, or topical), within 14 days prior to screening

6. Had received new or changed prescription for opioids, gabapentin, or pregabalin within 14 days prior to screening

7. Had received another investigational drug within 30 days prior to the start of screening or was planning to participate in another clinical study while enrolled in this study

8. In the opinion of the investigator, had pruritus attributed to a cause other than ESRD or its complications (e.g., subjects with concomitant pruritic dermatological disease or cholestatic liver disease)

Note: Subjects whose pruritus was attributed to ESRD complications, such as hyperparathyroidism, hyperphosphatemia, anaemia, or the dialysis procedure or prescription could be enrolled.

9. Had localised itch restricted to the palms of the hands

10. Had pruritus only during the haemodialysis session (by subject report)

11. Was receiving ongoing ultraviolet B treatment and anticipated receiving such treatment during the study

12. Had participated in a previous clinical study with difelikefalin

KALM-2 OLE

Inclusion criteria:

To be eligible for inclusion into the OLE phase of the study, each subject had to fulfil the additional following criteria at the time of entry into the OLE phase:

1. Had received at least 30 doses of the planned 36 doses of study drug during the double-blind phase of the study

2. Had a prescription dry body weight ≥40 kg

3. Continued to meet inclusion criteria 1 through 7 of the double-blind phase

Exclusion criteria:

A subject was excluded from the OLE phase of the study if any of the additional following criteria were met at the time of entry into the OLE phase:

1. Completed the double-blind phase of the study, but exhibited AEs during the course of the double-blind treatment period that might have precluded continued exposure to the study drug

2. Was noncompliant with protocol procedures during the double-blind phase of the study, which was indicative of an inability to follow protocol procedures

3. Had developed a concomitant disease or any medical condition that, in the opinion of the investigator, could have posed undue risk to the subject, impeded completion of the study procedures, or would have compromised the validity of the study measurements

CLIN3105

Inclusion criteria:

To be eligible for inclusion into the study, a subject had to meet the following criteria:

1. Was willing and able to provide written informed consent prior to participating in the study.

2. Was able to communicate clearly with the investigator and staff, was able to understand the study procedures, and was able and willing to comply with the study schedules and all study requirements.

3. Was between 18 and 85 years of age, inclusive.

4. Had ESRD and had been on haemodialysis 3 times per week for at least 3 months prior to the start of screening.

5. If female, was not pregnant or nursing during any period of the study.

6. If female:

a. Was surgically sterile; or

b. Had been amenorrheic for at least 1 year and was over the age of 55 years; or

c. Had a negative serum pregnancy test at screening and agreed to use acceptable contraceptive measures (eg, hormonal contraceptives, barrier with spermicide, intrauterine device, vasectomised partner, or abstinence from heterosexual intercourse) from the time of informed consent until 7 days after the last dose of study drug.

7. If male, had agreed not to donate sperm after the first dose of study drug until 7 days after the last dose of study drug, and agreed to use a condom with spermicide

or abstain from heterosexual intercourse during the study until 7 days after study drug administration.

8. Had a prescription dry body weight of ≥40.0 kg.

9. Over the last 3 months prior to screening, had at least one of the following:

a. At least two single-pool measurements of (dialyser clearance of urea × dialysis time) / (volume of distribution of urea) of ≥1.2 on different dialysis days.

b. At least two urea reduction ratio measurements of ≥65% on different dialysis days.

c. 1 single-pool measurement of (dialyser clearance of urea × dialysis time) / (volume of distribution of urea) of ≥1.2 and 1 urea reduction ratio measurement of ≥65% on different dialysis days.

10. Prior to treatment:

a. Had completed at least 3 WI-NRS questionnaires from the start of the run-in period, up to and including the assessment on Day 1.

b. Had a mean baseline WI-NRS score of ≥5, defined as the average of all non-missing scores reported from the start of the run-in period up to and including the pre-dose assessment on Day 1.

Exclusion criteria

A subject was excluded from the study if any of the following criteria were met:

1. Had known noncompliance with dialysis treatment that in the opinion of the investigator would have impeded completion or validity of the study.

2. Was scheduled to receive a kidney transplant during the study.

3. Had known history of allergic reaction to opiates, such as hives.

Note: Side effects related to the use of opioids, such as constipation or nausea, would not have excluded subjects from the study.

4. Had hypersensitivity to the active substance or any of the excipients in the investigational products.

5. Had a concomitant disease or a history of any medical condition that, in the opinion of the investigator, could have posed undue risk to the subject, impeded

completion of the study procedures, or compromised the validity of the study measurements, including but not limited to:

a. Known or suspected history of alcohol, narcotic, or other drug abuse, or substance dependence within 12 months prior to screening.

b. Significant systolic or diastolic heart failure (eg, New York Heart Association Class IV congestive heart failure).

c. Severe mental illness or cognitive impairment (eg, dementia).

d. Any other relevant acute or chronic medical or neuropsychiatric condition within 3 months prior to screening (eg, diagnosis of encephalopathy, coma, delirium).

6. Had received new treatment or changed treatment for itch, including antihistamines and corticosteroids (oral, IV, or topical) within 14 days prior to screening.

7. Had received a new prescription or had a change in prescription for opioids, gabapentin, or pregabalin within 14 days prior to screening.

8. Had received another investigational drug within 30 days or five half-lives (whichever was longer) prior to the start of dosing or was planning to participate in another interventional clinical study while enrolled in this study.

9. In the opinion of the investigator, had pruritus attributed to a cause other than ESRD or its complications (eg, subjects with concomitant pruritic dermatological disease or cholestatic liver disease).

Note: Subjects whose pruritus was attributed to ESRD complications, such as hyperparathyroidism, hyperphosphatemia, anaemia, the dialysis procedure, or prescription could be enrolled.

10. Had localised itch restricted to the palms of the hands.

11. Had pruritus only during the dialysis session (by subject report).

12. Was receiving ongoing ultraviolet B treatment and anticipated receiving such treatment during the study.

13. Participated in a previous clinical study with difelikefalin.

Appendix P: Concomitant medications in KALM-1 and KALM-2

Please see documents ‘Appendix P - KALM-1 concomitant medications’ and ‘Appendix P - KALM-2 concomitant medications’

Appendix Q: previous anti-itch medications used in KALM-1 and KALM-2

Please see documents ‘Appendix Q - KALM-1 previous anti-itch medications’ and ‘Appendix Q - KALM-2 previous anti-itch medications’

Appendix R: Opiate withdrawal scale tables

Please see document ‘Appendix R - Opiate withdrawal scale tables’.

Appendix S: ISE outputs by anti-itch medications

Please see document ‘Appendix S - ISE outputs by anti-itch medications’.

Appendix T: KALM-1 and KALM-2 subgroups

Please see document ‘Appendix T - KALM1_KALM2_by subgroups’.

Appendix U: AEs leading to discontinuation

Please see document ‘Appendix U - AEs leading to discontinuation’.

Appendix V: missingness patterns and frequencies

Please see documents:

- ‘Appendix V - KALM-1 - Reason for discontinuation’
- ‘Appendix V - KALM-2 - Reason for discontinuation’
- ‘Appendix V - Table 3c_1_3102’
- ‘Appendix V - Table 3c_1_3102’
- ‘Appendix V - KALM-1 OLE - Reason for discontinuation’

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis

[ID3890]

Clarification questions

25 August 2022

File name	Version	Contains confidential information	Date
		Yes/no	

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Clarification on effectiveness data

Literature searches

A 1. Priority question: The search strategy document reports only 2 sets of searches, one for treatment pathway and a second economic Systematic Literature Review (SLR). Please confirm if any additional clinical effectiveness searches were undertaken to identify randomised control trials (RCTs), observational studies or adverse events and provide full search strategies if appropriate.

A full systematic literature review was not performed for clinical effectiveness searches to identify randomised control trials, observational studies, or adverse events, as it is known that the number of RCTs in this disease area is limited.

Searches for 'CKD-aP' and 'Chronic Kidney Disease associated pruritis' on ClinicalTrials.gov.uk result in 13 trials, of which only 5 are completed (others are recruiting or not yet recruiting).

	Study Title	Status	Conditions	Interventions	URL
1	CKD-aP Among Adults on Dialysis in Switzerland	Not yet recruiting	Chronic Kidney Diseases|Dialysis|Chronic Kidney Diseases	Other: CKD-aP	https://ClinicalTrials.gov/show/NCT05415969

			Associated Pruritus		
2	Klotho in Chronic Kidney Disease-associated Pruritis (CKD-aP)	Unknown status	Chronic Kidney Disease-associated Pruritus	Other: skin biopsy|Radiation: narrowband ultraviolet B	https://ClinicalTrials.gov/show/NCT03532568
3	The MC2-25 Cream in Subjects wITh CHronic KIdNEy Disease-aSsociated pruritus (ITCHINESS) Trial	Recruiting	Chronic Kidney Disease-associated Pruritus	Drug: MC2-25 cream|Drug: MC2-25 vehicle	https://ClinicalTrials.gov/show/NCT05482698
4	Cross-sectional Study to Assess Prevalence and Burden of CKD-associated Pruritus in Haemodialysis Patients	Not yet recruiting	Chronic Kidney Disease-associated Pruritus		https://ClinicalTrials.gov/show/NCT05524467
5	CR845-310302: A Study to Evaluate the Safety and Efficacy of Difelikefalin in Advanced Chronic Kidney Disease Patients With Moderate-to-Severe Pruritus and Not on Dialysis	Not yet recruiting	Chronic Kidney Diseases|Pruritus	Drug: Difelikefalin 1 mg Oral Tablet|Drug: Placebo Oral Tablet	https://ClinicalTrials.gov/show/NCT05356403
6	A Study to Evaluate the Safety and Efficacy of Difelikefalin in Advanced Chronic Kidney Disease Patients With Moderate-to-Severe Pruritus and Not on Dialysis	Not yet recruiting	Chronic Kidney Diseases|Pruritus	Drug: Difelikefalin 1 mg Oral Tablet|Drug: Placebo Oral Tablet	https://ClinicalTrials.gov/show/NCT05342623
7	Intermediate-Size Patient Population Expanded Access Program for Intravenous Difelikefalin	Approved for marketing	Uremic Pruritus	Drug: Difelikefalin	https://ClinicalTrials.gov/show/NCT05031546
8	CR845-CLIN3105: A Study to Evaluate the Safety and Effectiveness of CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg	https://ClinicalTrials.gov/show/NCT03998163
9	CR845-CLIN3103: A Global Study to Evaluate the Safety and Efficacy of CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT03636269
10	A Study to Evaluate the Safety and Efficacy of CR845 in Chronic Kidney Disease Patients With Moderate-to-Severe Pruritus	Completed	Chronic Kidney Diseases|Pruritus	Drug: CR845 0.25 mg Oral Tablet|Drug: CR845 0.5 mg Oral Tablet|Drug: CR845 1 mg Oral Tablet|Drug: Placebo Oral Tablet	https://ClinicalTrials.gov/show/NCT03617536
11	A Study to Evaluate the Safety and Efficacy of CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus (KALM-1)	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT03422653
12	Extension Study to Evaluate IV CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845	https://ClinicalTrials.gov/show/NCT03281538
13	To Evaluate the Efficacy and Safety of Nemolizumab for 12 Weeks in Participants With Chronic Kidney Disease With Associated Moderate to Severe Pruritus	Recruiting	Chronic Kidney Disease Associated Moderate to Severe Pruritus	Drug: Nemolizumab Drug: Placebo	https://ClinicalTrials.gov/show/NCT05075408

An additional search using terms “Difelikefalin” and “CR845” on ClinicalTrials.gov returns 23 studies in total. These are predominately in CKD-aP., and uremic pruritis.

	Title	Status	Conditions	Interventions	URL
1	Study to Investigate the Effects of Single Intravenous Doses of Difelikefalin (CR845) on the QTc Interval in Healthy Subjects	Completed	Healthy	Drug: CR845 0.5 mcg/kg IV|Drug: CR845 3 mcg/kg IV|Drug: Moxifloxacin 400 mg Oral Tablet|Other: Placebo	https://ClinicalTrials.gov/show/NCT04019574
2	Intermediate-Size Patient Population Expanded Access Program for Intravenous Difelikefalin	Approved for marketing	Uremic Pruritus	Drug: Difelikefalin	https://ClinicalTrials.gov/show/NCT05031546
3	A Study to Evaluate the Safety and Efficacy of Difelikefalin in Advanced Chronic Kidney Disease Patients With Moderate-to-Severe Pruritus and Not on Dialysis	Not yet recruiting	Chronic Kidney Diseases|Pruritus	Drug: Difelikefalin 1 mg Oral Tablet|Drug: Placebo Oral Tablet	https://ClinicalTrials.gov/show/NCT05342623
4	Study to Evaluate the Efficacy and Safety of Oral Difelikefalin (CR845) for Moderate to Severe Pruritus in Subjects With Atopic Dermatitis	Completed	Pruritus|Atopic Dermatitis	Drug: difelikefalin 0.25 mg|Drug: difelikefalin 0.5 mg|Drug: difelikefalin 1.0 mg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT04018027
5	Study to Evaluate the Efficacy and Safety of Oral Difelikefalin (CR845) for Moderate to Severe Pruritus in Subjects With Notalgia Paresthetica (KOMFORT)	Completed	Pruritus|Notalgia Paresthetica	Drug: difelikefalin 2.0 mg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT04706975
6	Study to Evaluate the Efficacy and Safety of Oral Difelikefalin as Adjunct Therapy to a Topical Corticosteroid for Moderate to Severe Pruritus in Subjects With	Not yet recruiting	Pruritus|Atopic Dermatitis	Drug: difelikefalin 0.25 mg|Drug: difelikefalin 0.5 mg|Drug: TCS Cream|Drug: Placebo|Drug: Vehicle Cream	https://ClinicalTrials.gov/show/NCT05387707

	Atopic Dermatitis				
7	CR845-100303: Study to Assess the Potential of Physical Withdrawal From Intravenous CR845 (Difelikefalin) in Hemodialysis Patients	Completed	Hemodialysis	Drug: CR845 0.5 mcg/kg|Other: Placebo	https://ClinicalTrials.gov/show/NCT05533008
8	CR845-310302: A Study to Evaluate the Safety and Efficacy of Difelikefalin in Advanced Chronic Kidney Disease Patients With Moderate-to-Severe Pruritus and Not on Dialysis	Not yet recruiting	Chronic Kidney Diseases|Pruritus	Drug: Difelikefalin 1 mg Oral Tablet|Drug: Placebo Oral Tablet	https://ClinicalTrials.gov/show/NCT05356403
9	Study to Evaluate the Pharmacokinetics and Metabolism of [14C] CR845 (Difelikefalin) in Patients With End Stage Renal Disease on Hemodialysis and in Healthy Subjects	Completed	Hemodialysis|Healthy	Drug: [14C] CR845	https://ClinicalTrials.gov/show/NCT03947970
10	Study to Evaluate the Safety and Efficacy of Oral CR845 (Difelikefalin) in Patients With Primary Biliary Cholangitis (PBC) and Moderate-to-Severe Pruritus	Recruiting	Cholestatic Pruritus	Drug: CR845 1.0 mg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT03995212
11	Study to Evaluate IV CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg|Drug: CR845 1 mcg/kg|Drug: CR845 1.5mcg/kg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT02858726
12	A Study to Evaluate the Safety and Efficacy of CR845 in Chronic Kidney	Completed	Chronic Kidney Diseases|Pruritus	Drug: CR845 0.25 mg Oral Tablet|Drug: CR845 0.5 mg Oral Tablet|Drug: CR845 1 mg Oral Tablet|Drug: Placebo Oral Tablet	https://ClinicalTrials.gov/show/NCT03617536

	Disease Patients With Moderate-to-Severe Pruritus				
13	CR845-CLIN3103: A Global Study to Evaluate the Safety and Efficacy of CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT03636269
14	Extension Study to Evaluate IV CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845	https://ClinicalTrials.gov/show/NCT03281538
15	CR845-CLIN3105: A Study to Evaluate the Safety and Effectiveness of CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg	https://ClinicalTrials.gov/show/NCT03998163
16	A Study to Evaluate the Safety and Efficacy of CR845 in Hemodialysis Patients With Moderate-to-Severe Pruritus (KALM-1)	Completed	Uremic Pruritus	Drug: CR845 0.5 mcg/kg|Drug: Placebo	https://ClinicalTrials.gov/show/NCT03422653
17	A Study Evaluating Pain Relief and Safety of Orally Administered CR845 in Patients With Osteoarthritis of Hip or Knee	Completed	Osteoarthritis, Hip|Osteoarthritis, Knee|Arthritis|Joint Diseases|Musculoskeletal Diseases|Rheumatic Diseases	Drug: CR845 tablet 1 mg|Drug: CR845 tablet 2.5 mg|Drug: CR845 tablet 5 mg|Drug: Placebo tablet	https://ClinicalTrials.gov/show/NCT02944448
18	A Study Evaluating the Overall Pain Relief and Safety of Intravenous (IV) CR845 in Patients Undergoing Abdominal Surgery	Completed	Post Abdominal Surgery Pain	Drug: CR845 IV 1 mcg/kg|Drug: CR845 IV 0.5 mcg/kg|Drug: Placebo IV	https://ClinicalTrials.gov/show/NCT02542384

A 2. Both the treatment pathway search and economic SLR report a single search strategy for both Medline and Embase searches via Embase.com. Please confirm if this is a simultaneous search of both resources using a single strategy or a single search of the Embase database conducted on the understanding that it now contains all records from Medline.

Because Medline records are also indexed on Embase.com, a single search was performed covering both databases, to avoid duplication of records. It is understood that sometimes the MESH terms differ from EMTREE terms.

A different search string was also conducted for both the searches using Pubmed, covering MEDLINE in-process studies.

A 3. Section 2.2.1 (document Appendix D, G, H, and I - SLR results) lists a search of the Cochrane Database of Systematic Reviews (CDSR) via the Cochrane library, however the two Cochrane strategies provided in the search strategy document don't appear to report a search of this resource. The search for treatment pathway appears to be limited to Cochrane Clinical Answers (Table 3, document search strategy (SLR)), whilst the search reported in the economic SLR is limited to trials in CENTRAL. Please provide a full strategy for any searches of CDSR or confirm if this is a reporting error.

The CDSR was searched, but it is not part of the search strategy or search databases used for the review process. It was searched to retrieve the relevant systematic reviews for the validation of the searches.

A 4. The Embase.com search strategy reported in table 5 of the economic SLR appears to contain a number of line combination errors:

- line #19 appears to accidentally combine line #14 with the facets for the economic study designs, and;
- line #20 combines lines #13 and #18, rather than #14 and #19 which means that lines #1-7 and #15-17 are missed from the final line combination.

So instead of

CKD AND (pruritus or antipruritics) AND Economic study designs including utilities etc, the strategy retrieves:

- (pruritus or antipruritics) AND Economic terms from line #18 only.

Please see below for the full strategy and clarify if this was a reporting error or rerun and screen the results if this was an error in the strategy.

Table 1 (Table 5 from 'search strategy (SLR)' document): Economic SLR search strategy applied using Embase.com covering Embase and MEDLINE (Searched on April 18, 2022)

No.	Query	Hits
	'chronic kidney failure'/exp OR 'chronic kidney failure' OR 'chronic kidney failure' /syn	
	kidney disease' /exp OR 'kidney disease' OR 'kidney disease' /syn OR 'chronic kidney disease'	
	'end stage renal disease' /exp OR 'end stage renal disease' OR 'end stage renal disease' /syn OR 'end stage kidney disease'	
	esrd :ab,ti OR ckd:ab,ti OR esrf:ab,ti	
	(renal OR kidney) NEAR/3 (failure OR disease OR dialysis)	
	hemodialysis OR 'peritoneal dialysis' OR hemodialysis	
	pruritus'/exp OR 'pruritus' OR 'pruritus'/syn OR itch*:ab,ti OR 'itching pruritis'	
	'uremic pruritus' OR 'uraemic pruritus' OR pruritus OR 'renal itch' OR 'chronic kidney disease-associated pruritus' OR 'ckd-ap'	
	'antipruritic agent'/exp OR 'antipruritic agent'/syn OR 'antipruritic agent'	
	antiprurit*:ab,ti	
	'gabapentin'/syn OR 'μ-receptor antagonists' OR 'κ agonists'	
13		
14		
	((utilit* NEAR/2 (measure* OR outcome* OR state* OR health OR score* OR weight* OR analysis)):ab,ti) OR 'health utility index' OR 'hui':ab,ti OR (utilit* NEXT/1 (score* OR value* OR evaluation*)) OR (health NEXT/2 utilit*) OR (('health'/exp OR 'health') AND (state NEXT/1 utilit*)) OR hui:ab,ti OR ((health NEXT/1 state*) AND (state* NEXT/1 preference*)) OR 'quality adjusted life year'/exp OR 'quality adjusted life year' OR 'quality adjusted life' OR ('quality adjusted' NEXT/1 survival*) OR qaly:ab,ti OR qald:ab,ti OR qale*:ab,ti OR qtime*:ab,ti OR 'disability adjusted life' OR daly*:ab,ti OR 'health survey'/exp OR 'health survey' OR hye*:ab,ti OR health*year*equivalent OR (health NEAR/2 utility*) OR (willingness NEAR/2 pay) OR (standard NEAR/2 gamble) OR	

No.	Query	Hits
	disutili*:ab,ti OR (time NEAR/2 trade*off) OR tto:ab,ti OR ('discrete choice' NEXT/1 experiment*)	
	qwb:ab,ti OR 'short form 36'/exp OR 'short form 36' OR 'sf36':ab,ti OR 'sf-36':ab,ti OR 'sf 36':ab,ti OR 'short form 12'/exp OR 'short form 12' OR 'sf12':ab,ti OR 'sf-12':ab,ti OR 'sf 12':ab,ti OR 'short form 6' OR 'sf6':ab,ti OR 'sf-6':ab,ti OR 'sf 6':ab,ti OR 'euroqol' OR 'euro-qol' OR 'euro qol' OR 'eq5d':ab,ti OR 'eq-5d':ab,ti OR 'eq 5d':ab,ti	
	'cost effectiveness analysis'/syn OR 'cost utility analysis'/syn OR 'economic evaluation'/syn OR (('cost-effectiveness' OR 'cost-utility') NEXT/1 (evaluation* OR analys* OR model* OR intervention*))	
18	('economics'/exp OR 'economics'/de OR 'economic aspect'/exp OR 'economic aspect'/de OR 'cost'/exp OR 'cost'/de OR 'health care cost'/exp OR 'health care cost'/de OR 'drug cost'/exp OR 'drug cost'/de OR 'hospital cost'/exp OR 'hospital cost'/de OR imaging) AND cost:ab,ti OR 'blood test':ab,ti OR 'caregiver cost':ab,ti OR 'chemotherapy cost':ab,ti OR 'socioeconomics'/exp OR 'socioeconomics'/de OR 'health economics'/exp OR 'health economics'/de OR 'pharmacoeconomics'/exp OR 'pharmacoeconomics'/de OR 'fee'/exp OR 'budget'/exp OR 'hospital finance'/exp OR 'hospital finance'/de OR 'financial management'/exp OR 'financial management'/de OR 'health care financing'/exp OR 'health care financing'/de OR 'low cost' OR 'high cost' OR (health*care NEXT/1 cost*) OR ('health care' NEXT/1 cost*) OR fiscal OR 'funding'/exp OR funding OR financial OR 'finance'/exp OR finance OR (cost NEXT/1 estimate*) OR 'cost variable' OR (unit NEXT/1 cost*) OR economic*:ab,ti OR pharmacoeconomic*:ab,ti OR price*:ab,ti OR pricing:ab,ti OR ((cost* NEAR/3 (treat* OR therap*)):ab,ti) OR (((direct OR indirect) NEAR/2 cost*):ab,ti) OR ('health*care' NEXT/1 (utilisation OR utilization)) OR ('health care' NEXT/1 (utilisation OR utilization)) OR (resource NEXT/1 (utilisation OR utilization OR use))	
	#14 OR #15 OR #16 OR #17	
	#13 AND #18	
	#13 AND #18 AND [2012-2022]/py	

This was a reporting error due to shifting of one row. The original screening was based on the proposed strategy only.

Please see attached the printed version of the Embase search applied on 18th April 2022.

A 5. A search of NHS Economic Evaluation Database (NHS EED) is reported in section 2.2.1 (document Appendix D, G, H, and I - SLR results) and in the PRISMA flow chart for the treatment pathway review (3.1, Appendix D, G, H and I). This appears to be missing from the search strategy document, please provide full details.

A separate table for NHSEED searches has been added in the updated search strategy document

No.	Query	Hits
1	(pruritus) IN NHSEED FROM 2011 TO 2022	1
2	(chronic kidney disease) AND (pruritus) IN NHSEED FROM 2011 TO 2022	0
3	(chronic kidney disease) AND (itch) IN NHSEED FROM 2011 TO 2022	0
4	(haemodialysis) AND (chronic kidney disease) IN NHSEED FROM 2011 TO 2022	2
5	Total	3

A 6. Please provide the search date and keywords used for the search of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conference proceedings reported in Section 2.2.2. (document Appendix D, G, H, and I - SLR results).

The keywords used for conference searching were ‘pruritus’, ‘CKD’, ‘chronic kidney disease’, ‘chronic renal disease’, ‘itching’, ‘itch’, ‘dialysis’, ‘ESRD’, and ‘end stage renal disease’. The search was performed on 26th April 2022.

A 7. Section 2.2.3: Other literature sources (document Appendix D, G, H, and I - SLR results) mentions *“Hand searching of country-specific websites for relevant objectives”*. Please provide details of which sites were searched, the search dates, number of hits retrieved, and keywords used.

Websites such as British Association of Dermatologists (bad.org.uk) and Guidelines International Network (https://g-i-n.net/) were searched on 25th April 2022 using the keywords ‘pruritus’, ‘CKD’, ‘chronic kidney disease’, ‘chronic renal disease’, ‘itch’, ‘itching’, ‘dialysis’, ‘ESRD’, and ‘end stage renal disease’.

A 8. The summary protocol (document Appendix D, G, H, and I - SLR results, section 10 Appendices) mentions searches of HTA bodies. Please confirm whether these searches were carried out and if yes, please provide full details including date searched, keywords used, and hits retrieved.

Details of HTA searching are provided below. It was performed on 27th April 2022

HTA	Link	Keywords	Total hits screened	Total no
National Institute for Health and Care Excellence (NICE)	http://www.nice.org.uk/	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	37+149+176+76+2+50+67+1+39	552
Scottish Medicines Consortium (SMC)	https://www.scottishmedicines.org.uk/search/	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	0+11+34+7+2+22+1+11	88
Canadian Agency for Drugs and Technologies in Health (CADTH)	https://www.cadth.ca/	Pruritus, CKD, chronic kidney disease or chronic renal disease or ESRD or end stage renal disease, itch, itching, dialysis, ESRD	6+11+682+7+0+32+2	740
German Institute for Quality and Efficiency in Health Care (IQWiG)	https://www.iqwig.de/	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	192+67+154+195+9+12+46+27+139	841
Institute for Clinical and Economic Review (ICER)	https://icer-review.org/	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	1+10+7+1+9+39+3+0+0	70
Pharmaceutical Benefits Advisory Committee (PBAC)	http://www.pbs.gov.au/pbs/home	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	0	0
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear2n/search/search.aspx	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	1+33+17+2+5+137+4+2	201
EQ-5D Publications Database	https://euroqol.org/search-for-eq-5d-publications/	Pruritus, CKD chronic kidney disease chronic renal disease, itch itching, dialysis ESRD end stage renal disease	0	0
University of Sheffield Health Utilities Database	https://www.scharrhud.org/index.php?recordsN1&m=search&action=searchRecords	Pruritus, CKD, chronic kidney disease, chronic renal disease, itch, itching, dialysis, ESRD, end stage renal disease	1+4+7+1+1+0+8+1+0	23

A 9. There appears to be a discrepancy regarding the number of records retrieved by both the MEDLINE In-process (via Pubmed) and CENTRAL

(via Cochrane library) searches reported in both the PRISMA diagram for the cost resource use and utility review (Figure 3 of 'Appendix D, G, H and I – SLR results') and the corresponding search strategies recorded in the search strategy document (Tables 6 & 7 (from 'search strategy (SLR)' document). Please clarify which figures are correct.

Resource	Hits retrieved (PRISMA)	Hits retrieved (Search Strategy)
Medline in Process (Pubmed)	70	65
CENTRAL	15	19

The hits of CENTRAL and MEDLINE retrieved are correct and most recent as reported in PRISMA. The actual search for screening was performed on 19th April for CENTRAL and MEDLINE and on 18th April for Embase.

Decision problem

A 10. Priority question: The population in the decision problem is defined as patients where established clinical management is insufficient in reducing pruritus. Please clarify whether this implies a restriction to a later line of therapy than first line; if so, then which treatments as part of established clinical management need to have been tried before determining insufficiency in reducing pruritus? Which criteria would be applied in clinical practice for this determination?

Guidelines recommend ensuring adequate dialysis, normalising the calcium-phosphate balance, controlling parathyroid hormones (PTH) to acceptable levels, correcting any anaemia, and using simple emollients before employing other treatment strategies. If a patient is still suffering from pruritus the next stage is to use best supportive care, including creams and emollients, antihistamines, gabapentin and in some cases ultraviolet therapy or antidepressants. Those on no interventions are also deemed to be on best supportive care. If a patient has failed on first line treatment (best supportive care), difelikefalin will be offered for the duration of dialysis, as long as a sufficient reduction in itch score has been achieved within the first 12 weeks of treatment.

Insufficiency of reducing pruritus is determined by whether the patient is still experiencing moderate-to-severe itch (as reported by the patient) that has not been resolved with current management.

A 11. Priority question: In Section B.1.3, the company states that "If a patient has failed on best supportive care this is when difelikefalin will be offered for the duration of dialysis, as long as a sufficient reduction in itch score has been achieved within the first 12 weeks of treatment." (p. 20). This description is different from the definition of the population in the decision problem. Please clarify whether Difelikefalin is to be prescribed after best supportive care (BSC) has failed, in addition to BSC or instead of BSC.

Difelikefalin is to be prescribed after best supportive care has failed.

A 12. Priority question: Concerning the submission patient population,

a) Please clarify whether the submission population included only adults with Stage 5 CKD/end-stage renal disease (ESRD).

The submission population is the full population covered by the marketing authorisation for difelikefalin. Difelikefalin is indicated for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients on haemodialysis. Chronic kidney disease patients on haemodialysis have, by definition, end-stage renal disease (ESRD) and are, by definition, Stage 5 CKD.

b) If so, please clarify whether the submission population of ESRD patients with moderate-to-severe pruritus is narrower than the NICE final scope, which is 'adults with moderate-to-severe pruritus receiving haemodialysis'.

The submission population can be considered narrower than the final scope by virtue of the fact that difelikefalin treatment is only authorised for use in patients with chronic kidney disease and should be restricted for in-centre haemodialysis use only.

c) Please clarify whether moderate-to-severe pruritus is not observed in stage 4 CKD/ acute renal failure patients on haemodialysis.

Moderate to severe pruritus is observed in Stage 4 CKD and potentially in acute renal failure patients on haemodialysis. However, such patients are not included in the authorised indication for difelikefalin nor, therefore, the submission population.

A 13. Priority question: The NICE Final Scope defined the intervention as "*Difelikefalin*" and the comparator as "*Established clinical management without Difelikefalin, including gabapentin and pregabalin*" whereas the KALM-1 and KALM-2 RCTs included the respective definitions of "*Difelikefalin*" and "*Placebo*". Concerning the submission intervention and comparator, the CS states: "*It is proposed that difelikefalin be used as an adjunct to established clinical management where established clinical management is insufficient in reducing pruritus*."

a) It would appear from the list of concomitant medications that patients on both the intervention (difelikefalin) and comparator arm (placebo) of the KALM trials received established clinical management including gabapentin and pregabalin. Please clarify whether this was the case.

Yes. For the KALM trials there were no changes made to current established management. Where a patient was on no prior medication, this was still considered to be established clinical management.

b) Please clarify whether the submission intervention is therefore *Difelikefalin + established clinical management*, and the comparator is *established clinical management*.

Yes. The submission intervention is Difelikefalin + established clinical management, with a comparator of established clinical management

c) If yes to b), please justify this departure from the NICE final scope intervention (*Difelikefalin*).

The submission is in line with the licenced indication and reflects usage in the KALM-1 and KALM-2 trials (please see table below).

Table 2 defined populations

Final scope issued by NICE	Decision problem addressed in the company submission	Licenced indication	Draft scope	KALM-1	KALM-2
Adults with moderate-to-severe pruritus receiving haemodialysis	For the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis, including where established clinical management is insufficient in reducing pruritus	Difelikefalin is indicated for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients on haemodialysis	The population studied includes patients where other treatment options have failed. All patients with moderate-to-severe CKD-aP benefit in a similar way, and no subgroups have been identified where difelikefalin would be more or less beneficial	Adults (≥18 years of age) with end-stage renal disease who had been on HD at least three times per week for at least three months and who had moderate-to-severe CKD-aP, defined as a weekly mean score of >4 points on the 24-hour WI-Numerical Rating Scale (NRS) (Worst Itching Intensity Numerical Rating Scale).	Adults with moderate-to-severe pruritus receiving haemodialysis

d) If no to b), then evidence of Difelikefalin plus established clinical management vs. established clinical management would be inappropriate. In this case, could the company please provide evidence of the comparison of Difelikefalin only vs. established clinical management.

N/A - see response to A13b.

Systematic review

A 14. Priority question: Section B.2.1 of the CS suggests that Appendix D should provide "full details of the process and methods used to identify and select the clinical evidence relevant to the technology being evaluated." However, scrutiny of Appendix D suggests that whilst certain individual studies were described, there were no details of the methods used to identify and select clinical effectiveness evidence. Please provide full details of methods including: review question; study eligibility criteria; search strategy (please see questions on Literature Searches above); data extraction approach; critical appraisal (including a full bibliographic reference for the checklist used); and methods of pooling data.

Please see response to question A1

A 15. Priority question: 'Appendix D, G, H and I – SLR results' describes methods and results for literature reviews on: treatment pathway; utilities; costs and resource use; and economic modelling studies. Section 2.3.3 of 'Appendix D, G, H and I – SLR results' mentions specific critical appraisal checklists used to assess studies relating to utilities, costs and resource use and economic evaluation.

a) The population described in Table 1 ("PICOS eligibility criteria") of this appendix ("Adult patients with chronic kidney disease-associated pruritus (CKD-aP)") differs to that shown in the NICE Final Scope ("Adults having haemodialysis with moderate-to-severe pruritus") and in the decision problem (Table 1 of Document B) ("For the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis, including where established clinical management is insufficient in reducing pruritus"). Please explain the differences between these population characteristics.

Given the limited quantity of published literature in this area the PICOS criteria was kept broad to ensure all relevant papers with useful evidence were captured. See Table 2.

b) Please clarify whether studies relating to treatment pathway (i.e., practice guidelines and reviews) were also critically appraised. If so, please provide details of the checklist(s) used.

The guideline reviews were not critically appraised as the guidelines are broad. Only a small portion of the guidelines were extracted as not all sections were relevant to CKD-aP, meaning a critical appraisal of the guidelines was not appropriate. Furthermore, the guidelines are mainly issued by national bodies concerned with CKD or pruritis specifically, rather than CKD-aP.

c) For critical appraisal of all the above types of studies, please clarify the number of reviewers involved in the process and the methods used to resolve disagreements.

The critical appraisal conducted during the full-text review phase was assessed by two independent researchers. Three researchers were involved in this process, who each reviewed two-thirds of the articles. Any disagreement between the two researchers were resolved with the third researcher. Disagreements were discussed until a consensus formed over inclusion/quality assessment.

A 16. Priority question: The comparator in Table 1 ("*PICOS eligibility criteria*") of 'Appendix D, G, H and I – SLR results' is described as "*Pharmacological interventions only*" whereas the NICE Final Scope and decision problem (Table 1 of Document B) list the comparator as "*Established clinical management without difelikefalin, including gabapentin and pregabalin*". Please explain the differences between these comparator characteristics.

The SLR description of "Pharmacological interventions only" was deliberately broad so as to not restrict the results. The NICE Final Scope and Decision problem definition "Established clinical management without Difelikefalin, including

gabapentin and pregabalin" reflects the comparator definition as per the KALM-1 and KALM-2 trials. Established clinical management could also include "no treatment".

A 17. Date and language restrictions were applied to studies relating to treatment pathway, utilities, costs and resource use and economic evaluation (Table 1 of 'Appendix D, G, H and I – SLR results').

a) Please explain the impact on results of all reviews of omitting evidence published before 2012.

It was decided that date restrictions had a low impact on the omission of relevant information, and helped to provide the most recent evidence. Details have been provided of the 741 studies excluded prior to 2012 in an attached document. (See Excel file - SLR in CKD-aP Extended Searches 9th Sept 2022).

b) Please explain the impact on results of all reviews of restricting to studies published in English language or having and English abstract or summary available.

It was deemed that language restrictions had a low impact. Details have been provided of the 12 non-English papers excluded. (See Excel file - SLR in CKD-ap Extended Searches 9th Sept 2022)

A 18. Please clarify the number of reviewers involved in data extraction of all the above studies, giving precise details of how the process was operationalised. Please also outline the approach for resolving disagreements/errors.

The SLR followed a robust methodology that was fully compliant with PRISMA-P1 guidelines and meets the standards described by NICE. The SLR employed a standard two review process and quality control for evidence screening at first (Title/Abstract) and second stage (Full texts).

Two investigators, working independently, extracted data for the study population of interest for the final list of selected eligible studies. Any discrepancies observed between the data extracted by the two data extractors were resolved by discussion and coming to a consensus. If a consensus was not formed, the third independent reviewer provided arbitration. Details have been provided in section 2.3.2 of Appendix D, G, H and I.

A 19. Please provide a list of excluded studies for each of the above reviews, showing the full bibliographic details for studies excluded at the full text screening stage, together with reasons for exclusion.

See list of excluded studies provided.

Clinical effectiveness evidence

A 20. Priority question: The comparator according to NICE is "*established clinical management without difelikefalin, including gabapentin and pregabalin*". Furthermore, Table 12 of document B reports that the permitted concomitant medications include antihistamines, corticosteroids and opioids, as well as gabapentin, or pregabalin. On the other hand, the clinical study reports (CSRs) for both KALM-1 and KALM-2 state that patients who "*had received new or changed prescription for opioids, gabapentin, or pregabalin within 14 days prior to screening*" would be excluded for the double-blinded phases (section 9.3.2), thus implying that if a patient was already receiving any of them at a certain dose 14 days prior to screening, they would be allowed to enrol in the trial and continue using them at the same dose.

a) Please confirm whether subjects were allowed to be treated with one of the comparators within the trials.

Subjects were allowed to be treated with the comparators. The only restriction was that changes to current prescription should be avoid from screening to the end of the double-blind treatment period, unless needed for the acute treatment of AEs or emergent medical conditions (as per study protocols).

b) Please clarify whether opioid antagonists were permitted concomitant treatments as opposed to opioid agonists.

The opioid antagonists naloxone and naltrexone were not permitted to be used from the start of dosing of the double-blind treatment period to the end of the open-label treatment period, unless needed to for the acute treatment of an adverse event or

emergent medical condition. The same was true of mixed agonist-antagonists such as buprenorphine and nalbuphine.

c) Please tabulate use of these concomitant medications at baseline across all relevant trials.

A summary of the pooled concomitant medications from KALM-1 and KALM—2 is provided in the table below(3):

Medication	Placebo (n=425)	Difelikefalin (n=426)
Any baseline use of an anti-itch medication	163 (38.4%)	159 (37.3%)
Most commonly-used anti-itch medications at baseline (>2%)		
Diphenhydramine	100 (23.5%)	104 (24.4%)
Hydroxyzine	52 (12.2%)	42 (9.9%)
Hydrocortisone	16 (3.8%)	11 (2.6%)
Cetirizine	10 (2.4%)	7 (1.6%)
Clemastine	10 (2.4%)	7 (1.6%)

Please see appendix P for all concomitant medications split by KALM-1 and KALM-2 for further details.

d) Please report the number of patients who had prescription adjustments made due to adverse events or medical conditions by treatment in all relevant trials.

A review of Listing 16.2.10.1 (Prior and Concomitant Medications) indicates that 11 patients in each group had adjustments to their concomitant anti-itch medication during the double -blind period in the KALM-1 study and 6 patients in each group had adjustments to their concomitant anti-itch medication during the double -blind period in the KALM-2 study.

e) The subgroup analysis in Appendix E, reports that 39.7% of the subjects in KALM-1 and 34.9% in KALM-2 had received anti-itch medication at baseline. Please provide details on what specific anti-itch medication was received.

Please see A20 c) above.

For details on *all* previous anti-itch medications received in the 3 months prior to study commencement please see appendix Q

f) The CSR for KALM-1 in Table 29 and for KALM-2 in Table 33, present the rates of subjects using any concomitant anti-itch medication during the Double-blind Treatment Period by ingredients.

i) Please provide an alternative table including the details of the medication (mode of action and commercial name), also specifying if they are included in established clinical management.

KALM-1 CSR Table 29 – adaption

Medication Generic Name (drug class)	**Mode of action**	**Included in established clinical management**	**Placebo (N = 188) n (%)**	**DFK (N = 189) n (%)**	**All Subjects (N = 377) n (%)**
DIPHENHYDRAMINE (antihistamine)	Inhibits the effects of histamines in the body, providing symptomatic relief of itching(4)	Yes	71 (37.8%)	63 (33.3%)	134 (35.5%)
HYDROXYZINE (antihistamine)		Yes	19 (10.1%)	18 (9.5%)	37 (9.8%)
HYDROCORTISONE (corticosteroid)	Inhibits immune response by modifying the	Yes	8 (4.3%)	6 (3.2%)	14 (3.7%)

TRIAMCINOLONE (corticosteroid)	function of dermal cells, epidermal cells and leucocytes, reducing itch and inflammation(5)	Yes	3 (1.6%)	5 (2.6%)	8 (2.1%)
AMMONIUM LACTATE (topical emollient)	Promotes moisturisation and hydration of skin and provides symptomatic relief of itching(6)	Yes	4 (2.1%)	2 (1.1%)	6 (1.6%)

KALM-2 CSR Table 33 - adaption

Medication (drug class)	**Mode of action**	**Included in established clinical management**	**Placebo (N = 188) n (%)**	**DFK (N = 189) n (%)**	**All Subjects (N = 377) n (%)**
DIPHENHYDRAMINE (antihistamine)	Reduces the effects of histamines in the body, providing symptomatic relief of itching(4)	Yes	26 (11.0%)	45 (19.1%)	71 (15.1%)
HYDROXYZINE (antihistamine)		Yes	27 (11.4%)	22 (9.4%)	49 (10.4%)
CLEMASTINE (antihistamine)		Yes	10 (4.2%)	8 (3.4%)	18 (3.8%)
CETIRIZINE (antihistamine)		Yes	7 (3.0%)	4 (1.7%)	11 (2.3%)

LORATADINE (antihistamine)		Yes	4 (1.7%)	5 (2.1%)	9 (1.9%)
CHLORPHENAMINE (antihistamine)		Yes	5 (2.1%)	2 (0.9%)	7 (1.5%)
HYDROCORTISONE (corticosteroid)	Inhibits immune response by modifying the function of dermal cells, epidermal cells and leucocytes, reducing itch and inflammation(5)	Yes	8 (3.4%)	4 (1.7%)	12 (2.5 %)

ii) Please provide sub-group analysis for different anti-itch medication used during the Double-blind Treatment Period for all studies.

Please see 'Appendix S - ISE outputs by anti-itch medications' for sub-group analysis of the different anti-itch medications used.

iii) Document B contains several references to the use of antidepressants as part of ECM however, these are not mentioned in Table 12 (Summary of methodology) as being allowed or disallowed from the clinical effectiveness studies. Please confirm whether antidepressants were permitted as a concomitant intervention in the clinical effectiveness studies.

Antidepressants were permitted if they were part of established (>2 weeks) clinical management for a patient.

A 21. Priority question: KALM-2 and KALM-2 OLE appears to have recruited patients from study centres in the UK.

a) Please provide the number of patients enrolled in the UK from these centres.

There were 20 GBR subjects (6 DFK and 14 placebo) recruited from 5 centres in the UK

b) If ≥10 patients, please provide baseline characteristics for these patients.

Baseline characteristic	All subjects (N = 20)
Number of participants	20
Mean age, years (SD)	64.9 (11.11)
Male	9 (45.0%)
Female	11 (55.0%)
Ethnicity – n (%)	
Hispanic or Latino	0 (0.0%)
Not Hispanic or Latino	20 (100%)
Not reported	0 (0.0%)
Unknown	0 (0.0%)
Race – n (%)	
Asian	2 (10.0%)
Black or African American	3 (15.0%)
White	14 (70.0%)
Other	1 (5.0%)
Mean prescription dry body weight, kg (SD)	79.98 (21.658)
Baseline Worst Itching NRS, Mean (SD)	7.31 (1.624)
Baseline anti-itch medication use – [1] n (%)	
Yes	7 (35.0%)
No	13 (65.0%)
Specific medical conditions? – [1] n (%)	
Yes	4 (20.0%)
No	16 (80.0%)
Mean duration of pruritus, years (SD)	2.72 (3.407)
Mean years since diagnosis of ESRD, years (SD)	5.94 (6.552)
Years since diagnosis of CKD	
n	20
Mean (SD)	14.29 (11.918)
Years on chronic haemodialysis, mean (SD)	5.67 (6.444)
Aetiology of CKD [2]	
Diabetes	10 (50.0%)

Baseline characteristic	All subjects (N = 20)
Hypertension	11 (55.0%)
Large vessel disease	0
Glomerulonephritis	3 (15.0%)
Vasculitis	0
Interstitial nephritis	0
Pyelonephritis	1 (5.0%)
Cystic	0
Hereditary	0
Congenital	0
Neoplasms	0
Tumours	0
Urologic	0
Nephrotic syndrome	4 (20.0%)
Unknown	2 (10.0%)
Other	2 (10.0%)

CKD = chronic kidney disease; ESRD = end-stage renal disease; max = maximum; min = minimum; NRS = Numerical Rating Scale; SD = standard deviation.

[1] Observed stratum values.

[2] More than one item may have been checked.

A 22. Priority question: When considering whether the results of a trial are relevant to a particular health service, the population characteristics of the trial need to be seen to be comparable to those of the target population. This is vital for characteristics that may be outcome modifiers, such as ethnicity, gender and age (see p22 in Document B, CS). Tables 7 and 8 provide clear data on the ethnic sub-groups (race), gender and age in the KALM 1and 2 studies. To evaluate comparability with the UK target population (all those with CKD and pruritis in the UK) it is necessary to know the ethnicity (race), age and gender characteristics of the UK target population. However, these data are unavailable in the CS.

a) Please provide data on the proportions of people in different ethnic sub-groups (for example, Asian, Black, White, etc), the mean age, and the proportions of males and females in the UK population of people with CKD and pruritis

The table below includes demographic data from the UK Renal Registry of adult patients on in-centre haemodialysis in England; approximately 70% of which are somewhat bother by pruritus and 47% experience moderate to extreme itching (Sukul et 2021), along with equivalent data from the pooled KALM studies (Topf et al 2022).

	Ethnicity (Race)			Gender	Age
	White	**Black**	**Asian/other**	**Male**	**Median**
UKRR Adults ICHD*	67.6%	12.8%	19.6%	62.3%	66.5
KALM pooled dataset	60.8%	29.2%	10.0%	59.6%	60.0

* UK Renal Registry (2022) UK Renal Registry 24th Annual Report – data to 31/12/2020, Bristol, UK. Available from https://ukkidney.org/audit-research/annual-report

b) Please explain how these data demonstrate the representativeness (or not) of the trial data to the UK population of people with CKD and pruritis

The UK population is slightly older and consists of slightly more white and fewer black patients than the population of patients participating in the KALM studies. The trial data is considered to be representative of a UK population with CKD-aP, verified by clinicians. Please see Appendix N: Clinical Opinion and consensus report:
'Advisors were clear that the KALM-1 and KALM-2 are high quality studies'
'The majority (7) of the group agreed that the KALM-1 and KALM-2 studies were broadly generalisable to the UK population'.

A 23. Priority question: If the trial and target population have different characteristics (as discussed in A23) then this may influence the validity of inferring any effects from the trial to the target population, if the characteristics are effect modifiers. Effect modification may be inferred from sub-group analysis. A sub-group analysis including the characteristics of race, age and gender is only presented in Figure 6 of the pre-proof of the Topf et al., 2022 paper, which presents a pooled analysis of all four KALM-1 and KALM-2 studies. Please provide subgroup analysis for race, gender and age for each study individually for the primary efficacy results accompanied by a discussion.

Please see ‘Appendix T - KALM1_KALM2_by subgroups’
Heterogeneity was observed across studies with respect to age, sex, and race subgroups. Although generally response is greater to the <65 subgroup it was numerically higher in older patients in the KALM-2 study.

A 24. Priority question: Please provide an illustration of all the subgroup analysis results reported in Appendix E of document B, for ease of comparison.

Updated in Appendix E in Document B

A 25. Priority question: Multiple Imputation (MI) has been used throughout the CS. The use of MI was built in the statistical analysis plan (SAP) for handling missing numerical rating scale (NRS) data. It is not clear why the SAP presupposed that the proportion and nature of missing data would justify its use.

a) Please provide the rationale for using MI over other available methods for handling missing data.

The choice of multiple imputation for the treatment of missing data was suggested by the Food and Drug Administration (FDA) during a meeting held on September 6th 2017 to discuss the Phase 3 clinical development program for IV difelikefalin. Specifically, the FDA stated that “*The efficacy analyses should be based on the intent-to-treat (ITT) population (i.e., all randomized subjects)”* and added that the protocol *“should pre-specify a scientifically sound primary imputation method (e.g. multiple imputation) to handle missing data.* Multiple imputation is also one of the analytical methods recommended by the National Research Council Committee on National Statistics in their 2010 report on the prevention and treatment of missing data in clinical trials. Based on these regulatory and technical recommendations, the sponsor decided to use multiple imputation as the primary method for the treatment of missing data in the pivotal studies KALM-1 and KALM-2.

b) Please elaborate on the specific methods used within the MI process.

The specific methods used in the MI process are detailed in the Statistical Analysis Plan (SAP) for each of KALM-1 and KALM-2 (Section 8). Specifically, the SAP stated the following:

- Intermittent missing weekly mean WI-NRS scores were first imputed using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data.

- The monotone missing weekly mean WI-NRS values were then multiply imputed with the SAS MI procedure using the monotone regression method.

- For each stage, MI was performed within treatment group with covariates for baseline WI-NRS score, both randomisation stratification factors, region (in CLIN3103 only), and the non-missing WI-NRS scores for each week. Should convergence issues occur due to small cell size for the categorical covariates corresponding to strata (at either stage), those specific covariates will be removed from the model. For study CLIN3103, the handling of convergence issues related to the region covariate were described in section 8.1.4 of the SAP.

- The proportion of subjects who have an improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score ≥3 points will be calculated for each imputed dataset. Differences between DFK 0.5 mcg/kg and placebo with respect to the primary endpoint were compared using a logistic regression model containing terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and presence of specific medical conditions. For KALM-2, the handling of convergence issues with the region covariate were described in section 8.1.4 of the SAP.

- Twenty imputations were performed.

- Results of the logistic regression on the multiply imputed data sets will be summarised by the SAS MIANALYZE procedure.

The above MI process was implemented independently among subjects contributing to the interim results and those following the interim analysis. Likewise, the logistic regression and results described above were generated independently for both samples, with the samples combined and adjusted using the methodology proposed by Cui, Hung, Wang (1999).(7)

Sections 13 and 14 of the SAP for each of the pivotal studies provided sample SAS code in additional to the seeds to be used in the multiple imputation methodology. This pre-specification ensured that the analysis was not data driven.

c) Please provide an overview of missingness patterns observed and their frequencies, both for the intermediate missings and the monotone missings.

The pattern of missing data of the weekly mean WI-NRS in studies KALM-1 and KALM-2 are provided in Tables 3c-1 and 3c-2, respectively in Appendix V as well as the reasoning for this missing data. In both studies, none of the covariates were missing. There were few intermittent missing WINRS scores. In KALM-1, 77.8% of the placebo patients had complete data; this percentage was 74.6% for patients randomised to DFK. A similar result was observed in KALM-2, with 80.9% of the placebo patients having complete data compared to 73.8% of DFK patients. 12.8% of placebo patients and 16.9% of DFK patients in CLIN3102 had a missing weekly WI-NRS score at Week 12, the primary timepoint. Similarly, 12.3% of the patients randomised to placebo had a missing weekly mean WI-NRS at Week 12 compared to 19.4% of the patients randomised to DFK**.**

A 26. Priority question: The covariates used in the MI analysis for KALM-1 and KALM-2 are not reported in document B.

a) Please confirm that the covariates for KALM-1 MI analysis were: baseline WI-NRS score, both randomization stratification factors (use of anti-itch medication during the week prior to randomization and presence of specific medical conditions), and the non-missing NRS scores for each week, as stated in the SAP.

We can confirm that the covariates for the KALM-1 (CLIN3102) MI were baseline WI-NRS score, both randomisation stratification factors (use of anti-itch medication during the week prior to randomisation and presence of specific medical conditions), and the non-missing NRS scores for each week.

b) Please confirm that the covariates for KALM-2 MI analysis were: baseline WI-NRS score, both randomization stratification factors (use of anti-itch medication during the week prior to randomization and presence of

specific medical conditions), region, and the non-missing NRS scores for each week, as stated in the SAP.

We can confirm that the covariates for the KALM-2 (CLIN3103) MI analysis were baseline WI-NRS score, both randomisation stratification factors (use of anti-itch medication during the week prior to randomisation and presence of specific medical conditions), region, and the non-missing NRS scores for each week.

c) Please provide the rationale for the use of specific covariates over other potential prognostic variables correlated to the outcome of interest.

The adjustment of the analysis based on the use of anti-itch medication was suggested by the FDA, as there could be potential for differential placebo and DFK response depending on the status of this covariate. In addition, both stratification factors were included as covariates based on recommendations in ICH E9. Region was added to account for possible differences in the patient's responses to treatment based on regional differences. The baseline WI-NRS was included as a covariate because the primary endpoint, based on an improvement from baseline, would be correlated with the baseline level. No other prognostic factors were identified as important in the Phase 2 studies that were used to plan the pivotal Phase 3 studies.

A 27. Priority questions: Please provide a formal presentation of the MI analysis.

It would be helpful if the company could report the following:

- ***Number of participants per arm with complete data for the variables of interest (i.e., complete cases for the data being analysed)***
- ***Number of participants per arm with missing data for each variable of interest together with reasons for missing data***

Primary and secondary endpoints based on the weekly mean WI-NRS score

- The weekly mean WI-NRS score was computed as the average of daily e-diary entries (scored on an integer scale from 0 to 10). If there were more than 3 missed e-diary entries in a week, the weekly mean WI-NRS score was set to missing.
- The reason for a missed e-diary entry was not collected.
- CR845-CLIN3102

  - There were 147 complete cases among 189 placebo subjects (77.8%); 19 subjects (10.1%) were missing one weekly mean WI-NRS score.
  - There were 141 complete cases among 189 treated subjects (74.6%); 16 subjects (8.5%) were missing one weekly mean WI-NRS score.
  - The most common (highest frequency) pattern of incomplete data was a missing Week 12 mean WI-NRS score, which occurred in 3.7% of placebo subjects and 2.1% of treated subjects.
  - Among the 20 patterns of missing data for placebo subjects (excluding missing only Week 12), none was exhibited by more than 5 subjects.
  - Among the 23 patterns of missing data for treated patients (excluding missing only Week 12), none was exhibited by more than 4 subjects.
- CR845-CLIN3103
  - There were 191 complete cases among 236 placebo subjects (80.9%); 20 subjects (8.5%) were missing one weekly mean WI-NRS score.
  - There were 175 complete cases among 237 treated subjects (73.8%); 18 subjects (7.6%) were missing one weekly mean WI-NRS score.
  - The most common (highest frequency) pattern of incomplete data was a missing Week 12 mean WI-NRS score, which occurred in 4.2% of placebo subjects and 4.6% of treated subjects.
  - Among the 26 patterns of missing data for placebo subjects (excluding missing only Week 12), none was exhibited by more than 2 subjects.
  - Among the 27 patterns of missing data for treated subjects (excluding missing only Week 12), none was exhibited by more than 5 subjects.
- See Table NICE 3c-1 and Table NICE 3c-2 for details of missing data patterns.
- There was a wide range of missing data patterns, most experienced by very few subjects (≤1%).

- ***Number of participants excluded per arm because of missing data***
- ***It would be helpful to see a per arm table of the above together with a discussion about the differences between complete and incomplete cases***

In compliance with the intent-to-treat principle, multiple imputation was used in the primary analysis to ensure that no subjects would be excluded due to missing data.

- ***Assumptions used for the MI analysis (e.g., missing at random, missing not at random)***

As pre-specified in the statistical analysis plans for the pivotal studies:

- The primary efficacy analysis used imputed data based on a missing at random (MAR) assumption, that is, that subjects who discontinued double-blind treatment early would have similar WI-NRS scores as subjects in the same treatment group who had complete data.
- Sensitivity analysis 1 treated subjects who discontinued study drug early as non-responders.
- Sensitivity analysis 2 used multiple imputation based on a missing not at random (MNAR) assumption. A pattern mixture model was used to draw from different populations based on the reason for discontinuation.
    - Intermittent missing WI-NRS scores were imputed using the Markov Chain Monte Carlo (MCMC) method.
    - WI-NRS scores that were missing after a subject discontinued due to an adverse event were imputed using the distribution of the baseline value of all subjects' daily WI-NRS score using a trimmed normal distribution (from 4 to 10).
    - WI-NRS scores that were missing after a subject discontinued due to reasons other than an adverse event were multiply imputed using data from subjects in the same treatment group who had complete data at that time, including subjects who discontinued due to an adverse event. Terms in the model were baseline WI-NRS score, randomization stratification factors, and the prior weeks' mean WI-NRS scores.
- Sensitivity analysis 3 was a tipping point analysis that used multiple imputation with MNAR for treated subjects and MAR for placebo subjects. This analysis was used to assess the robustness of the MAR assumption. Departures from the MAR assumption were investigated by progressively decreasing the treatment differences for WI-NRS scores over the missing visits in the active treatment group until the conclusion from the primary analysis was overturned. This was applied only to Week 12 values.
    - Intermittent missing WI-NRS scores were imputed using the MCMC method.

- The monotone missing WI-NRS scores were multiply imputed using the monotone regression method.
- For each stage, multiple imputation was performed within treatment group with terms for baseline WI-NRS score, randomization stratification factors, and the non-missing WI-NRS scores for each week.
- For subjects in the active treatment group, a shift parameter ranging from 0 to 5 points in 0.25-point increments were progressively applied to impute the missing data at Week 12 until the p-value exceeded 0.05.

- ***•Software used for analysis and any specific settings/options pertaining to MI analysis***

SAS software version 9.4, proc mi was used for imputation.

- ***The number of imputed datasets and variables included***

Intermittent missing weekly mean WI-NRS scores were imputed separately for each treatment group using the MCMC method. There were 20 imputed datasets. The minimum imputed value was 0 and the maximum imputed value was 10, consistent with the range of the WI-NRS. The maximum number of iterations to impute values in the specified range was 1 million. A multiple chain imputation was used, the EM algorithm was used to generate the initial parameter estimates, the convergence criterion was 0.001, the maximum number of iterations used by the EM algorithm was 100,000, and there were 500 burn-in iterations and 100 iterations between imputations in a single chain. Since this was intended to fill in intermittent missing values, a monotone imputation was performed. The covariates were the baseline mean WI-NRS score, randomization stratification factors, and the prior weeks' mean WI-NRS scores. The SAS code from study CR845-CLIN3102 follows. Study CR845-CLIN3103 included an additional covariate for geographical region.

```
proc mi data=cp3 seed=8392857 nimpute=20  MAXIMUM=10 MINIMUM=0 out=step1  minmaxiter=1000000 ;
      mcmc chain=multiple initial = EM(CONVERGE=0.001  maxiter=100000) NBITER=500 NITER=100 impute = monotone displayinit ;
      var aimblfln smcblfln w1 w2 w3 w4 w5 w6 w7 w8 w9 w10 w11 w12 w13 ;
by trt01pn;

run;
```

The monotone missing WI-NRS values were then multiply imputed using a monotone regression method. Multiple imputation was performed within each treatment group using the baseline mean WI-NRS score, randomization stratification

factors, and the prior weeks' mean WI-NRS scores as covariates. There was one imputed dataset. The minimum imputed value was 0 and the maximum imputed value was 10. The maximum number of iterations to impute values in the specified range was 1 million. The SAS code from study CR845-CLIN3102 follows. Study CR845-CLIN3103 included an additional covariate for geographical region.

```
proc mi data=step1 seed=2985729 nimpute=1 MAXIMUM=10 MINIMUM=0 out=step2 minmaxiter=1000000 ;
monotone reg;
      var aimblfln smcblfln w1 w2 w3 w4 w5 w6 w7 w8 w9 w10 w11 w12 w13 ;
by trt01pn;
run;
```

- ***•How different types of variables are handled (e.g., binary, categorical, non-normal distribution)***

The randomization stratification factors were binary or categorical and were coded as integers prior to their use in imputing missing weekly WI-NRS scores. The values to be imputed were means and non-normality was not a concern.

- ***•Results for complete cases versus multiple imputation together with discussion of differences/interpretation of this***

Analysis Type		Placebo (N=189)	CR845 (N=189)	Odd Ratio
	Observed data [a]	30.9% (51/165)	52.2% (82/157)	
Primary	MI with MAR	27.6% CI: 20.2%, 36.6%	51.0% CI: 42.9%, 58.9%	3.31 CI: 1.67, 6.57
Sensitivity	Early discontinuation as non-responder [b]	26.0% CI: 19.0%, 34.5%	44.6% CI: 35.4%, 54.2%	2.29 CI: 1.46, 3.60
Sensitivity	MI with MNAR	27.6% CI: 20.2%, 36.4%	47.0% CI: 37.1%, 57.3%	2.33 CI: 1.47, 3.71
Sensitivity	Tipping point [c]	29.1% CI: 21.5%, 38.1%	42.8% CI: 33.7%, 52.4%	1.82 CI: 1.16, 2.86

OR = odds ratio; MI = multiple imputation; MAR = missing at random; CI = (95%) confidence interval; MNAR = missing not at random

Estimated proportions and odds ratios are from a logistic regression model with terms for treatment group, baseline WI-NRS score, randomization stratification factors (use of anti-itch medication during the week prior to randomization and presence of specific medical conditions). Interim analysis and post-interim analysis results were combined to generate an adjusted overall estimate using Lawrence, Hung and Cui, Hung and Wang methodology.

a Counts and percentages are based on non-missing data. No model was fit.

b Subjects who discontinued study drug were imputed as non-responders.

c A tipping point was not reached.

Discussion:

Analysis of the primary endpoint, ≥3-point improvement in from baseline to Week 12 in weekly mean WI-NRS, was conducted using multiple imputation assuming a missing at random mechanism. This analysis showed a statistically significant effect

of CR845 relative to placebo. This conclusion was confirmed by a conservative sensitivity analysis in which subjects who discontinued were analyzed as non-responders. The conclusion was also confirmed by an analysis that used imputed data from a pattern mixture model, using a missing not at random mechanism, to confirm the robustness of the MAR assumption in the primary analysis. In a third sensitivity analysis, which used a tipping point method, none of the scenarios resulted in a reversal of the conclusion that CR845 is superior to placebo, even with additional shifts (to 6.5 rather than the planned limit of 5.0).

Analysis Type		Placebo (N=236)	CR845 (N=237)	Odd Ratio
	Observed data [a]	37.2% (77/207)	49.7% (95/191)	
Primary	MI with MAR	42.2% CI: 32.5%, 52.5%	54.0% CI: 43.9%, 63.9%	1.61 CI: 1.08, 2.41
Sensitivity	Early discontinuation as non-responder [b]	37.2% CI: 27.8%, 47.6%	43.7% CI: 33.4%, 54.7%	1.31 CI: 0.89, 1.94
Sensitivity	MI with MNAR	39.9% CI: 30.6%, 50.1%	50.7% CI: 41.2%, 60.1%	1.55 CI: 1.05, 2.28
Sensitivity	Tipping point [c]	41.8% CI: 31.9%, 52.3%	51.3% CI: 42.0%, 60.5%	1.47 CI: 0.98, 2.19

OR = odds ratio; MI = multiple imputation; MAR = missing at random; CI = (95%) confidence interval; MNAR = missing not at random

Estimated proportions and odds ratios are from a logistic regression model with terms for treatment group, baseline WI-NRS score, randomization stratification factors (geographical region, use of anti-itch medication during the week prior to randomization, and presence of specific medical conditions). Interim analysis and post-interim analysis results were combined to generate an adjusted overall estimate using Lawrence, Hung and Cui, Hung and Wang methodology.

a Counts and percentages are based on non-missing data. No model was fit.

b Subjects who discontinued study drug were imputed as non-responders.

c The tipping point was reached when the shift was 1.00.

Discussion:

Analysis of the primary endpoint, ≥3-point improvement in from baseline to Week 12 in weekly mean WI-NRS, was conducted using multiple imputation assuming a missing at random mechanism. This analysis showed a statistically significant effect of CR845 relative to placebo. The conservative sensitivity analysis in which subjects who discontinued were analyzed as non-responders did not reach statistical significance. Results of the primary analysis were confirmed by an analysis that used imputed data from a pattern mixture model, using a missing not at random mechanism, to confirm the robustness of the MAR assumption in the primary analysis. In a third sensitivity analysis, which used a tipping point method, the imputed values in the active treatment group were decremented by 1 unit (on a 0 to 10 scale) to reverse the conclusion of the primary analysis.

Clarification questions

In conclusion, the several sensitivity analyses confirm the conclusions of the primary analyses, which were conducted according to the intent-to-treat principle. The MAR assumption was shown to be robust. The tipping point analysis showed that the primary analysis was resistant to substantial perturbations of the missing data imputation algorithm.

A 28. Priority question: The covariates used in the logistic regression analysis were: trial group, baseline WI-NRS score, baseline use of antipruritic medication, and history of prespecified medical conditions in KALM-1 plus region in KALM-2.

a) Please provide the rationale and validity of using these variables.

Please see answer to 26c.

b) Please discuss if other baseline characteristics were considered for use in the logistic regression models, such as gender, race and age (also see question A 23 on subgroup analysis).

The subgroup analyses based on gender, race and age were planned to be conducted in the integrated summary of efficacy (ISE) rather than at the study level. The sample size for the pooled data was determined to be more appropriate to ensure the validity of these analyses.

A 29. Priority question: Please provide a formal presentation of the logistic regression analysis.

- ***•Clear definition of response and predictor variables as used in the analysis***

The primary endpoint was analyzed using logistic regression. The dependent variable (outcome) was binary:
≥3-point improvement in mean weekly WI-NRS from baseline to Week 12 or
<3 point improvement in mean weekly WI-NRS from baseline to Week 12.
The predictor variables were treatment (placebo or CR845), baseline weekly mean WI-NRS score, randomization stratification factors (use of anti-itch medication during

the week prior to randomization, presence of specific medical conditions, and, in CR845-3103, geographical region).

- ***•Description of how variables were selected for inclusion in the model (e.g., pre-defined significance level in univariate analyses)***

The variables to be included in the logistic model were specified a priori in the study protocol and statistical analysis plan. There was no additional selection or de-selection of variables in the model.

- ***•Software used and modelling/selection methods (e.g., single step, forward selection, backward elimination)***

The logistic regression was run with SAS software using proc genmod. No variable selection methods were employed.
A logistic regression model was fit to each of the imputed datasets. SAS software proc mianalyze was used to combine the estimates of the response rates, odds ratio, and p-value from each of the logistic regressions.

- ***•Table of statistics showing model output (regression coefficient, standard error, p-value and associated statistics such as z-score) for the intercept and each predictor variable, with definition of the reference value for each predictor variable***
- ***•Odds ratio (OR) estimates for each predictor variable***
- ***•Interpretation of relationship between each predictor variable and the response variable assuming other variables held constant (e.g., quantity of increase/decrease of estimated value for a 1-unit increase in the predictor variable)***

The objective of the primary analysis was to estimate the treatment effect, rather than to develop a predictive model. To that end, the regression coefficients from the logistic regression on each imputed dataset were not combined.

- ***•Adjusted (for all relevant predictors) and unadjusted overall OR estimates***

Please see the tables ‘Subjects with ≥3-point improvement from baseline to Week 12 in WI-NRS, ITT population (CR845-CLIN3102)’ and ‘Subjects with ≥3-point

improvement from baseline to Week 12 in WI-NRS, ITT population (CR845-CLIN3103)' in the response to A27 for this information.

A 30. Priority question: In the effectiveness conclusions the company states that "At Week 12, a majority of the subjects reported at least a 3-point (73.7%) or 4-point (59.3%) improvement from baseline in the weekly mean of the daily 24-hour WI-NRS score, which was previously established as a clinically meaningful threshold for this patient population (Vernon et al., 2021)." (p. 85). Please report how these data were calculated. The referenced abstract does not contain the reported results.

'At Week 12, a majority of the subjects reported at least a 3-point (73.7%) or 4-point (59.3%) improvement from baseline in the weekly mean of the daily 24-hour WI-NRS score' is referring to the results taken from the KALM trial. The reference to a 3 (or more) point improvement being a clinically meaningful threshold is taken from Vernon et al., 2021 page 1133: *'These analyses demonstrated that a reduction of ≤3 points on the WI-NRS marks an appropriate threshold for defining a clinically meaningful change in pruritus in patients with CKD-aP.'*

A 31. Priority question: Table 75 in Appendix E of document B reports on the subgroup analysis for KALM-2, as n=246 in the placebo arm and n=247 in the Difelikefalin arm which is different from the n=236 and n=237 reported in the results of KALM-2. In addition, the subjects not receiving anti-itch medication at baseline don't add up. Please correct the table accordingly.

This was a typo, and has been updated in Document B**.**

A 32. Priority question: The interim analysis was built in the study protocol as a means of sample size re-estimation, should it be required. Regarding KALM-1, the CSR (3102) states that "… ***there were no changes to the original enrolment target of 350 subjects"*** **(p. 82). Nevertheless, the primary efficacy analysis was conducted separately for interim analysis and post-interim analysis subjects.**

a) Please provide the rationale for executing a "*separate*" interim analysis.

The interim analysis was performed to inform the IDMC recommendation that no change was required to the original enrolment target. This triggered the prespecified requirement for the multiple imputation approach and logistic regression to be implemented independently for subjects contributing data to the interim analysis and subjects contributing data following the interim analysis. The primary analysis was also conducted separately for interim analysis subjects and post-interim analysis subjects to evaluate the potential impact of the interim analysis on the properties of statistical inference at the end of the study.

b) Please provide full results without splitting of the data with respect to interim versus post-interim status which are partially reported in Table 21 of document B.

Splitting of the data with respect to interim and post-interim status was only applied to the primary efficacy variable i.e. the Worst Itching Intensity Numerical Rating Scale. All other results are presented without splitting the data.

A 33. Priority question: In section B.2.6, pooled efficacy results are reported for all four KALM-1 and KALM-2 studies (double-blinded and open-label extension). Please provide the rationale, methods and analysis of how pooling was executed.

KALM 1 and 2 were similarly designed studies with similar inclusion and exclusion criteria and similar endpoints. In both studies, significantly greater proportions of participants in the difelikefalin group achieved ≥3- and ≥4-point reductions in weekly means of daily Worst Itching Intensity Numerical Rating Scale (WI-NRS) scores versus the placebo group. Pooled data from the KALM-1 and KALM-2 studies was analysed to obtain a combined estimate of the treatment effects of difelikefalin in HD participants with moderate to severe pruritus, including QoL endpoints. The statistical methods of how pooling was executed are included in the, now fully published, Topf et al. manuscript and are as follows.

Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants. Differences between placebo and difelikefalin were analysed using a logistic regression model containing terms for the treatment group, baseline WI-NRS score,

use of an anti-itch medication during the week before randomization, presence of specific medical conditions, and geographic region. For the analysis of the proportions of participants who achieved ≥3-point or ≥4-point reductions in the weekly mean WI-NRS scores, missing weekly WI-NRS scores were imputed by multiple imputation under a missing-at-random assumption. Participants who reported <4 daily WI-NRS scores at week 12 or who discontinued treatment early were considered non-responders in the analysis of the complete WI-NRS response. Proportions of participants achieving a ≥5-point improvement in the 5-D Itch total score and a ≥15-point improvement in the Skindex-10 total score were analysed without imputation for missing values. Proportions of participants achieving a ≥5-point improvement in 5-D Itch total score are reported for the pooled population during the placebo-controlled, double-blind period (12 weeks) and the open-label extension period (up to 52 weeks).

Continuous efficacy endpoints were analysed by a mixed model for repeated measures, with terms for treatment, visit, treatment-by-visit interaction, baseline score, use of an anti-itch medication during the week before randomization, the presence of specific medical conditions, and geographic region. An unstructured covariance structure was applied to model the within participant errors. Missing values were not imputed. The mean improvements from baseline in 5-D Itch total score are reported for the pooled population during the placebo-controlled, double-blind period (12 weeks) and the open label extension period (up to 52 weeks).

The subgroup analyses of ≥3-point and ≥4-point reductions from baseline in the weekly mean WI-NRS scores were performed using the same methodology as that employed for the full intent-to-treat population.

A 34. The double-blinded period of KALM-1 was followed by a 2-week discontinuation period before the OLE phase started, during which the patients were evaluated for signs of physical dependence.

a) Please discuss why the same design was not also followed by KALM-2.

Results from KALM-1 were considered robust enough not to warrant interrupting patient treatment in KALM-2

b) Please discuss whether the 2-week discontinuation period had an effect on the efficacy and safety results.

Comparing the results from KALM-1 with KALM-2 suggests there is no obvious effect of the 2-week discontinuation period on the efficacy and safety of difelikefalin over the 64 weeks of treatment.

A 35. Priority question: It is unclear if the patient-reported, single item Worst Itch Numerical Rating Scale (WI-NRS) was used to determine the severity of chronic kidney disease (CKD)-associated pruritus across all relevant trials.

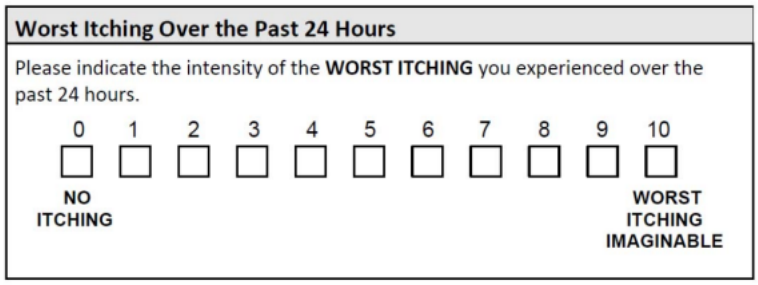
a) Please clarify if this was so or explain the measure used to determine the severity of CKD associated pruritus.

All trials measured itch severity using both the WI-NRS and 5-D itch scores.

b) Please provide information on the methodology of the WI-NRS score

The Worst Itching Intensity Numerical Rating Scale (WI-NRS) is a simple, single-item patient-reported outcome measure to assess the intensity of the worst itching a patient has experienced over the past 24 hours, as described in 'Clinically Meaningful Change in Itch Intensity Scores: an Evaluation in Patients with Chronic Kidney Disease associated Pruritus' (Figure 1) (8).

Figure 1 Worst Itching Intensity Numerical Rating Scale (WI-NRS)



Adapted from Phan NQ, Blome C, Fritz F, Gerss J, Reich A, Ebata T et al. Assessment of pruritus intensity: prospective study onvalidity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronicpruritus. Acta Derm Venereol. 2012;92:502-507.

c) Please discuss the validity of this scale in capturing the severity of CKD-associated pruritus.

Various pieces of literature (9-11) have found WI-NRS to be a reliable, reproducible,

and valid measure of itch intensity in moderate-to-severe CKD-aP patients, and therefore a reasonable choice.

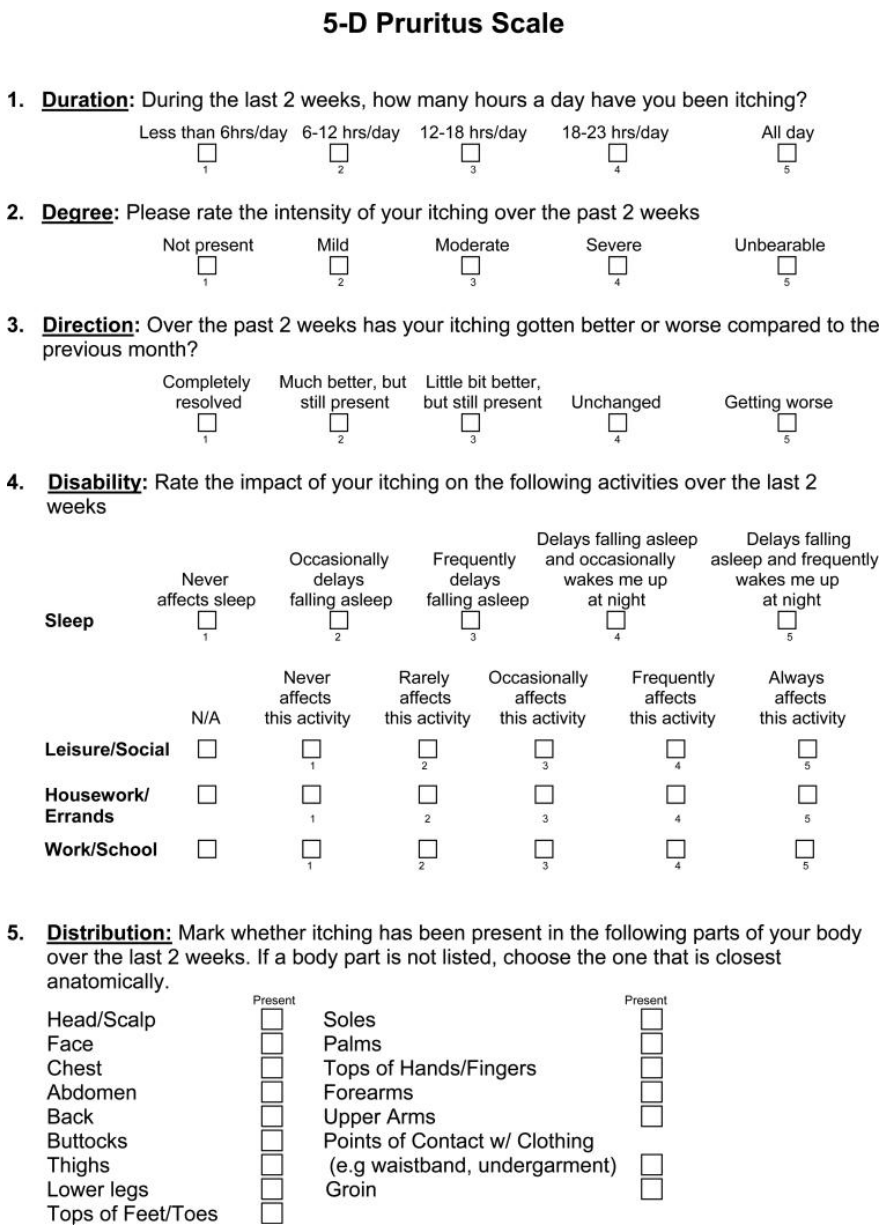
d) Please justify the choice of >4 points weekly mean, as a benchmark for moderate-to-severe pruritus.

The paper 'A Randomized Controlled Trial of Difelikefalin for Chronic Pruritus in Haemodialysis Patients' states *'Itching severity scores collected via the WI-NRS have been categorized in the literature (12) as mild (<4), moderate (≥4 to <7), or severe (≥7).' (13)*

A 36. Priority question: Please also provide evidence of the methodology and validation of the 5-D Itch score outcome used in the CS.

The 5-D Itch scale is a multidimensional questionnaire which assesses itch severity and itch-related quality of life over the previous 2 weeks. The questionnaire covers 5 dimensions of itch, including the degree, duration of itch/day, direction (improvement/worsening), disability (impact on activities such as work), and body distribution of itch (Figure 2). The total 5-D Itch scale score ranges from 5 to 25, with higher scores indicating worse responses. The scale has been validated in patients with chronic pruritus, including haemodialysis patients, and has been shown to be sensitive to changes in pruritus over time(14). Additionally, with limited options for itching scales (Appendix N: Clinical Opinion and consensus report), 5D-Itch is both commonly used and produces valid and reproducible results. It is therefore an appropriate choice for measuring itch in CKD-aP patients for this submission(14, 15).

Figure 2 5-D Itch scale



5-D Pruritus Scale

1. **Duration**: During the last 2 weeks, how many hours a day have you been itching?

Less than 6hrs/day	6-12 hrs/day	12-18 hrs/day	18-23 hrs/day	All day
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

2. **Degree**: Please rate the intensity of your itching over the past 2 weeks

Not present	Mild	Moderate	Severe	Unbearable
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

3. **Direction**: Over the past 2 weeks has your itching gotten better or worse compared to the previous month?

Completely resolved	Much better, but still present	Little bit better, but still present	Unchanged	Getting worse
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

4. **Disability**: Rate the impact of your itching on the following activities over the last 2 weeks

	Never affects sleep	Occasionally delays falling asleep	Frequently delays falling asleep	Delays falling asleep and occasionally wakes me up at night	Delays falling asleep and frequently wakes me up at night
Sleep	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

	N/A	Never affects this activity	Rarely affects this activity	Occasionally affects this activity	Frequently affects this activity	Always affects this activity
Leisure/Social	☐	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Housework/ Errands	☐	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Work/School	☐	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

5. **Distribution:** Mark whether itching has been present in the following parts of your body over the last 2 weeks. If a body part is not listed, choose the one that is closest anatomically.

	Present		Present
Head/Scalp	☐	Soles	☐
Face	☐	Palms	☐
Chest	☐	Tops of Hands/Fingers	☐
Abdomen	☐	Forearms	☐
Back	☐	Upper Arms	☐
Buttocks	☐	Points of Contact w/ Clothing	
Thighs	☐	(e.g waistband, undergarment)	☐
Lower legs	☐	Groin	☐
Tops of Feet/Toes	☐		

A 37. The number of participants recruited to the double-blind phase of KALM-1 is shown as N=378 in both Table 3 and Figure 21 and N=377 (189 + 188 = 377) in Table 7 of Document B. Please provide the correct number of participants (overall and per treatment arm) or explain the discrepancy.

Figure 21 in the document provides an explanation of the discrepancy between the ITT population (378) and those who received the allocated intervention (378). There was one patient in the placebo group who did not meet the entry requirement and therefore did not receive an allocated intervention and was therefore excluded from

the double-blind safety population in Table 7. An updated clarification on this has also been included in Table 3 of the amended Document B.

A 38. Table 7 of Document B shows a blank cell for the group receiving placebo for the row entitled "*Specific medical condition?*". Please provide these data.
Table 7 has been updated in Document B.

A 39. The number of participants recruited to the double-blind phase of KALM-2 is discrepant between Table 5 (N=474), Figure 23 (N=473) and Table 8 (N=471; 236 + 235 = 471) in Document B. Please provide the correct number of participants (overall and per treatment arm) or explain the discrepancy.

Figure 23 in the document provides an explanation of the discrepancy between the ITT population (473) and those who received the allocated intervention (471). There were two patients in the difelikefalin group who did not receive the allocated intervention and were therefore excluded from the double-blind safety population in Table 8. There was a typo in Table 5. Further clarification on the double-blind safety population in Table 5 has been updated in Document B.

A 40. There seem to be some instances of p-values being the only information provided about estimation. For example, at the top of page 63 of document B in the CS it is stated that: "*the findings for the per protocol population were also in favour of difelikefalin and statistically significant (p<.001)*", with no further information given. Even if the full estimate is in the CSR, full results should be provided in the CS. Please provide 95% CIs for all between-arm estimates in the CS.

Please see updates in Document B (Section B2.6)

A 41. The end of page 66 of Doc B discusses discontinuation in relation to the KALM-1 OLE study, stating that it was not due to lack of efficacy, but no positive reasons for discontinuation are listed. Please provide reasons for discontinuation.

Reason for discontinuation include adverse events (26 patients), withdrawn consent (17 patients), non-compliance (5 patients) and lost to follow-up (5 patients). Please see 'Appendix V - KALM-1 OLE - Reason for discontinuation' for further details.

A 42. Page 96 of document B mentions early termination of KALM-2 saying that this was due to an administrative decision and not related to efficacy or safety. Please provide further information about the administrative reason for the early termination.

The open-label portion of the KALM-2 study was halted by the sponsor to enable analysis and reporting of the results to be completed for inclusion in the marketing authorisation application to the FDA

A 43. Table 74 on page 182 of document B shows numbers with 'medical conditions' at baseline per treatment arm, but all conditions have been lumped together. Given that different conditions may have very different effects on outcome, this is potentially misleading. Please disaggregate these medical conditions.

The specific medical conditions are of interest as they are typically associated with the pharmacology of kappa opioid receptor agonists. Stratification was performed to ensure balanced groups for those pre-existing medical conditions so as not to confound the assessment of the safety profile of difelikefalin. However, as they were not identified as potential modifiers of treatment response no analysis of the individual conditions has been performed.

A 44. Please explain how the data from the CLIN3105 trial were used in the submission. For example, how did these data supplement the data from the KALM-2 and KALM-2 trials?

CLIN3105 was included in the submission, but not included in the economic model because it did not contain a relevant comparator arm.

CLIN3105 gathered data on sleep quality using the Sleep Quality Questionnaire. CKD-aP patients often report restless and poor-quality sleep as a result of their itch, causing considerable burden on quality of life(10, 16); effect on sleep quality is therefore considered an important outcome of difelikefalin. This outcome was not

investigated in both KALM-1 and KALM-2, so supplementary data from CLIN3105 is used.

Furthermore, CLIN3105 provides real world evidence for difelikefalin with patients in full knowledge of the treatment, as opposed to a blinded trial.

For these reasons, CLIN3105 was included in the submission to supplement data provided by KALM-1 and KALM-2.

Adverse events

A 45. The CS states that, “Of the 143 who reported a treatment-emergent adverse event (TEAE), 68 patients (30.6%) had a maximum severity of mild, 56 (25.2%) had a maximum severity of moderate, and 19 (8.6%) had a maximum severity of severe.”

b) Please provide the scale used to judge the severity of TEAEs.

The Investigator assessed the severity (i.e., intensity) of each adverse event (serious and non-serious) reported during the study based on his/her clinical judgment. The severity of each adverse event was assigned to one of the following categories:

Mild: Transient, requires no special treatment, is easily tolerated by the patient, causes minimal discomfort, and does not interfere with the patient’s daily activities

Moderate: Introduces a level of inconvenience or concern to the patient that may interfere with daily activities, but usually is ameliorated by simple therapeutic measures

Severe: Interrupts a patient’s usual daily activity and requires systemic drug therapy or other treatment

c) Please provide a list of TEAEs by severity.
Please see Table 3 below

Table 3 Incidence of TEAEs During the Treatment Period by MedDRA System Organ Class and Severity Population: Safety

System Organ Class	Severity	CR845 (N=222)
Number of subjects with an event (N=143)	Mild	68 (30.6%)
	Moderate	56 (25.2%)
	Severe	19 (8.6%)
Blood and lymphatic system disorders (N=7)	Mild	4 (1.8%)
	Moderate	2 (0.9%)
	Severe	1 (0.5%)
Cardiac disorders (N=19)	Mild	10 (4.5%)
	Moderate	5 (2.3%)
	Severe	4 (1.8%)
Ear and labyrinth disorders (N=3)	Mild	2 (0.9%)
	Moderate	1 (0.5%)
	Severe	0
Endocrine disorders (N=1)	Mild	0
	Moderate	1 (0.5%)
	Severe	0
Eye disorders (N=4)	Mild	4 (1.8%)
	Moderate	0
	Severe	0
Gastrointestinal disorders (N=36)	Mild	20 (9.0%)
	Moderate	14 (6.3%)
	Severe	2 (0.9%)
General disorders and administration site conditions (N=20)	Mild	11 (5.0%)
	Moderate	6 (2.7%)
	Severe	3 (1.4%)
Hepatobiliary disorders (N=1)	Mild	0
	Moderate	1 (0.5%)
	Severe	0
Infections and infestations (N=49)	Mild	27 (12.2%)
	Moderate	16 (7.2%)
	Severe	6 (2.7%)
Injury, poisoning and procedural complications (N=30)	Mild	15 (6.8%)
	Moderate	14 (6.3%)
	Severe	1 (0.5%)
Investigations (N=8)	Mild	5 (2.3%)
	Moderate	2 (0.9%)
	Severe	1 (0.5%)
Metabolism and nutrition disorders (N=20)	Mild	7 (3.2%)
	Moderate	9 (4.1%)
	Severe	4 (1.8%)
Musculoskeletal and connective tissue disorders (N=19)	Mild	9 (4.1%)
	Moderate	9 (4.1%)
	Severe	1 (0.5%)
Nervous system disorders (N=39)	Mild	23 (10.4%)
	Moderate	16 (7.2%)
	Severe	0
Product issues (N=2)	Mild	1 (0.5%)
	Moderate	1 (0.5%)
	Severe	0
Psychiatric disorders (N=5)	Mild	0
	Moderate	5 (2.3%)
	Severe	0

Renal and urinary disorders (N=3)	Mild	2 (0.9%)
	Moderate	1 (0.5%)
	Severe	0
Respiratory, thoracic and mediastinal disorders (N=21)	Mild	12 (5.4%)
	Moderate	7 (3.2%)
	Severe	2 (0.9%)
Skin and subcutaneous tissue disorders (N=9)	Mild	3 (1.4%)
	Moderate	6 (2.7%)
	Severe	0
Surgical and medical procedures (N=1)	Mild	1 (0.5%)
	Moderate	0
	Severe	0
Vascular disorders (N=21)	Mild	10 (4.5%)
	Moderate	8 (3.6%)
	Severe	3 (1.4%)

d) Please discuss AEs that lead to dose reductions, interruptions or discontinuation of difelikefalin treatment.

A total of 14 subjects (6.3%) experienced at least 1 TEAE that led to study drug discontinuation during the Treatment Period. The most common preferred term of TEAE leading to study drug discontinuation was somnolence (2 subjects [0.9%]).

Four subjects (840008011, 840012004, 840018009, 840028005) experienced TEAEs leading to study drug discontinuation that were assessed as related to study drug. These subjects experienced the following study-drug related TEAEs: somnolence (840008011 and 840012004), nausea (840018009), and dizziness (840028005). A supplementary table has been provided detailing adverse events resulting in study drug discontinuation during the study period (please see Appendix U – AEs leading to discontinuation'.

Of the 14 subjects who experienced a TEAE that led to study drug discontinuation, 2 subjects (348001001 and 840034003) had events with fatal outcomes (Table 23), and the events for the remaining subjects were reported as recovered/resolved

A 46. Table 37 presents the pooled adverse reactions results for the two double-blinded and the two open-label studies. Please specify how pooling was executed.

The placebo-controlled cohort included participants (848) from the 12-week, pivotal studies (KALM-1 and KALM-2) who received (at least 1 dose of) IV difelikefalin at 0.5

mcg/kg or placebo 3 times per week. The all-difelikefalin-exposure cohort included all participants who received 1 or more doses of IV difelikefalin at 0.5 mcg/kg for up to 64 weeks from the placebo-controlled periods of the pivotal studies (if randomized to difelikefalin) and from the open-label extension periods (up to 52 weeks) of these studies. Safety was evaluated based on adverse events (AEs) and safety assessments (ie, physical examinations, vital signs, clinical laboratory tests, and electrocardiograms). Safety analyses were summarized descriptively.

Other

A 47. Section 4.7 of the summary of product characteristics (SmPC) states that, *"has minor influence on the ability to drive and use machines."* As difelikefalin is approved for in-centre use only, please provide more details on what measures will be put in place to manage dizziness/somnolence symptoms in patients who drive in for their thrice-weekly haemodialysis appointments, most especially within their first 3 weeks of treatment.

As per normal prescribing practice, with any medication associated with potential to cause dizziness or somnolence it is expected that clinicians would advise patients in the standard way until the effect of difelikefalin on the patient's ability to drive or operate machinery is known.

A 48. The anti-asthma, leukotriene receptor antagonist (LTRA), Montelukast, was listed in Table 55 of the CS as 'established clinical management' for CKD-associated pruritus'.

a) Please provide supporting evidence to support Montelukast being an established CKD-associated pruritus intervention.

Please refer to Appendix D, G, H, and I 'Systematic Literature Review of Chronic kidney disease associated pruritus'. Evidence to support Montelukast being an established CKD-associated pruritis intervention is summarised below:

- Hercz et al., conducted a systematic literature review reviewing 92 studies assessing all topical and systemic interventions for the treatment of uraemic

itch. Here it is stated that Montelukast may slightly reduce symptoms of uraemic itch (Hercz et al., 2020)

- The European S2k Guideline on Chronic Pruritus was published in 2019 (Weisshaar et al., 2019). The guideline covers the diagnosis and management of several different types of pruritus, including hepatic/cholestatic pruritus, aquagenic pruritus and CKD-aP, According to this guideline, Montelukast at 10 mg/day is listed as a 'therapeutic option' stating that antipruritic effects in patients with CKD-aP have been demonstrated in controlled studies.

b) Please clarify whether Montelukast was a permitted concomitant trial medication.

Yes Montelukast was a permitted concomitant trial medication

A 49. Page 91 of the CS states that potential signs and symptoms of opioid withdrawal were measured with the Short Opiate Withdrawal Scale (ShOWS) and Objective Opiate Withdrawal Scale (OOWS). Please supply the relevant data.

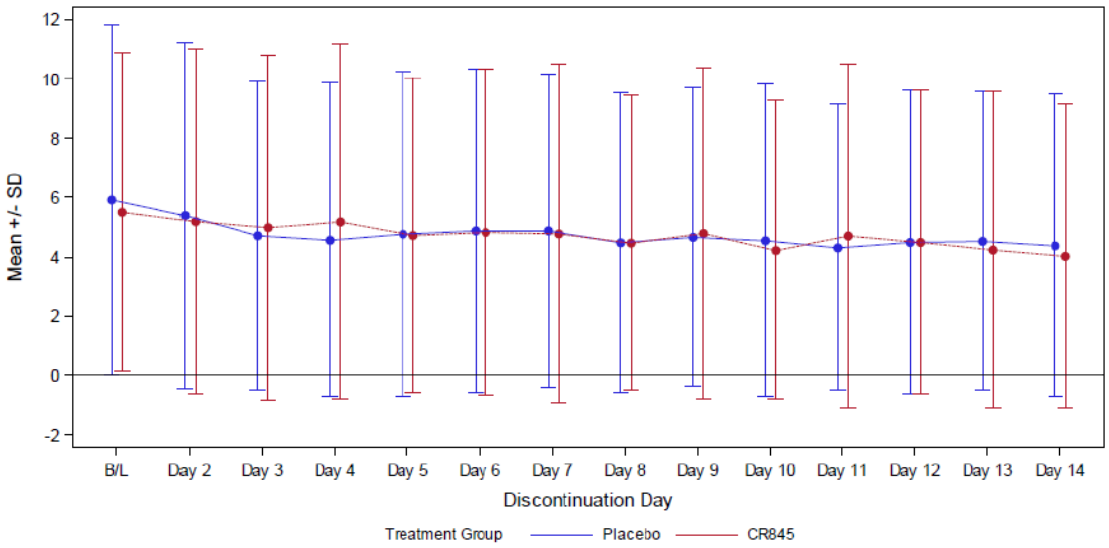
Data on both ShOWS and OOWS scores for the Double-blind Discontinuation Population of KALM-1 is summarised below. Please see the provided supplementary tables 14.3.7.5.1 - 14.3.7.6.3 for raw ShOWS and OOWS data of both the Double-blind Discontinuation Population and the Double-blind Discontinuation Safety Population of KALM-1.

ShOWS:

Figure 3 presents mean total ShOWS scores over time for the Double-blind Discontinuation Population. Both treatment groups showed a slight decrease in mean ShOWS score over time. At baseline, subjects in the difelikefalin and placebo groups reported mean ShOWS scores of 5.5 and 5.9, respectively. On Discontinuation Day 14, difelikefalin and placebo subjects reported mean ShOWS scores of 4.0 and 4.4, respectively, with mean changes from baseline of -1.1 and -1.2, respectively. The largest LS mean treatment group difference in the change in ShOWS score from baseline was 0.9, which was observed at Discontinuation Day 4

and was significant (P = .044). No other treatment group difference in change in ShOWS score from baseline was significant for Discontinuation Days 1 through 6 or for Days >6.

Figure 3 Total ShOWS Score Over Time During the Double-blind Discontinuation



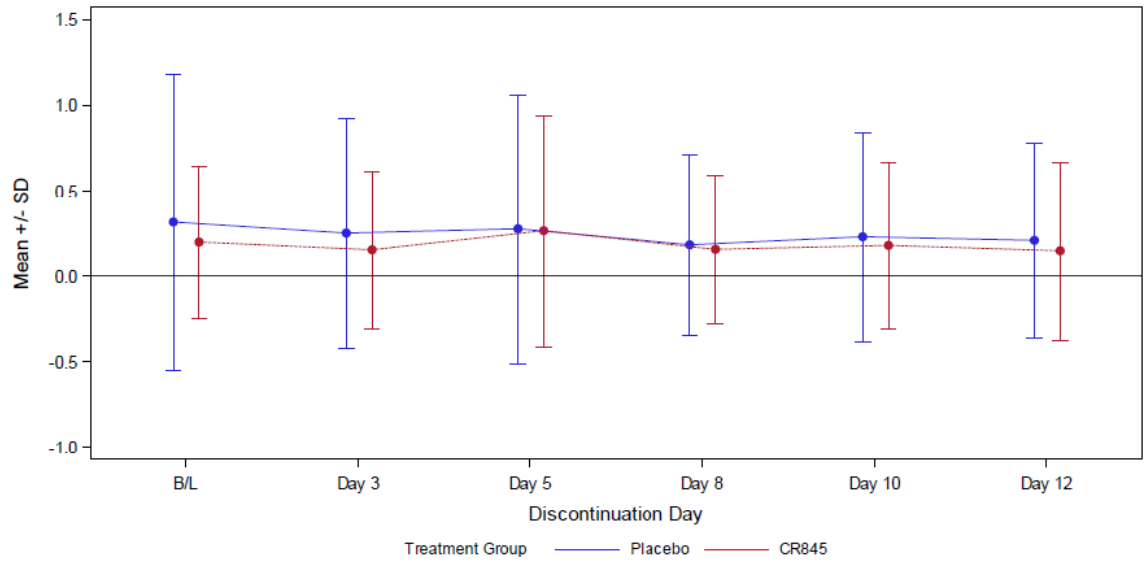
SD = standard deviation; ShOWS = Short Opiate Withdrawal Scale

Note: The mean and standard deviation for the total observed daily scores are displayed.

OOWS:

Figure 4 presents the total OOWS score over time for the Double-Blind Discontinuation Population. Mean OOWS scores were generally stable over time in both treatment groups. At baseline, subjects in the difelikefalin and placebo groups reported mean OOWS scores of 0.2 and 0.3, respectively. On Discontinuation Day 12 (the last OOWS assessment), difelikefalin and placebo subjects reported mean OOWS scores of 0.1 and 0.2, respectively, with mean changes from baseline of -0.1 in both treatment groups. The largest LS mean treatment group difference in change in OOWS score from baseline was -0.1, which was observed at Discontinuation Day 3 (P = .255). No treatment group difference in change in OOWS score from baseline was significant for Discontinuation Day 3 through Discontinuation Days >6.

Figure 4 Total OOWS Score Over Time During the Double-blind Discontinuation Period (No Imputation) – Line Graph (Population: Double-blind Discontinuation)



OOWS = Objective Opiate Withdrawal Scale; SD = standard deviation

Note: Least squares means, SEs, and 95% CIs come from an ANCOVA model fit at each time point, with treatment group and baseline (Day 85) value as a covariate.

Note: The mean and standard deviation for the total observed daily scores are displayed.

Please see 'Appendix R – Opiate withdrawal scale tables for further information'.

Clarification on cost-effectiveness data

Model Structure/Assumptions

B 1. Priority question: When explaining the reasoning behind the selection of the model structure, the CS states that this approach is deemed appropriate as it is consistent with the 3 appraisals for atopic dermatitis and 2 appraisals for CKD identified in the extended SLR. Please provide the references to these appraisals and explain how many and if other appraisals for atopic dermatitis and CKD used alternative model structures. Please provide a brief overview of the alternative model structures used in the other appraisals identified to be relevant to this one.

Due to a lack of economic evaluations in CKD-aP, disease criteria were extended to include CKD and pruritus analogues. In total, 6 health technology assessments

(HTA) were identified in the extended SLR, with 3 appraisals for atopic dermatitis, 2 appraisals for CKD, and 1 appraisal for progressive familial intrahepatic cholestasis. References to, and details of the model structure used in the appraisals considered appropriate for this evaluation are provided in the Table 4: below. Further detail is provided in Section 7 of Appendix D, G, H and I (SLR results).

Table 4: CKD and atopic dermatitis appraisal details

HTA ID	Title	Indication	Model structure
TA775	Dapagliflozin for treating chronic kidney disease [ID3866] (17)	CKD	Cohort Markov model
TA807	Roxadustat for treating anaemia in people with chronic kidney disease [ID1483] (18)	CKD	Cohort Markov model
TA681	Baricitinib for treating moderate to severe atopic dermatitis [ID1622] (19)	Atopic dermatitis	Cohort Markov model
TA534	Dupilumab for treating moderate to severe atopic dermatitis after topical treatments [ID1048] (20)	Atopic dermatitis	One-year decision tree followed by Markov model
GID-TA10856	Upadacitinib, abrocitinib and tralokinumab for dermatitis [ID3960] (21)	Atopic dermatitis	One-year decision tree followed by Markov model

Both CKD appraisals and 1 atopic dermatitis appraisal used a cohort Markov model structure, whilst the remaining 2 atopic dermatitis appraisals used a combined decision tree and Markov model structure.

The combined decision tree and Markov model structure was considered as an alternative structure that may be appropriate for the evaluation of cost-effectiveness of difelikefalin compared with established clinical management. As noted in the CS, the first 3 cycles of the Markov model reflect an initial run-in period whereby all patients eligible for difelikefalin are treated, with a clinical assessment undertaken at week 12 (end of cycle 3) to determine response to treatment. The dupilumab model structure was designed to reflect a similar pathway, using a decision tree rather than a Markov model to reflect the short-term treatment period, with a clinical assessment undertaken at week 16 to determine response to treatment. The company believe a combined decision tree and Markov models structure would result in similar outcomes.

B 2. Priority Question: The baseline characteristics applied in the model are based on the KALM trial populations, but only KALM-2

included some patients from the UK. Baseline characteristics of data from the UK Renal Registry (UKRR) seem to show some differences especially in terms of the starting age of the patients (Table 41 in the CS).

b) Please indicate if the KALM trial populations are representative of the UK population and explain the reason behind the discrepancies between the KALM data and the UKRR.

During clinical validation of the model inputs, it was highlighted that the KALM data are most appropriate given that pruritus is not regularly coded in current UK clinical practice, and therefore the UKRR data would be more reflective of the wider haemodialysis population.

c) Please include a scenario analysis in which the baseline characteristics in the model are informed from the UKRR and discuss the impact.

Table 5: Baseline demographics scenario

Scenario	Inc. costs	Inc. LYs	Inc. QALYs	ICER
Revised base case Using KALM data for demographics	[redacted]	0.06	[redacted]	£23,277
Scenario Using UKRR data for demographics	[redacted]	0.06	[redacted]	£23,392

d) Please specify the number of patients used to inform baseline characteristics in Table 41 of the CS.

Data on the baseline characteristic in Table 41 of the CS for the KALM trials for age, proportion male, and weight were based on the all-difelikefalin-exposure cohort (n=1,306) as presented in Fishbane et al., (22). The mean length of time on dialysis was estimated using the pooled KALM-1 and KALM-2 patient level data which included 848 observations (PBO n = 424 and DFK n = 424). For reference, Fishbane et al., report the median length of time on dialysis to be 4.0 (IQR = 5.2).

B 3. In section B.3.2 a 7-day run-in period during the week prior to randomisation has been reported for KALM-1 and KALM-2 trials,

followed by a 12-week double-blind treatment period. In section B.3.2.4. a 4-week cycle length is used for the first 3 cycles, which is labelled as the 'run-in' period, followed by a 52-week cycle length used from Cycle 4 onwards. Please confirm whether the term 'run-in' had different definitions in the clinical effectiveness and cost-effectiveness sections of the submission.

In section B.2.3 of the CS, the company summarise the structure of the KALM-1 and KALM-2 trial designs as:

Both studies included a double-blind phase and OLE phase. The double-blind phase consisted of a screening visit, a 7-day run-in period during the week prior to randomisation and a 12-week double-blind treatment period where difelikefalin was evaluated relative to placebo.

The purpose of the 7-day run-in period was to confirm that each subject did have moderate-to-severe pruritus, and to establish baseline itch intensity.

In section B.3.2.4 of the CS, the company summarise the model structure as:

A 4-week cycle length is used for the first 3 cycles (the 'run-in' period) with a 52-week cycle length used from Cycle 4 onwards…

The EAG is correct in noting that the term 'run-in' is used to describe different time periods in the clinical and cost-effectiveness sections of the submission and do not reflect the same data periods within the trial. The KALM trial run-in period reflects the 7-day period prior to randomisation to confirm eligibility for the double-blind period of the trials. The model run-in period refers to the short-term treatment decision period (first 3 model cycles) and refers directly to the double-blind 12-week period of the KALM trials.

B 4. In the base case analysis, only patients with moderate-to-severe CKD-aP are considered. The CS states that a scenario analysis is presented considering only patients with severe or very severe CKD-aP at baseline (Table 43 of the CS). However, this scenario is not presented in the scenario analysis section (B.3.9.3). Please provide the additional scenario analysis. Table 43 of the CS also provides the distribution of

patients at model entry based on the WI-NRS scores. Please clarify the purpose of presenting the patient distribution at model entry based on the WI-NRS scores.

This is a reporting error in the submission. The relevant analysis was provided as a subgroup analysis in section B.3.11.

The patient distribution at model entry based on the WI-NRS scores has been removed from the submission.

B 5. Priority Question: In section B.3.2.2 it is noted that the KALM trials did not directly include any comparator treatments, although patients using anti-itch medication at baseline were allowed to continue doing so. Please specify explicitly what anti-itch medications were used in both trial arms and on what percentage of patients in each arm.

Please see answer to question A.20.

B 6. Priority Question: The company presents no clinical evidence from the KALM-1 and KALM-2 trials showing that DFK treatment may potentially improve survival in moderate-to-severe CKD-aP patients, but the economic model accounts for DFK survival benefit due to DFK treatment. Please explain through what causal relationship a reduction in itching score due to treatment with difelikefalin would be expected to reduce mortality in these patients.

The increased mortality risk data for the very severe, severe, and moderate CKD-aP population used in the model is informed by Sukul et al., 2021 (23). They report that extreme pruritus is an independent predictor of all-cause and case-specific mortality when adjusting for influential confounders such as patient demographic and clinical characteristics. The authors acknowledge that the possible bidirectionality of the relationship between pruritus and cross-sectional patient-reported outcomes limits the inferences that can be made and does not allow conclusions about cause-effect relationships, however increased depression, missed dialysis sessions, poor sleep

quality, and skin lesions susceptible to infection are outcomes that could mediate the relationship between extreme pruritis and mortality.(24, 25)

B 7. Priority Question: The efficacy data used in the model focused on a 5-point reduction in the total 5-D Itch score from baseline as a clinically meaningful improvement in patients with CKD-aP undergoing haemodialysis. The company states that the "*scale has been validated in patients with chronic pruritus, including haemodialysis patients, and has been shown to be sensitive to changes in pruritus over time (Elman et al., 2010)*". However, Elman et al. do not indicate what size reduction could be considered clinically meaningful. Please provide further support on the reasoning of the 5-point clinically meaningful cut-of value and compare to other relevant appraisals/studies that used the same rating scale of itching.

As reported in the KALM-1 and KALM-2 clinical study report for the open-label extension phase (26, 27), psychometric analysis of the Phase 2 study CR845-CLIN2101 dataset showed that a 5-point decrease in the total 5-D Itch score from baseline represented a clinically meaningful improvement to the subjects.

The phase 2 study to assess the efficacy and safety of difelikefalin over an 8-week treatment period in haemodialysis patients with moderate-to-severe pruritus, demonstrated that improvement in itch-related quality-of-life measures were highly correlated with a reduction in the WI-NRS score at week 8, with a Pearson coefficient (r) of 0.71 for the 5-D itch total scores ($P < 0.0001$). A 5-point reduction in 5-D itch was associated with a 4-point reduction in WI-NRS, which in turn has been defined as a clinically important reduction in the severity of CKD associated pruritus.

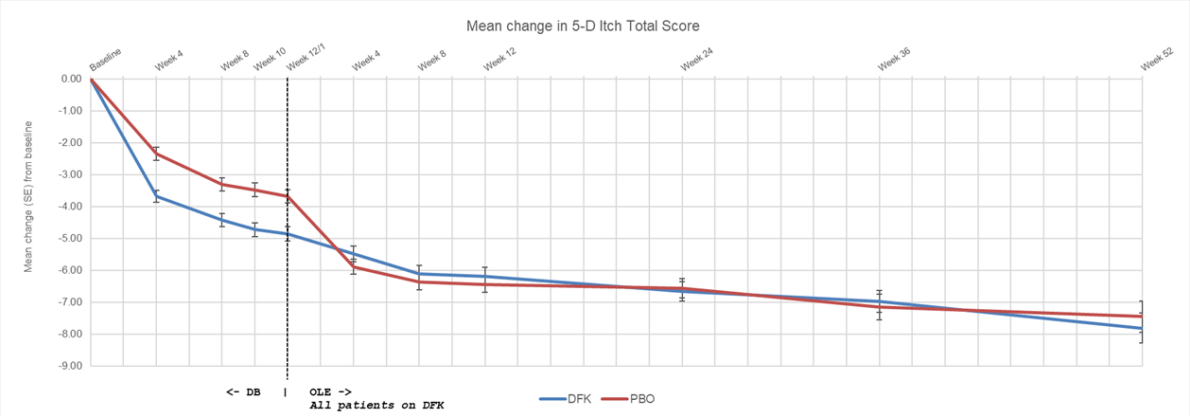
No literature searching was conducted to identify studies or appraisals that used the 5-D Itch Scale.

B 8. Priority Question: It is mentioned that "*as no data was collected beyond the 52-week OLE phase, in the base case, efficacy remains unchanged after Week 64*". Please discuss the validity of

this assumption that beyond the 52-week efficacy remains unchanged

From March to April 2022, a modified Delphi panel was conducted to collect expert opinion from eight consultant nephrologists from across England who treat patients with CKD-aP (results presented in Appendix N: Clinical opinion and consensus report). When asked about the potential waning of difelikefalin over time, participants were unable to comment on any suspected waning effect. During the clinical validation of model inputs, it was noted that the data (referring to Figure 5) would support the assumption of no drop-off in effect.

Figure 5: Pooled KALM-1 and KALM-2 data - mean change in total 5-D Itch score



B 9. In section B.3.3.1.4 the company states that "in line with the clinical opinion that the placebo affect would wane over time, a waning effect is applied in the established clinical management arm equal to a 5% probability for patients to gain a health state (deteriorate) each year following Week 64". However, in the text above it is mentioned that Soro et al (2022) shows that "over the 18 months, a general trend was observed, with the prevalence of moderate pruritus remaining stable, mild/none increasing, and severe/overwhelming decreasing".

a) Please provide the complete source of Soro et al (2022), as the current version in the reference package is only an abstract and the patterns mentioned above cannot be validated.

Please see Soro et al., 2022 in 'Reference pack (1)' for the poster 'A methodological approach to assess the economic value of difelikefalin to treat chronic kidney disease associated pruritus (CKD-aP)'.

b) Apart from the clinical expert opinion, is there any further support around the waning pattern of the ECM arm? Please discuss the validity of this assumption.

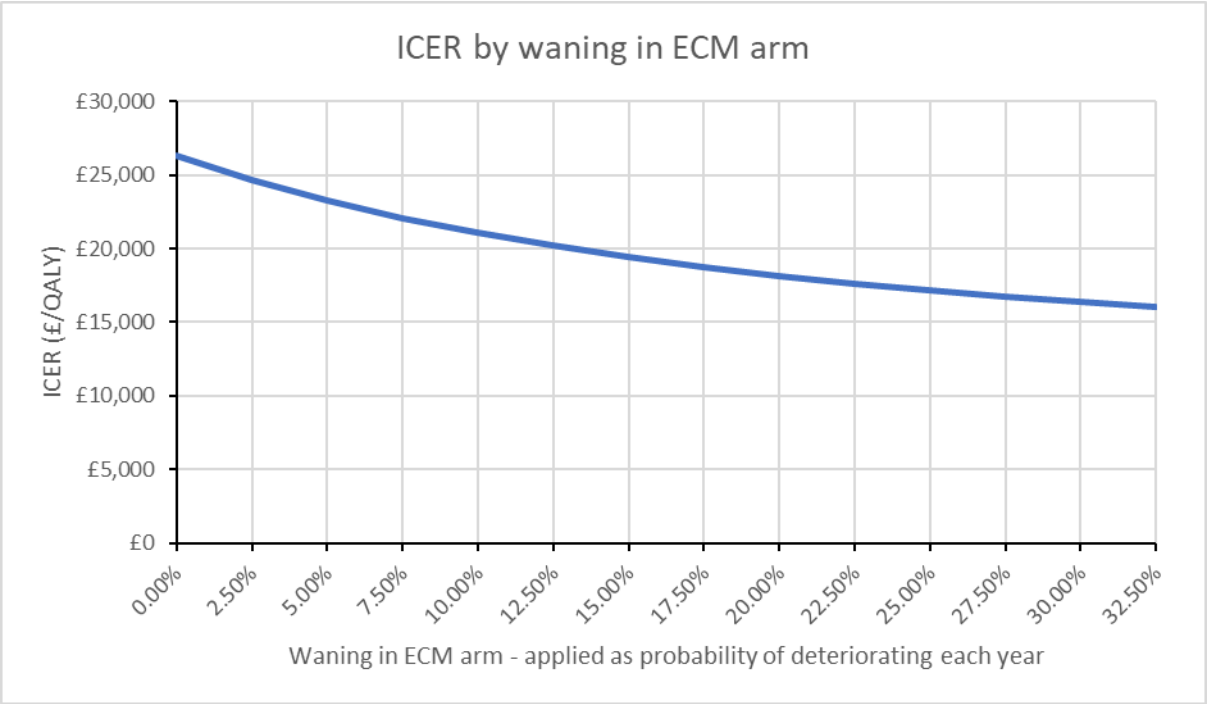
The waning effect was applied in the model to mitigate the long-term impact of the improved outcomes observed for the placebo arm in the KALM trials. Clinical experts advised that maintained long-term benefits of placebo would not be observed in clinical practice and lacked clinical face validity. Hence, improvement in itch for the ECM group (informed by efficacy estimates from the placebo arm of the trials) may overestimate those observed in standard clinical practice. However, it is unclear if similar trial effects would affect the outcomes observed for the difelikefalin treatment arm. No additional real-world evidence was identified that could further inform the extrapolation and long-term outcomes for patients receiving ECM.

c) Please justify why the value of 5% was used to model reduction in efficacy of ECM.

As noted in answer b), there was no quantitative evidence that could be used to inform a value for waning in efficacy for the ECM model arm. Instead, a simple assumption was made to reflect the probability for a patient to gain a health state each year.

Figure 6 plots the change in the ICER for the waning effect applied to the ECM arm.

Figure 6: Change in ICER by waning in ECM arm



B 10. In section 3.3.1.5 it is stated that "Multiple imputation was used to fill in missing data values in the patient-level data set for the total 5-D Itch scale scores. This was carried out in R. …. which was verified via a Missing Completely at Random test in the Misty package (28). The MICE package was used to perform multiple imputation with the Predictive Mean Matching approach; the number of imputations and maximum iteration were set to 5 and 40, respectively (29)." However, in the SAP for the KALM studies it was mentioned that the multiple imputation was done using the procedure MI in SAS, with 20 imputations. Please explain the discrepancy between the text in the submission and the SAPs.

The MI reported in the SAPs for the KALM studies was conducted separately to the MI conducted during the modelling to estimate transition matrices.

The answer to question A25 provides further detail on the MI analysis reported in the SAPs for the KALM studies.

The MI conducted and reported in the company submission (section B.3.3.1.5) was used to fill in missing data values that were present in the patient-level data set for the total 5-D Itch scores. Missing values were estimated for the full intent-to-treat

(ITT) population up to the maximum duration of the trial (i.e. regardless of withdrawal or death).

B 11. Table 46 of the CS includes 279 patients in the difelikefalin model arm for Cycle 4 onwards reflecting the "population included in previous cycles that achieved a clinically meaningful treatment response at Week 12 and that entered the OLE period of the KALM-1 and KALM-2 trials, plus the population that received placebo in the DB period who were eligible to enter the OLE period and achieved a clinically meaningful treatment response at Week 12 of the OLE period". Please explain the difference between this group of patients and the 74 reported in the "observed" column of the DFK arm in Table 46. Please also clarify how many of the 279 patients that achieved a clinically meaningful response entered the OLE phase from the placebo arm and how many from the DFK arm.

For reference, Table 46 from the CS is provided below.

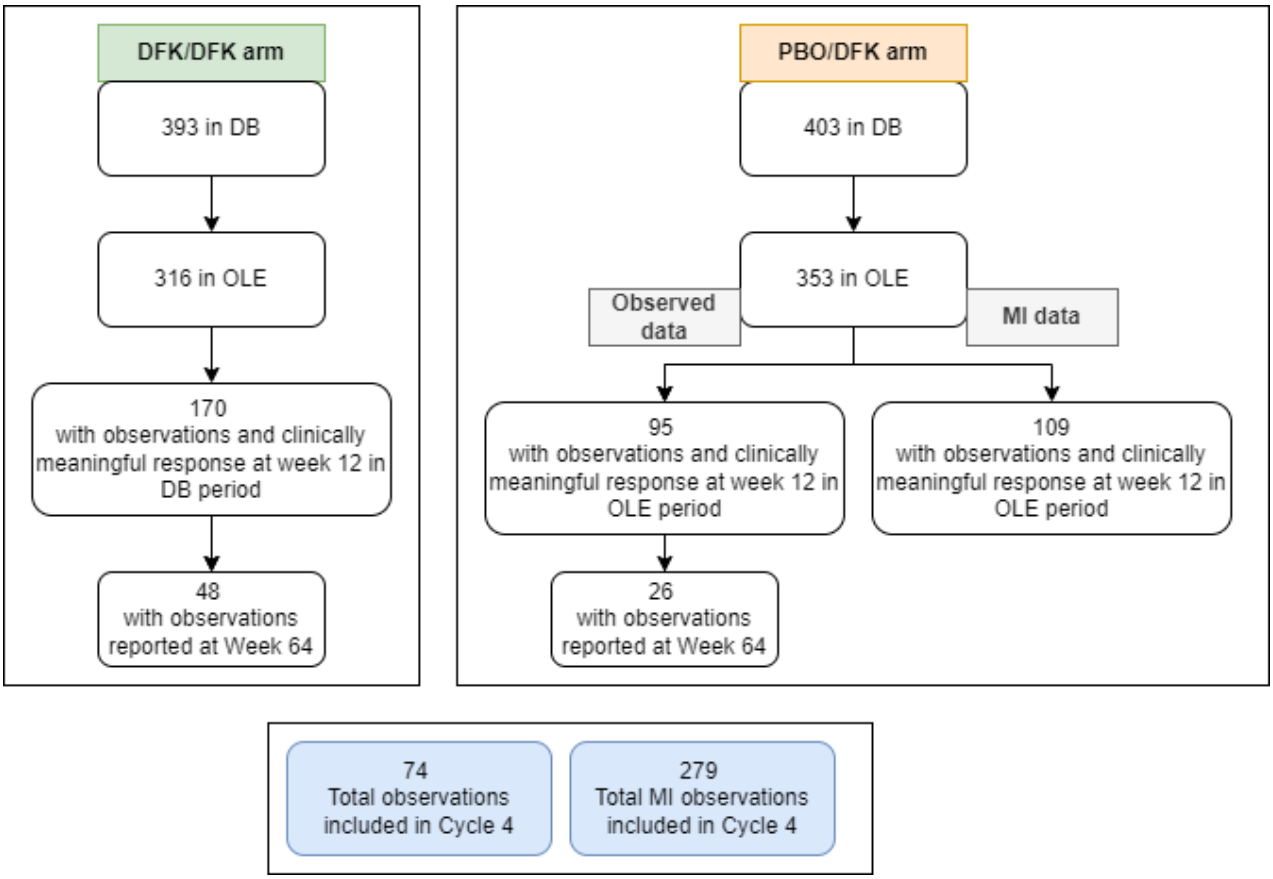
Table 46 : Number of data observations included in the analysis at each model cycle from KALM-1 and KALM-2

5-D Itch Scale total scores	**Difelikefalin**		**Placebo**	
Cycle	Observed only	Missing data imputation	Observed only	Missing data imputation
Baseline count	393	393	403	403
Cycle 1 (baseline to Week 4)	356	393	371	403
Cycle 2 (Week 4 to Week 8)	333	393	357	403
Cycle 3 (Week 8 to Week 12)	330	393	359	403
Cycle 4 (Week 12 to Week 64)	74	279	N/A	N/A

For the Cycle 4 observations, only patients who 1) entered the OLE period, and 2) achieved a clinically meaningful response are included. It is important to note that for the PBO/DFK arm of the KALM trials, clinical response was measured at week 12 of the OLE period, compared with week 12 of the DB period for the DFK/DFK arm.

Error! Reference source not found. outlines the number of observations split by treatment arm from the pooled KALM-1 and KALM-2 data.

Figure 7: Summary of observations included in Cycle 4 for observed and MI data set



B 12. Priority question: For the derivation of the transition matrices, it is assumed that the probability of improving or deteriorating CKD-aP in each cycle is equal regardless of current health state.

b) Please justify the assumption used to estimate transition probabilities

Creating a matrix which calculates the probability of moving from any one state to each of the other states can result in small observation numbers estimating a single probability value, which may lead to unrealistic outcomes. Furthermore, because an extrapolation of the trial data was required to estimate the long-term efficacy for patients receiving placebo (Cycle 4 onwards), unless assuming no change in efficacy, using estimates of a mean change in itch score from baseline would result in all placebo patients remaining with mild or moderate CKD-aP.

c) What are the limitations of this assumption?

When using the 'change in state' transitions by assuming that the probability of improving or deteriorating CKD-aP in each cycle is equal, it is implied that the rate of response to treatment is averaged across the population. By estimating treatment response by CKD-aP severity at baseline, the average treatment response is weighted by the distribution of patients at baseline (i.e. the number of patients with moderate, severe, or very severe CKD-aP). The numerical benefit of treatment with DFK was larger in more severe patients, therefore, the rate of transitions from more severe states to less severe states may be underestimated whilst the rate of transitions from less severe states to more severe states may be overestimated.

d) Please validate this assumption by estimating the transition probabilities directly from the patient-level data, so that they can be compared to those currently used.

The transition probabilities, directly populated by observed transitions are presented in the Tables below. Please note, no data are available beyond cycle 3 (week 12) for the PBO arm. As noted above, unless assuming no change in efficacy, using estimates of a mean change in itch score from baseline would result in all PBO patients remaining as mild or moderate CKD-aP.

Count data and the relevant probabilities are provided below. The states are defined as none = 1, mild = 2, moderate = 3, severe = 4, and very severe = 5.

DFK ARM

Cycle 1 - count		After					
		1	2	3	4	5	
Before	1	0	0	0	0	0	**0**
	2	0	0	0	0	0	**0**
	3	29	64	126	5	0	**224**
	4	7	21	79	19	3	**129**
	5	1	4	15	15	5	**40**
		37	**89**	**220**	**39**	**8**	**393**

Cycle 1 - probability		After					
		1	2	3	4	5	
Before	1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	3	12.95%	28.57%	56.25%	2.23%	0.00%	100.00%
	4	5.43%	16.28%	61.24%	14.73%	2.33%	100.00%

	5	2.50%	10.00%	37.50%	37.50%	12.50%	100.00%

Cycle 2 - count		After					
		1	2	3	4	5	
Before	1	17	17	3	0	0	**37**
	2	30	35	22	0	2	**89**
	3	10	53	142	14	1	**220**
	4	1	1	12	21	4	**39**
	5	0	1	0	4	3	**8**
		58	**107**	**179**	**39**	**10**	**393**

Cycle 2 - probability		After					
		1	2	3	4	5	
Before	1	45.95%	45.95%	8.11%	0.00%	0.00%	100.00%
	2	33.71%	39.33%	24.72%	0.00%	2.25%	100.00%
	3	4.55%	24.09%	64.55%	6.36%	0.45%	100.00%
	4	2.56%	2.56%	30.77%	53.85%	10.26%	100.00%
	5	0.00%	12.50%	0.00%	50.00%	37.50%	100.00%

Cycle 3 - count		After					
		1	2	3	4	5	
Before	1	37	21	0	0	0	**58**
	2	22	58	27	0	0	**107**
	3	13	44	113	7	2	**179**
	4	2	2	21	10	4	**39**
	5	1	1	0	3	5	**10**
		75	**126**	**161**	**20**	**11**	**393**

Cycle 3 - probability		After					
		1	2	3	4	5	
Before	1	63.79%	36.21%	0.00%	0.00%	0.00%	100.00%
	2	20.56%	54.21%	25.23%	0.00%	0.00%	100.00%
	3	7.26%	24.58%	63.13%	3.91%	1.12%	100.00%
	4	5.13%	5.13%	53.85%	25.64%	10.26%	100.00%
	5	10.00%	10.00%	0.00%	30.00%	50.00%	100.00%

Cycle 4 - count		After					
		1	2	3	4	5	
Before	1	43	10	3	1	0	**57**
	2	49	26	3	1	0	**79**
	3	41	37	17	1	0	**96**
	4	9	14	10	6	0	**39**
	5	1	1	3	3	0	**8**
		143	**88**	**36**	**12**	**0**	**279**

Cycle 4 - probability		After					
		1	2	3	4	5	
Before	1	75.44%	17.54%	5.26%	1.75%	0.00%	100.00%

	2	62.03%	32.91%	3.80%	1.27%	0.00%	100.00%
	3	42.71%	38.54%	17.71%	1.04%	0.00%	100.00%
	4	23.08%	35.90%	25.64%	15.38%	0.00%	100.00%
	5	12.50%	12.50%	37.50%	37.50%	0.00%	100.00%

PBO ARM

Cycle 1 - count		After					
		1	2	3	4	5	
Before	1	0	0	0	0	0	**0**
	2	0	0	0	0	0	**0**
	3	15	45	131	19	6	**216**
	4	2	19	78	37	7	**143**
	5	0	2	19	13	10	**44**
		17	**66**	**228**	**69**	**23**	**403**

Cycle 1 - probability		After					
		1	2	3	4	5	
Before	1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	3	6.94%	20.83%	60.65%	8.80%	2.78%	100.00%
	4	1.40%	13.29%	54.55%	25.87%	4.90%	100.00%
	5	0.00%	4.55%	43.18%	29.55%	22.73%	100.00%

Cycle 2 - count		After					
		1	2	3	4	5	
Before	1	9	5	3	0	0	**17**
	2	14	31	17	3	1	**66**
	3	6	48	153	19	2	**228**
	4	0	3	38	22	6	**69**
	5	0	2	6	4	11	**23**
		29	**89**	**217**	**48**	**20**	**403**

Cycle 2 - probability		After					
		1	2	3	4	5	
Before	1	52.94%	29.41%	17.65%	0.00%	0.00%	100.00%
	2	21.21%	46.97%	25.76%	4.55%	1.52%	100.00%
	3	2.63%	21.05%	67.11%	8.33%	0.88%	100.00%
	4	0.00%	4.35%	55.07%	31.88%	8.70%	100.00%
	5	0.00%	8.70%	26.09%	17.39%	47.83%	100.00%

Cycle 3 - count		After					
		1	2	3	4	5	
Before	1	19	8	2	0	0	**29**
	2	13	48	27	1	0	**89**
	3	4	52	131	29	1	**217**
	4	2	2	19	20	5	**48**

	5	1	1	0	11	7	**20**
		39	**111**	**179**	**61**	**13**	**403**

Cycle 3 - probability		After					
		1	2	3	4	5	
Before	1	65.52%	27.59%	6.90%	0.00%	0.00%	100.00%
	2	14.61%	53.93%	30.34%	1.12%	0.00%	100.00%
	3	1.84%	23.96%	60.37%	13.36%	0.46%	100.00%
	4	4.17%	4.17%	39.58%	41.67%	10.42%	100.00%
	5	5.00%	5.00%	0.00%	55.00%	35.00%	100.00%

e) Please include these directly derived transition matrices in the model in a scenario analysis.

As the current structure of the model has not been designed to conduct these analyses, a separate version of the model has been saved in which the above transition matrices are hard coded into the model. This model will provide the results to the requested scenario; however, efficacy values can no longer be included in the DSA and PSA in this adapted model version.

The results of this scenario are presented in Table 6:

Table 6: 'Change in state' transitions vs Directly observed transitions

Scenario	Inc. costs	Inc. LYs	Inc. QALYs	ICER
Revised base case 'Change in state' transitions	[redacted]	0.06	[redacted]	£23,277
Scenario Directly observed transitions	[redacted]	0.05	[redacted]	£25,792

B 13. Please provide a variant of Figure 16, based on MD instead of RoM.

See Figure 7 below for the variant of Figure 16 from the CS. Please also see Figure 8 for a direct comparison of placebo MD extrapolation and placebo ROM extrapolation

Figure 7: Mean change in 5-D Itch scale total score from baseline by baseline itch severity for difelikefalin and placebo (MD extrapolation)

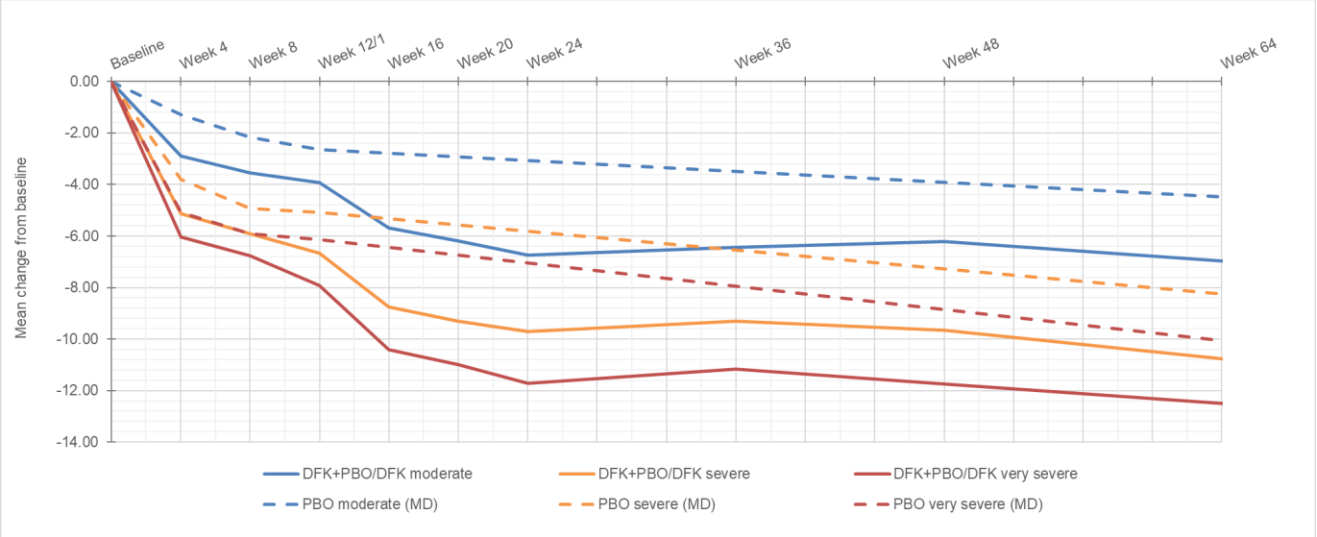
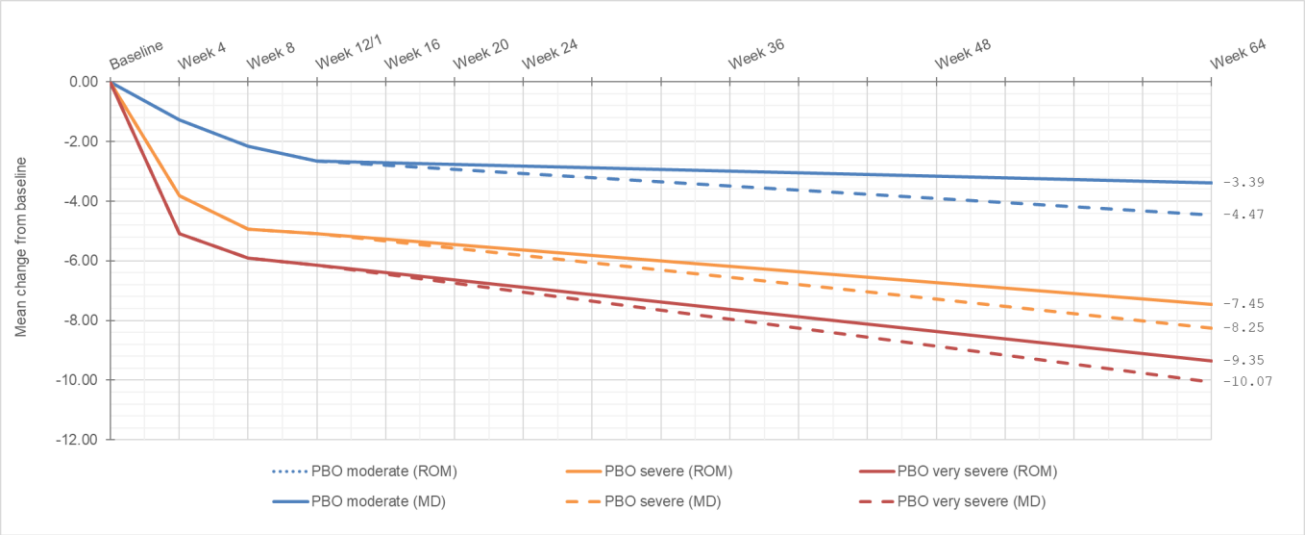


Figure 8: Mean change in 5-D Itch scale total score from baseline by baseline itch severity for placebo (ROM extrapolation) and placebo (MD extrapolation)



B 14. Priority question: In the base case analysis, transition probability matrices were reported to be estimated from a "simulated data set using the mean change from baseline in itch scores by CKD-aP severity at baseline: moderate, severe or very severe".

b) Please provide further details on the simulation methods used to derive these data, especially why this approach was taken and why it was

deemed the best approach as the explanation in the submission is unclear.

The simulation method uses the mean change from baseline in itch scores by CKD-aP severity at baseline: moderate, severe or very severe. These are provided in Table 47 from the CS and copied here for reference.

Table 47: Mean change in 5-D Itch scale total score from baseline

Mean change from baseline (5-D Itch scale total scores)	Week 4	Week 8	Week 12	Week 64
Difelikefalin treatment arm				
Moderate (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Very severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Established clinical management arm				
Moderate (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]
Very severe (SE)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SE; standard error
Note: This table corresponds with the curves presented in **Error! Reference source not found.**

The mean change values are multiplied by the baseline itch scores to estimate the mean itch score for the correlating week. For example, for a patient who receives difelikefalin with a baseline itch score of 15 (moderate CKD-aP), at week 8 their itch score would be 11.46 which is equal to 15 (baseline score) + -3.54 (mean change in itch score). The simulated scores are then grouped into health states and used to estimate the change in health state from the previous cycle. Continuing with the previous example, at week 4 the patient was defined as moderate (score of 12.10), and at week 8, their new health state is mild (score of 11.46). The change in health state values is then multiplied by the distribution of patients at baseline to estimate the weighted proportion of health state transitions for each cycle.

Using the mean change in itch score values to estimate the transition matrices in the model was the preferred option, as this best aligns with the trial outcomes which looked to measure an improvement in itch by the change from baseline at the end of week 12. This approach was also deemed more appropriate for probabilistic sensitivity analysis whereby the standard errors could be used to indicate variation from the mean. When using the directly observed data, a Dirichlet distribution is required to generate a fully probabilistic transition matrix which may be associated with increased uncertainty.

c) The model includes the option to use "*Observed*" instead of "*Simulated*" data on the Settings sheet. Please provide further details on the transition matrices as estimated from the observed data as this is currently not included in the CS.

The methods used to model efficacy and estimate transition matrices can be summarised as follows:

<u>Direct observed transitions</u>

- This is the scenario presented in question B.12. with the results presented in B.12. d).
- Here, matrices are directly populated by the observed transitions (e.g. in cycle 1, the proportion of patients who moved from the moderate health state to the mild health state)

<u>'Change in state' transitions</u>

- This is how the model structure is currently built.
- Here, matrices are derived from the per-cycle probability of losing or gaining between 0 and 3 health states. The per-cycle probabilities of losing or gaining health states can be estimated in 2 ways:
    - Using simulated data, or
    - Using observed data

The rationale for using 'change in state' transitions over direct observed transitions in the model base case is provided in the answer to question B.12.

When selecting “Observed” data on the settings sheet, the transition matrices used in the model are informed by values on the TM-Observed sheet. The bank of count data has been derived from the same data that is used to estimate the mean change in 5-D itch scores from baseline that are presented on the TM-Simulated sheet and used in the model base case. As with the simulation approach, the count data are used to estimate the probability of improving or deteriorating CKD-aP each cycle. However, as noted in response to B.14.a), when using the directly observed data, a Dirichlet distribution is required to generate a fully probabilistic transition matrix.

B 15. Priority Question: The transition probability matrices in Table 48 and Table 49 show that patients can only move to inferior health states. For example, a patient in the moderate health state can only go to the ‘mild’ or ‘none’ health state and cannot move to the ‘severe’ health state. The same applies for the ECM patients. Please explain the reasoning and the validity of this assumption.

When using the simulated data to inform transitions, the mean change value is used which will always result in patients either remaining in the same itch state, or improving.

Although a simplification of the data observed in the trial, this is modelling assumption is likely to reflect clinical practice whereby only patients who have a clinically significant improvement in itch will remain on treatment with difelikefalin. Furthermore, as noted in the answer to question B.9, on average, no treatment waning was observed in the KALM trials.

This assumption may overestimate the number of patients in more severe itch health states for the first 3 cycles of the model but will not have an impact on the proportion of patients who continue on treatment as this is still modelled directly from the observed data. Please see the answer to question B.34. which presents the comparison of the directly observed transitions (scenario for B.12) with the simulated transitions as used in the base case.

Utilities/HRQoL

B 16. In section B.3.4.2. "The data collected was used to estimate EQ-5D-3L mapping functions from 5-D Itch scale scores, WI-NRS, and 5-D Itch scale scores and WI-NRS combined" but the company used only the 5D-itch scale scores. Please provide a scenario analysis using the mapping function based on the combined data.

It is not possible to predict on the full KALM 1 and 2 datasets using the model with 5D-itch and WI-NRS as covariates because the KALM datasets did not collect WI-NRS data in the open label extension. Furthermore, this model was estimated using the crosswalk algorithm published by van Hout et al. - NICE's recommended mapping function to convert EQ-5D-5L to EQ-5D-3L at that time. Thus, there are two differences between Table 7 below and the utilities used in the model base case (Table 8 in the mapping study report (Appendix J)):

1. the table below is based on a smaller number of observations as it has been computed only for the Double-blind Phase of the KALM datasets
2. It is based on EQ-5D-3L values obtained from EQ-5D-5L using the van Hout cross walk

Table 7: EQ-5D-3L model predictions in the KALM trials using the 5-D itch and WI-NRS combined 3 component mapping model based on van Hout et al. crosswalk - Double-blind Phase

	Full sample		Severe/unbearable at baseline subsample	
	Mean EQ-5D-3L	95% confidence interval	Mean EQ-5D-3L	95% confidence interval
Not present	0.6185	0.5575, 0.6794	0.5942	0.5161, 0.6723
Mild	0.5815	0.5323, 0.6306	0.5768	0.5263, 0.6273
Moderate	0.5101	0.4634, 0.5567	0.5015	0.4535, 0.5496
Severe/unbearable	0.4104	0.3437, 0.4772	0.4076	0.3398, 0.4754
Sample size	n=3,386		n=1,628	

Scenario analyses are presented in Table 8 for basecase and 'Subgroup Analysis C. Severe only at baseline' using the mean EQ-5D-3L values presented in Table 7.

Table 8: Results using alternative mapping models to estimate utility values

ICER (£/QALY) results	Original EQ-5D-3L values	Table 7 EQ-5D-3L values
Base case analysis	£23,277	£21,915
Subgroup analysis C: Severe only at baseline	£18,642	£18,798

B 17. Please explain why in the mapping study the overall 5-D score was used as independent variable in the mapping function, rather than the scores on each of the 5 dimensions separately.

The mapping function is based on a sample of 377 observations. The three-component model requires the estimation of 23 parameters; including each dimension of the 5D-itch separately would increase the number of parameters for the same model to 43. Such a large number of parameters relative to the sample size makes model convergence difficult. Attempts at estimating models with a higher number of parameters resulted in many insignificant parameters at standard significance levels and problems of convergence.

B 18. In section B.3.4.5 it is stated that health state utility values for transplant were informed by Lee et al., 2005, which was identified in NICE TA775, while also other NICE HTAs that were identified in the expanded SLR were reviewed. Please provide a table with all identified utility values and the alternative sources and give an explanation why the Lee et al. 2005 was the preferred source for the base case analysis.

The Vifor Utility Mapping study was used for utility scores from none to very severe in the model. Lee et al., 2005 as identified from NICE TA775 was used as a source for transplant utility (0.712) as none of the other 5 HTAs listed included transplant utility, other than HST17 for liver failure which was excluded on bases of relevance to CKD.

B 19. In section B 3.4.4. it is explained that no HRQoL values for adverse events in patients with CKD-aP were identified in the SLR. However, it

appears that no search was done for AE-induced disutilities in other disease areas. Please justify why no such search was done.

Adverse events included diarrhoea, dizziness, nausea, gait disturbance(falls), hyperkalmia, headache and somnolence. Inputs for which were sourced from the KALM 1 and KALM 2 pooled data (Fishbane et al 2022). Given the type and frequency of AEs it was not feasible to perform literature searches for utilities for each AE due to the size of the published literature base describing them. Additionally, as incidence rates of AEs for patients treated with DFK and PBO are consistent in the pooled analysis of KALM 1 and 2, any incremental effect due to the utility impact of AEs is likely to be negligible.

B 20. Priority question: In section B 3.4.4 it is explained that the model does not account for a utility decrement due to AEs, since any utility decrements associated with AE are expected to be implicitly captured in health state utility values. However, the heath state utilities are based on responses to the 5D itching score, which only contains questions explicitly about itching. Thus, these scores are unlikely to capture the effect of the various AE reported. Also, most AEs usually occur in the early stage of treatment, whereas the 5D questionnaire was administered for the first time (after baseline) after 4 weeks, thus missing the period of the AE.

Furthermore, it is stated that the incremental incidence of adverse events reported in Fishbane et al., (2022) for the results of the pooled KALM-1 and KALM-2 trials were small and in general lower in those patients treated with DFK, suggesting that observed AEs are likely to be a feature of underlying disease.

a) **Please provide further evidence on the timing of the occurrence of AEs.**

b) **Please justify comparing the incidence rate of AE of the control group to the combined observations in the 12-week controlled study and the extension study where all patients received DFK, rather than the incidence rate of AE of DFK in the 12-week controlled study.**

c) Please provide disutility estimates for all AEs currently included in the model.

Most patients treated with DFK who experienced a treatment emergent AE did so within the first 12 weeks on treatment. The incidence rate of experiencing ≥1 AE per 1,000 patient years was 10,863 in the first 12 weeks of treatment with DFK, reducing to 8,116 per 1,000 patient years over the entire 64-week follow-up period, based on the pooled analysis of KALM 1 and 2 presented by Fishbane et al (22). This reduction implies that for the period between 12 and 64 weeks, the incidence rate was lower than in the first 12 weeks of the trials.

The model utilises all available data describing the incidence of AEs for patients treated with DFK from the pooled analysis of KALM 1 and 2. This is consistent with its application in the model, where patients treated with DFK are at the same risk of experiencing all modelled adverse events regardless of their time on treatment. As no data beyond 12 weeks was available for patients treated with PBO, it was assumed that the long-term incidence of AEs for patients treated with ECM would be consistent with the first 12 weeks. This assumption is believed to be appropriate, as the AE profile associated with ECM in the model is based on patients in receipt of PBO, and as such AEs experienced by these patients are likely a feature of underlying chronic comorbidity which may not change over time.

As the incidence rate and type of AEs for patients treated with DFK and PBO are consistent, it is believed that the AEs observed in the trial are largely a feature of underlying comorbidity in the enrolled patient population. As such, the health state utility estimates from the mapping study are likely to capture the impact of AEs on patient quality of life implicitly. Although the health state utility values are based on a predictive model that used 5D Itch as an independent variable, which as noted, may not be sensitive to the quality-of-life impact of AEs, the estimated health state utility values are derived from EQ-5D-5L questionnaires (mapped to EQ-5D-3L using the crosswalk algorithm published by van Hout et al.) which is sensitive to these events. As such, the utility impact of AEs will be captured in the model intercept, resulting in a corresponding reduction in the health state utilities that are incorporated in the model.

To assess the robustness of model results to these assumptions, an additional scenario has been included where patients treated with DFK and ECM will experience AEs for the first 12 weeks of the model only, with AE incidence based on the 12-week data from the pooled analysis of KALM 1 and 2. This scenario also included an additional event specific utility decrement as presented in Table 9. The results of this scenario are presented in Table 10.

Table 9: Utility decrements for AEs in the model

Adverse event	Utility decrement (SE)	Source (30-33)
Diarrhoea	0.0753 (0.0209)	Sullivan et al.; non-infectious gastroenteritis
Dizziness	0.1500 (0.0581)	Agrawal et al.; vestibular loss
Nausea	0.0753 (0.0209)	Sullivan et al.; non-infectious gastroenteritis
Gait disturbance (falls)	0.1500 (0.0581)	Agrawal et al.; vestibular loss
Hyperkalaemia	0.0300 (0.0030*)	Palaka et al. 2019
Headache	0.0439 (0.0090)	Sullivan et al.; Migraine
Somnolence	0.1130 (0.0113*)	Katz et al.
Abbreviations: CKD, chronic kidney disease; SE, standard error. * SE assumed 10% of mean		

Table 10: Adverse event scenario

	Description	**Inc. Costs**	**Inc. LYs**	**Inc. QALYs**	**ICER**
Base case analysis	No additional AE disutility modelled. AE rates based on DFK-all exposure,	[redacted]	0.06	[redacted]	£23,277

Scenario for AEs	AE disutility applied as per Table 9. AE rates based on DFK 12-week analysis and applied only in first 3 cycles of model (upto 12 weeks).	[redacted]	0.06	[redacted]	£25,807

a. Resource Use/Costs

B 21. Please update all NHS cost process from 2019/2020 to 2020/2021

Updated in model.

B 22. Please clarify how you derived the resource use for the two merged groups (severe and very severe). Is it an average of the two merged groups, post merging?

Resource use for the very severe state were set equal to the resource use values for the severe state as these groups were merged in the mapping study.

B 23. The resource use for severe and very severe patient populations were merged in the mapping study. This is justified by the small size of each population group. Please include a description on how the estimation of the resource use was derived for this merged group. Were any assumptions made for the derivation?

[Company: please enter your answer to this question here]

B 24. Please provide the resource use for the severe and very severe patient groups separately.

[Company: please enter your answer to this question here]

B 25. Please clarify the sentence “As patients in the KALM trials were not discontinued in this way, it is not possible to exclusively measure efficacy and model outcomes for this patient group.” Does that mean, if in the run-in to the trial the ‘no clinically meaningful response’ patients would not continue the DFK and be on ECM, we therefore have no evidence on

the patients who experienced no effect of DFK in the run-in period? Did they then join the placebo arm, or were they dropped?

A stopping rule has been implemented in the analysis whereby patients on difelikefalin not achieving clinically significant improvement in itch at 12 weeks will discontinue treatment. The model assumes that these patients will progress at the same rate as the ECM arm for the remaining duration of the model.

This assumption had to be made as the KALM trials did not discontinue patients in this way. Level of treatment response was not specified in the inclusion criteria for the OLE phase of the KALM trials. Therefore, there is no data on patients who would discontinue treatment with difelikefalin following an insufficient response to treatment with difelikefalin.

B 26. The estimations for proportion receiving anti-itch medication by health states (Table 55) are stated to be detailed in section B.3.4.2, and then in appendix J. However, both this section and the appendix only describe data collection on HRQoL. Please provide information on how the data as presented in Table 55 was collected. Please also provide patient numbers besides the percentages in Table 55.

The case report form for the study asked for the relevant medications which have been used in management of CKD-aP including dose and frequency of medications.

The total number of patients per state were as follows:

- None, n = 164
- Mild, n = 117
- Moderate, n = 123
- Severe/Very severe, n = 81

Note, numbers of patients and proportions do not correctly sum up in the original Mapping study report.

B 27. The weighted total treatment costs for moderate CKD-aP were set to be equal to the weighted total treatment costs for mild CKD-aP.

a) **Please provide a justification for choosing mild health state costs instead of using an average between mild and severe health state cost or applying the costs of severe also to moderate.**

In the base case, the moderate health state treatment costs were set equal to the mild health state treatment costs as this was seen to be conservative for the cost-effectiveness for difelikefalin. However, the impact on the ICER is negligible. Please see results for these scenarios in Table 11:

Table 11: Moderate health state treatment cost scenarios

Scenario	Inc. Costs	ICER (£/QALY)
Updated base case ICER Treatment costs; Moderate health state set to Mild health state	[redacted]	[redacted]
Treatment costs; Moderate health state set to Severe health state	[redacted]	[redacted]
Treatment costs; Moderate health state set to average between Mild and Severe health state	[redacted]	[redacted]

b) Please comment on the assumption that the moderate health state cost from the mapping study is unrealistic and whether this is backed up by clinical opinion. Does this reflect on the certainty of the other health state cost estimates?

As highlighted in the submission, it was noted that the total weighted treatment cost for moderate CKD-aP was lower than that for 'mild' and 'none' severity due to the greatly reduced proportion of people with moderate CKD-aP using antidepressants.

This assumption was not checked with clinical opinion. As shown in Table 7, the company do not believe that this assumption has a material weighting on the ICER and cost-effectiveness of difelikefalin.

c) Is it possible that patients are in the moderate health state rather than the mild health state because they receive insufficient treatment for their itching, hence the lower costs (i.e., the moderate health state is caused by the lower resource use and thus costs)? Please clarify whether this is the case.

The proportion of patients receiving anti-itch medication with severe CKD-aP is greater than both those without, and with mild CKD-aP. This would suggest that

treatment for CKD-aP in these patients is itself insufficient, rather than patients receiving a lack of treatment.

B 28. Regarding average annual cost for ECM:

Please note that an error was identified in the model regarding the calculation of the weekly treatment costs. This has been corrected in the updated model.

a) **Please include units for table 55. For example, in the dose column 1.5 units of what? For oral corticosteroids, is 1.00 =10 mg as mentioned in the source column? Also, is a unit in the unit costs a pack, or a single dose?**

This has been updated in the model.

Dose refers to the dose required per day. For topical corticosteroids the dose per mg of active ingredient is unspecific. For topical corticosteroids, following BNF dosage guidance for hydrocortisone (10mg per 1 gram) for mild inflammatory skin disorders to be applied 1-2 times a day, the dose in the model assumes application 1.5 times per day at 10mg per gram. Upon review, the NHS indicative pack price for hydrocortisone 1% cream was updated to £1.26.

b) **Please also clarify the dosage mentioned in Table 55, are these dosages per day? What source of information was used to determine these dosages?**

As stated above, this has been updated in the model and refers to the dose per day. The BNF was used to determine the relevant dosages and pack prices.

c) **It appears that the assumption is made that the dosages of the various drugs will be the same for all health states; please provide a justification for this assumption.**

This assumption was made given paucity of data on current management of CKD-aP in current clinical practice.

Given the cheap costs of anti-itch medications, the company do not believe that any increase in dosages across health states would have a material impact on the cost-effectiveness of difelikefalin.

B 29. Does the pooled trial physical weight reflect the UK CKD-aP population physical weight? In Table 41 of the CS, UKRR median age, male proportion, and length of time on dialysis is reported as the median, while weight is not reported. Are there any figures in the literature? Does the UKRR data not report for means? It would be preferrable to see the comparison of age, sex, weight, and length of time of dialysis between the trial and UKRR populations for mean (SD).

No additional information on the demographics of patients with CKD-aP in the UK were identified in the literature.

All data on the demographics of the prevalent adult in-centre haemodialysis population as reported by the UKRR is presented as the median.

B 30. In B 2.5.2 it is mentioned that patients will visit a nephrologist once per 3 months.

a) Please clarify if this schedule is the same for all health states.

In the modified Delphi panel, it was noted that patients would have a 3-monthly patient review conducted by the consultant nephrologists. This consultation was highlighted as being part of the clinical pathway for people receiving haemodialysis and would therefore not be an exclusive review of any associated pruritus.

It is not expected that additional reviews would be conducted if the associated pruritus was more severe.

b) Please clarify the correct frequency, as the submission states once per 3 months, but in the electronic model a frequency of once per 3 weeks is used.

This was an error in the model and has been updated as per submission, with nephrologist visits occuring once every 3 months.

c) Are the nephrologist visits dependent on severity of condition of patient? The patient review visit is based on first attendance consultation (WF01B). Please add clarification on whether the first attendance consultation is appropriate given that these patients see the nephrologist on a regular basis.

Adjusted in model to follow-up consultation (WF01A)

B 31. The 2019-2021 inflation rate is used for the post-transplant cost, which was stated to be sourced for the cost year 2009. Please use the appropriate inflation rate 2009-2021 for post-transplant cost.

This was a reporting error in the CS, and has been inflated using the correct rate.

B 32. The adverse events cost is based on a single GP visit. This is justified by the statement "*Given that no relevant or appropriate costs for AEs were identified in either the SLR or the adapted SLR*". Please describe and substantiate the assumption behind the AE cost, is this based on clinical opinion? Are there no tariffs or appraisals for CKD-aP or CKD which could inform the AE cost?

As noted in the answer for B.19, given the wide range and generalisability of the AEs it was not feasible to perform literature searches for either costs or utilities for each AE due to the size of potential published evidence base.

None of the appraisals identified in the expanded literature search included costs which could be used to inform the costs for the AEs included in the model.

Furthermore, as noted in the CS, for the all-difelikefalin-exposure and placebo cohort of the difelikefalin trials, AEs were mild or moderate in severity (≥65% of any of the events) in the majority of patients (22). As such, it was assumed that the costs of AEs could be considered manageable through a standard GP appointment, if not absorbed in the HRG code for patients receiving haemodialysis .

B 33. The % male patients as reported in Table 41 differs from the percentage used in the electronic model. Please clarify which input is correct. If the 59.58% in the model is the correct input, please add the source for this value and add justification for choosing this input.

The model is incorrect, and the value should be 58.7% as noted in Fishbane et al., 2022. This has been updated.

b. Validation

B 34. Priority question: Please provide an internal validation to show to what extent the model results match the observed data for the first 64 weeks.

Figure 9 and Figure 10 present the distribution of patients in the DFK arm of the model at Cycle 3 (Week 12) and Cycle 4 (Week 64) when using the 'change in state' transition estimates (base case) and the directly observed transitions (scenario for question B.12) as estimated from the pooled patient level data for the KALM trials.

Figure 9: Cycle 3 - health state distributions for alternative model transition estimates

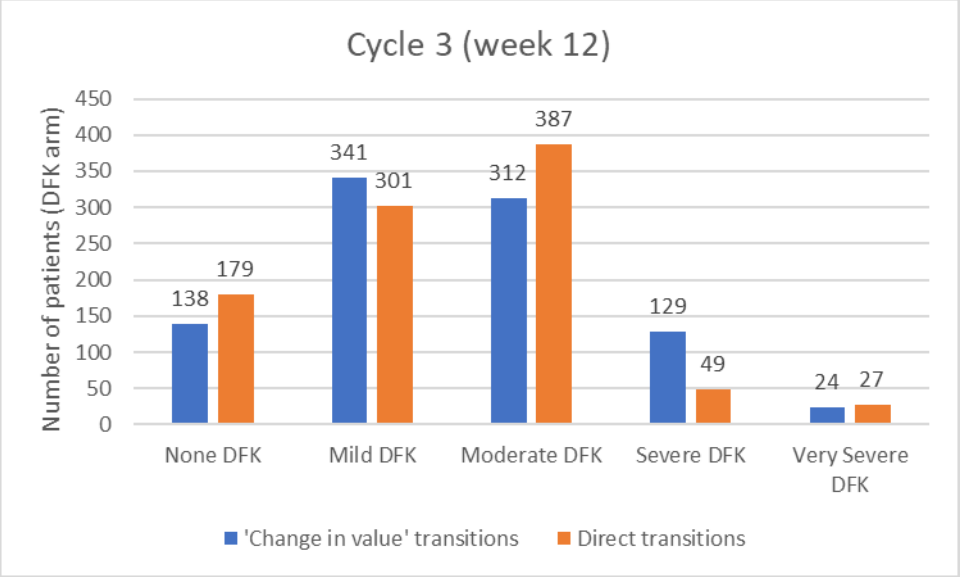
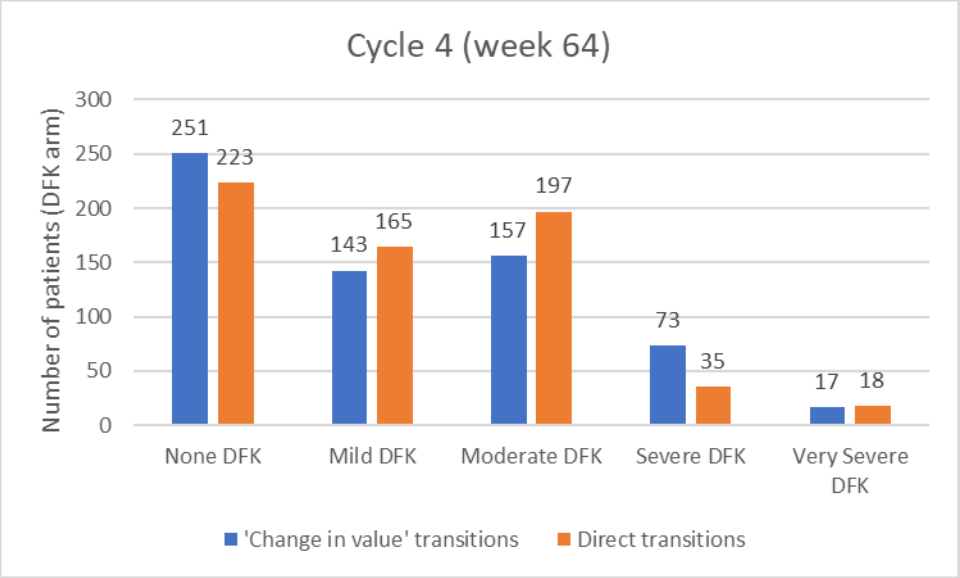


Figure 10: Cycle 4 - health state distributions for alternative model transition estimates



Textual clarification and additional points

C 1. The company submission (Document B) has used a referencing system based on authors' surnames and year of publication (e.g., Harvard). Please provide a reformatted version of Document B, using a numbered referencing system (e.g., Vancouver).

See revised Document B with Vancouver referencing.

C 2. LH odds ratio test in Table 24 of the CS is reported as a %. Please correct.

See corrected Table 15 below:

Table 12*:* Subjects with a ≥4-point improvement from baseline at Week 12 in Worst Itching Intensity NRS Score – MI with MAR assumption (population: ITT)

Combined assessments (Week 12)	Placebo (n=189)	DFK (n=189)
Observed ≥4-point NRS improvement [1] – n (%)		
Yes	35 (21.2%)	64 (40.8%)
No	130 (78.8%)	93 (59.2%)
Missing	24	32
LS means estimate of percent with improvement [2]		
Percent (95% CI)	18.0% (21.1%, 26.0%)	38.9% (29.8%, 48.7%)
LH odds ratio (95% CI)		2.89 (1.75, 4.76)

CHW p-value		<.001

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = numerical rating scale. [1] Counts and percentages were based on non-missing data. [2] Estimated percent, odds ratio, and p-value use a logistic regression model with terms for treatment group, baseline Worst Itching Intensity NRS score, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.

C 3. The total management costs in Table 62 of the base case results do not match with the total management costs in the model results sheet. Please explain the discrepancy.

The company does not see any difference between the values in Table 62 of the CS and the total management costs in the model results sheet.

This has been updated with the results of the revised model.

C 4. In Table 65 of the scenario analysis, it is reported that the 'no change in efficacy applied in long-term extrapolation for ECM arm' results in an ICER of £26,443/QALY gain. However, this ICER value corresponds to scenario 1b of the model on the ratio of means (RoM). Please confirm if that is the case and provide the value for the scenario 'no change in efficacy applied in long-term extrapolation for ECM arm'.

This was a reporting error and scenario 1.b. in Table 65 of the CS should state:

"Ratio of means long-term extrapolation for ECM arm".

C 5. Please provide more details on the subgroup analyses in Section B3.11 and explain how the user can run these analyses in the model.

The relevant subgroup analyses have been included in the scenario selection dropdown on the settings sheet in the model. Detail on the changes made in the model when selecting the relevant subgroup analysis is provided in section B.3.11 of the CS.

C 6. Priority question: Please adjust the model such that the hazard ratios for hospitalisation ('Management costs' worksheet) are no longer hard-coded, but are named input parameters that are also included in the PSA and DSA.

Updated in revised model.

Please note, in the parameters sheet, the disease management health state costs are no longer included in the PSA and DSA – these have been unselected.

C 7. There are several error messages throughout Document B, indicating that hyperlinks between different parts of the document are no longer working (i.e., "*Error! Reference source not found*"). Please provide a version of Document B with all of these links restored or correctly deleted.

The submission document provided does not contain any error messages indicating broken hyperlink

List of model changes

This list can also be found in the revised model on sheet ‘Updates’.

Question reference	Description of error	Location of change	Detail of change
B.20 c)	N/A	Adverse event sheet	An additional scenario has been included where patients treated with DFK and ECM will experience AEs for the first 12 weeks of the model only, with AE incidence based on the 12-week data from the pooled analysis of KALM 1 and 2. This scenario also included an additional event specific utility decrement. Data informing this scenario have been included in the off-piste section of the sheet from Row 151 onwards
B.21	NHS National schedule of costs informing management costs was an old version.	Management costs sheet; Rows 160 to 201	Updated all NHS cost process from 2019/2020 to 2020/2021, and removed HRG currencies that were for patients aged 18 or under.
B.28	There was an error in the calculation of the established clinical management weekly treatment costs.	Treatment costs sheet; Cells S29:S36	The weekly cost was adjusted to correctly calculate the weekly cost of treatment as informed by the daily dose, pack size and pack cost.
B.30 b)	The frequency of nephrologist visits is once per 3-weeks in the model. This is incorrect and should be once per 3-months, as stated in the submission.	Management costs sheet; Cells H25:L25	Frequency of nephrologist visits changed to 4 per year (equal to once per 3-months).
B.30 c)	Nephrologist attendance costed as first visit (WF01B)	Management costs sheet; Cells I175:J175	Nephrologist visit adjusted to follow-up cost (WF01A).
C.5	No functionality to run sub-group analysis in the model	Settings sheet; Dropdown cell G31	Added in the functionality to run the 3 subgroups as detailed in the company submission. Additional switches were added to the treatment costs and settings sheet to update as appropriate. AdditionalAdditional data were input in the TM-Observed and TM-Simulated sheet.
C.6	Hazard ratios for all-cause hospitalisation rates hard-coded into parameter values	Management costs sheet; Cells H23:L24	Added in row below hospitalisation rates to reflect the relevant hazard ratios as provided by Sukul et al., 2021. These have been appropriately updated and applied in the parameters sheet to work in the DSA and PSA.
Additional change	External sheets linked to model. No impact on model.	TM-Simulated sheet; Cells E89:F114, O89:R100, and AB89:AE100	Copy and pasted cell values for E89:F114 to remove the link Copied formulas from row O101:R101, upwards to row O89:R89 Copied formulas from row AB101:AE101, upwards to row AB89:AE89

Additional change	DSA Tornado diagram labels incorrect	DSA sheet; Tornado diagram	Source of labels adjusted.
Additional change	Simulated 5-D itch scores not correctly bounded between 5 and 25. This was identified in the DSA plot when looking at the resampled efficacy for the DFK treatment arm. This error has no impact on the deterministic results.	TM-Simulated sheet; Cells K89:N114 and	Old formula; IFERROR(($D89+INDEX($D$32:$H$38,MATCH($E89,$D$32:$D$38,0),MATCH(N$88,$D$32:$H$32,0))),"") New formula; =IFERROR(**MAX**($D89+INDEX($D$32:$H$38,MATCH($E89,$D$32:$D$38,0),MATCH(N$88,$D$32:$H$32,0))**,5**),"")

References

1.Informedhealth.org. When is the best time to start dialysis? 2018 [Available from: https://www.informedhealth.org/when-is-the-best-time-to-start-dialysis.html#:~:text=Dialysis%20is%20usually%20started%20as%20soon%20as%20the,also%20referred%20to%20as%20end-stage%20renal%20%28kidney%29%20disease.
2.Foundation NK. Dialysis 2022 [Available from: https://www.kidney.org/atoz/content/dialysisinfo#:~:text=Dialysis%20is%20a%20treatment%20that%20does%20some%20of,kidney%20function%20and%20have%20a%20GFR%20of%20%3C15.
3.Topf J, Wooldridge T, McCafferty K, Schomig M, Csiky B, Zwiech R, et al. Efficacy of Difelikefalin for the Treatment of Moderate to Severe Pruritus in Hemodialysis Patients: Pooled Analysis of KALM-1 and KALM-2 Phase 3 Studies. Kidney Med. 2022;4(8):100512.
4.Dermatographia. What antihistamines do you use for itchy skin? 2022 [Available from: https://dermatographia.com/antihistamines-for-itchy-skin/.
5.Kragballe K. Topical corticosteroids: mechanisms of action. Acta Derm Venereol Suppl (Stockh). 1989;151:7-10; discussion 47-52.
6.Ogbru O. ammonium lactate (Amlactin, Geri-Hydrolac) Last review: 2020 [
7.Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics. 1999;55(3):853-7.
8.Vernon M, Stander S, Munera C, Spencer RH, Menzaghi F. Clinically meaningful change in itch intensity scores: An evaluation in patients with chronic kidney disease-associated pruritus. J Am Acad Dermatol. 2021;84(4):1132-4.
9.Vernon MK, Swett LL, Speck RM, Munera C, Spencer RH, Wen W, et al. Psychometric validation and meaningful change thresholds of the Worst Itching Intensity Numerical Rating Scale for assessing itch in patients with chronic kidney disease-associated pruritus. J Patient Rep Outcomes. 2021;5(1):134.
10.Mathur VS, Lindberg J, Germain M, Block G, Tumlin J, Smith M, et al. A longitudinal study of uremic pruritus in hemodialysis patients. Clin J Am Soc Nephrol. 2010;5(8):1410-9.
11.Storck M, Sandmann S, Bruland P, Pereira MP, Steinke S, Riepe C, et al. Pruritus Intensity Scales across Europe: a prospective validation study. J Eur Acad Dermatol Venereol. 2021;35(5):1176-85.
12.Reich A, Heisig M, Phan NQ, Taneda K, Takamori K, Takeuchi S, et al. Visual analogue scale: evaluation of the instrument for the assessment of pruritus. Acta Derm Venereol. 2012;92(5):497-501.
13.Fishbane S, Mathur V, Germain MJ, Shirazian S, Bhaduri S, Munera C, et al. Randomized Controlled Trial of Difelikefalin for Chronic Pruritus in Hemodialysis Patients. Kidney Int Rep. 2020;5(5):600-10.
14.Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. Br J Dermatol. 2010;162(3):587-93.
15.Manenti L, Leuci E. Do you feel itchy? A guide towards diagnosis and measurement of chronic kidney disease-associated pruritus in dialysis patients. Clin Kidney J. 2021;14(Suppl 3):i8-i15.

16.Rayner HC, Larkina M, Wang M, Graham-Brown M, van der Veer SN, Ecder T, et al. International Comparisons of Prevalence, Awareness, and Treatment of Pruritus in People on Hemodialysis. Clin J Am Soc Nephrol. 2017;12(12):2000-7.
17.Dapagliflozin. Dapagliflozin for treating chronic kidney disease [ID3866]. 2021.
18.Roxadustat. Roxadustat for treating anaemia in people with chronic kidney disease [ID1483]. 2021.
19.NICE. Baricitinib for treating moderate to severe atopic dermatitis [ID1622]. 2020.
20.Dupilumab. Dupilumab for treating moderate to severe atopic dermatitis after topical treatments [ID1048]. 2018.
21.NICE. Upadacitinib, abrocitinib and tralokinumab for dermatitis [ID3960]. 2021.
22.Fishbane S, Wen W, Munera C, Lin R, Bagal S, McCafferty K, et al. Safety and Tolerability of Difelikefalin for the Treatment of Moderate-to-Severe Pruritus in Hemodialysis Patients: Pooled Analysis From the Phase 3 Clinical Trial Program. Kidney Medicine. 2022.
23.Sukul N, Karaboyas A, Csomor PA, Schaufler T, Wen W, Menzaghi F, et al. Self-reported Pruritus and Clinical, Dialysis-Related, and Patient-Reported Outcomes in Hemodialysis Patients. Kidney medicine. 2021;3(1).
24.Elder SJ, Pisoni RL, Akizawa T, Fissell R, Andreucci VE, Fukuhara S, et al. Sleep quality predicts quality of life and mortality risk in haemodialysis patients: results from the Dialysis Outcomes and Practice Patterns Study (DOPPS). Nephrol Dial Transplant. 2008;23(3):998-1004.
25.Weisbord SD, Mor MK, Sevick MA, Shields AM, Rollman BL, Palevsky PM, et al. Associations of depressive symptoms and pain with dialysis adherence, health resource utilization, and mortality in patients receiving chronic hemodialysis. Clin J Am Soc Nephrol. 2014;9(9):1594-602.
26.Cara Therapeutics I. CR845-CLIN3102 Clinical Study Report – Open-label Extension Phase (AIC). 2020 10/12/2020.
27.Cara Therapeutics I. CR845-CLIN3103 Clinical Study Report – Open-label Extension Phase (AIC). 2020 12/11/2020.
28.Yanagida T. Package ‘misty’ - Miscellaneous Functions 2022 2022-06-08.
29.van Buuren S. Package ‘mice’ - Multivariate Imputation by Chained Equations. 2021 24/11/2021.
30.Agrawal Y, Pineault KG, Semenov YR. Health-related quality of life and economic burden of vestibular loss in older adults. Laryngoscope Investig Otolaryngol. 2017;3(1):8-15.
31.Katz JN, Smith SR, Collins JE, Solomon DH, Jordan JM, Hunter DJ, et al. Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities. Osteoarthritis Cartilage. 2015;24(3):409-18.
32.Palaka EJ, James; Moon, Rebecca; Grandy, Susan. HEALTH STATE UTILITY OF CKD PATIENTS WITH HYPERKALEMIA: ANALYSIS OF EQ-5D IN A REAL WORLD POPULATION ACROSS THE EU-5, CHINA AND USA. Nephrology Dialysis Transplantation. 2019;34.
33.Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. Med Decis Making. 2011;31(6):800-4.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

NICE National Institute for Health and Care Excellence

About you

1.Your name	[redacted]
2. Name of organisation	Kidney Research UK
3. Job title or position	Head of Policy and External Affairs
4a. Brief description of the organisation (including who funds it). How many members does it have?	Kidney Research UK is the leading kidney research charity in the UK. We fund and promote research into kidney disease and related topics; bring together patients and renal researchers in networks and clinical study groups; raise awareness of kidney disease; campaign for the adoption of best practice by the NHS; and campaign for improved health outcomes for kidney patients. Our latest annual report 2020/21 shows the majority of our income is from donations, gifts and legacies (78%). The remainder is from trusts, partnerships, investments, trading and government funding. We are not a membership organisation but have an extensive supporter base and a significant number of active volunteers.
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	Vifor Fresenius / Pharma - Clinical Research Projects £319,183.00

4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	Kidney Research UK commissioned the University of Hertfordshire to survey patients with kidney disease on their experiences of pruritus. One-to-one in-depth semi-structured interviews were conducted with nine individuals with a variety of treatment modalities, ethnicities, ages and genders.

Living with the condition

What is like to e with ndition? at do ers perience en ing for meone h the ndition?	Living with kidney disease makes every day a challenge. It is a life-threatening condition that never goes away. Kidney disease affects, and often governs, every aspect of people's lives, requiring extensive medical treatment time and time again. [redacted] [redacted] [redacted] Caring for a person with kidney disease can be exhausting, both physically and emotionally. Multiple trips to hospital every week if the patient is using in-unit dialys or managing complex medical equipment at home for home dialysis; extensive medicines regimes; and often managing mental health conditions such as anxiety a depression. Dealing with pruritus on top of this can be extremely challenging.

Current treatment of the condition in the NHS

[redacted]

[redacted]

ysis patients are very much or extremely troubled by itching, but up-to 18% receive no treatment for this symptom. In addition, 17% had not reported itching to a healthcare pro

Advantages of the technology

[redacted]

Disadvantages of the technology

hat do ts or think e vantages ology?	[redacted] An oral medicine would address this. This treatment would provide another treatment option where other interventions have failed to manage the condition

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Kidney disease disproportionally affects people from deprived communities and ethnic minority groups and people in these cohorts progress faster to end stage renal failure*. Evidence shows that fewer kidney patients from deprived communities are treated with peritoneal dialysis, with more treated with haemodialysis. There are therefore likely to be proportionally more people from these cohorts on haemodialysis, experiencing pruritus and likely to benefit from this treatment. * Kidney Health Inequalities in the UK: Reflecting on the past, reducing in the future. Kidney Research UK 2018

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	See above – people from ethnic minority groups and deprived communities may be more likely to benefit from this treatment as they are more likely to live with kidney disease and more likely to be treated with haemodialysis when they reach renal failure.

Other issues

. Are there y other sues that u would e the mmittee consider?	People on haemodialysis have to contend with an extremely burdensome treatment regime to stay alive. On top of this, pruritus can significantly impact their quality of life. [redacted]

. To be ded by chnical am at ope sign f. Note that pic-ecific estions ll be ded only if e eatment thway or ely use of e chnology mains certain er oping nsultation, r example here were fferences opinion; s is not pected to required r every praisal.] here are ne delete ghlighted	

ws and number low	

Key messages

[redacted]

People from deprived communities and ethnic minority groups are more likely to require renal replacement therapy and may be more likely to benefit from this treatment

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - **YES** or NO

For more information about how we process your personal data please see our privacy notice.

1.4

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]

Produced by

Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC+)

Authors

Susan O'Meara, Reviews Manager, KSR Ltd, United Kingdom (UK)
Venetia Qendri, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands (NL)
Mark Perry, Reviews Manager, KSR Ltd, UK
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Ingelin Kvamme, Health Economics Researcher, Institute for Medical Technology Assessment (IMTA), EUR, the NL
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Caro Noake, Information Specialist, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Maiwenn Al, Health Economics Researcher, ESHPM, EUR, the NL
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to Susan O'Meara
Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 19/10/2022

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 135355.

Declared competing interests of the authors None

Acknowledgements None

[redacted]

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

O'Meara S, Qendri V, Perry M, Danopoulos E, Kvamme I, McDermott K, Posadzki P, Noake C, Armstrong N, Al M, Wolff R. Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2022.

Contributions of authors

Susan O'Meara acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Venetia Qendri, Ingelin Kvamme and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Mark Perry, Evan Danopoulos, Kevin McDermott and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence, and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Maiwenn Al acted as health economics project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Robert Wolff critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

β	Regression coefficient
5-D Itch scale	Five-dimension Itch scale
AE	Adverse events
AGREE II	Appraisal of Guidelines for Research and Evaluation, second version
AiC	Academic in confidence
ANCOVA	Analysis of covariance
BAD	British Association of Dermatologists
BNF	British National Formulary
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Database of Systematic Reviews
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Clinical Trials
CMHP	Committee for Medicinal Products for Human Use
CHW	Cui, Hung, Wang
CI	Confidence interval
CiC	Commercial in confidence
CKD	Chronic kidney disease
CKD-aP	CKD-associated pruritis
CNS	Central nervous system
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
DFK	Difelikefalin
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
EAG	Evidence Assessment Group
ECM	Established clinical management
EMA	European Medicines Agency
EOT	End of treatment
EQ-5D	European Quality of Life-5 Dimensions
ED-5D-3L	European Quality of Life-5 Dimensions 3 Levels
EQ-5D-5L	European Quality of Life-5 Dimensions 5 Levels
ESKD	End stage kidney disease
ESHPM	Erasmus School of Health Policy & Management
ESRD	End stage renal disease
EUR	Erasmus University Rotterdam
FE	Fixing errors
FDA	Food and Drug Administration
FV	Fixing violations
GIN	Guidelines International Network
GP	General Practitioner
HD	Haemodialysis
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility value
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ICER	Institute for Clinical and Economic Review
ICHD	In-centre haemodialysis
iDBC	Disease Burden Calculator

iMTA	Institute for Medical Technology Assessment
Incr.	Incremental
INHS	Italian National Health Service
IQWiG	German Institute for Quality and Efficiency in Healthcare
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
KALM	KALM-1 and KALM-2 were randomised trials to study the safety and efficacy of difelikefalin in haemodialysis patients with moderate-to-severe pruritus
Kg	Kilogram
KSR	Kleijnen Systematic Reviews Limited
Kt/V	Clearance of urea multiplied by dialysis duration, normalised for urea distribution volume
LH	Lawrence, Hung
LS	Least squares
LTRA	Leukotriene receptor antagonist
LYG	Life years gained
MAR	Missing-at-random
Max	Maximum
Mcg	Microgram
MCMC	Markov Chain Monte Carlo
MD	Mean difference
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical subject headings
MI	Multiple imputation
Min	Minimum
MJ	Matters of judgement
MMRM	Mixed-effects model for repeated measures
MNAR	Missing not at random
N/A	Not applicable
NHS	National Health Service
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NL	The Netherlands
NMA	Network meta-analysis
No.	Number
NRS	Numerical rating scale
OLE	Open label extension
ONS	Office for National Statistics
OOWS	Objective Opiate Withdrawal Scale
OR	Odds ratio
PAS	Patient Access Scheme
PBAC	Pharmaceuticals Benefits Advisory Committee
Pbo	Placebo
PGIC	Patient Global Impression of Change
PICOS	Population, Intervention, Comparison, Outcomes and Study
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTH	Parathyroid hormone

QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
ROBIS	Risk of Bias in Systematic Reviews
RoM	Ratio of means
RR	Relative risk; Risk ratio
RRT	Renal replacement therapy
SAE	Serious adverse event
SAP	Statistical analysis plan
ScHARR	School of Health and Related Research
ScHARRHUD	University of Sheffield Health Utilities Database
SD	Standard deviation
SE	Standard error
ShOWS	Short Opiate Withdrawal Scale
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
STA	Single Technology Appraisal
TEAE	Treatment emergent adverse events
UK	United Kingdom
UKRR	UK Renal Registry
UMC+	University Medical Centre
US	United States
USA	United States of America
VAS	Visual analogue scale
WI-NRS	Worst Itching Intensity Numerical Rating Scale
WTP	Willingness-to-pay
Y	Years

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Assessment Group (EAG) as being potentially important for decision making. Where possible, it also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 relates to the clinical effectiveness evidence, and Section 1.5 covers issues related to the cost effectiveness (CE) evidence. Other key issues are discussed in Section 1.6 while a summary is presented in Section 1.7.

Background information on the condition, technology, evidence and information on key as well as non-key issues are in the main EAG report, see Sections 2 (decision problem), 3 (clinical effectiveness), and 4 and 5 (CE) for more details.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1.1: Summary of key issues

ID1457	Summary of issue	Report Sections
1	The population in the decision problem and the included trials appears narrower than that in the National Institute for Health and Care Excellence (NICE) final scope. The decision problem and trial populations preclude first line treatment and are restricted to people receiving in-centre haemodialysis (ICHD). The NICE scope makes no restrictions in terms of ICHD and treatment line.	2.1
2	The comparison in the included trials is difelikefalin plus established clinical management (ECM) versus placebo plus ECM, whereas the comparison in the NICE final scope and the decision problem is difelikefalin versus ECM. The nature of the treatment comparison in the trials may lead to a more optimistic impression of the study drug's benefits compared with the NICE final scope/decision problem.	2.2 and 2.3
3	A systematic literature review (SLR) of clinical effectiveness evidence was not carried out. This made it difficult to determine whether all relevant studies were included in the clinical effectiveness part of the submission. The Evidence Assessment Group (EAG) identified two potentially relevant randomised controlled trials (RCTs) that had not been considered within the company submission (CS). Data from these RCTs have been added to the report by the EAG.	3.1
4	Differences between ECM in the included trials and the United Kingdom (UK) target population may limit the generalisability of clinical effectiveness evidence from the trials. The company did not provide the results of sub-group analyses in relation to specific anti-itch medications other than difelikefalin. This hindered the evaluation of the impact of differences in the use of non-difelikefalin anti-itch medications between the included trials and the UK target population.	3.2.1.1
5	The included trials recruited a larger proportion of Black participants relative to those seen in the UK target population. Results from sub-group analyses suggested that Black participants tend to have better difelikefalin outcomes than other ethnic groups. This may further affect the generalisability of the overall trial results.	3.2.1.1

ID1457	Summary of issue	Report Sections
6	The rationale for the statistical analysis in the included trials (specifically, multiple imputation (MI) and logistic regression) is not transparent. Lack of information about model inputs and outputs in both instances has hindered the EAG's assessment of the quality of the statistical analyses.	3.2.2
7	Clinical effectiveness data from the KALM-1 and KALM-2 trials were pooled without adjusting for differences between the trials. This may have resulted in biased estimates of treatment effectiveness.	3.2.5.1.1
8	The company assumed in the base case and an alternative scenario that transitions can be modelled by only looking at the probability of shifting between one, two or three health states up or down, regardless of the current health state. This assumption did not seem to be supported by the directly estimated transition probabilities.	4.2.6.1 and 4.2.6.2
9	In the estimation of the transition probabilities, use has been made of data that was imputed multiple times to account for missing data. It is unclear to the EAG how all transition matrices were derived in light of the multiple imputation.	4.2.6.2
10	In the base case, treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In the absence of further real-world evidence to support the waning impact of ECM and/or the lack of waning over time with difelikefalin, the EAG considers this assumption uncertain.	4.2.6.2
11	The company applies an increased risk of death for patients in the moderate, severe and very severe health states of the chronic kidney disease-associated pruritis (CKD-aP) population, based on an observational study. The EAG considers the evidence presented not substantial enough to establish a causal relationship between pruritus and mortality of these patients.	4.2.6.3

CKD-ap = chronic kidney disease-associated puritis; CS = company submission; EAG = Evidence Assessment Group; ECM = established clinical management; ICHD = in-centre haemodialysis; MI = multiple imputation; NICE = National Institute for Health and Care Excellence; RCT = randomised controlled trial; SLR = systematic literature review; UK = United Kingdom

1.2 Overview of key model outcomes

National Institute for Health and Care Excellence technology appraisals compare how much a new technology improves length (overall survival) and quality of life (QoL) in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increasing the number of patients in better health states (lower itch score), thus improving their health-related quality of life (HRQoL)

Overall, the technology is modelled to affect costs by:

- The costs of difelikefalin, which are added to current treatment
- Increasing the number of patients in better health states (lower itch score), thus lower the costs of management of pruritus

The modelling assumptions that have the greatest effect on the ICER are:

- How the transition probabilities should be estimated from the clinical data
- Alternative assumptions after 64 weeks regarding waning of the treatment effect
- Inclusion of a relationship between level of itching and mortality
- Inclusion of the costs of mortality (only in combination with the previous assumption).

1.3 The decision problem: summary of the EAG's key issues

The decision problem addressed in the company submission (CS[1]) is broadly in line with the final scope issued by NICE. However, there are discrepancies in terms of the breadth of population (Table 1.2) as well as the actual intervention and comparators (Table 1.3).

Table 1.2: Key issue 1: Population in decision problem is narrowed relative to the NICE scope

Report Section	**2.1**
Description of issue and why the EAG has identified it as important	The population in the trials appears to be narrower than the National Institute for Health and Care Excellence (NICE) scope population. The decision problem, as indicated by the place in the care pathway "*where established clinical management is insufficient in reducing pruritus*", precludes first line treatment, and is restricted to in-centre haemodialysis (ICHD), whereas the NICE scope makes no restrictions in terms of ICHD or treatment line. The trial population included both patients who were currently on anti-itch medication, which they could continue, as well as those who were not. Narrowing of the decision problem scope relative to the NICE scope means that the evidence in the submission is unlikely to be applicable to the whole population receiving haemodialysis for chronic kidney disease (CKD).
What alternative approach has the EAG suggested?	The company needs to confirm that they are seeking a recommendation for the narrower population i.e., ICHD after trial of established clinical management (ECM). Otherwise, the company need to extend their evidence base to encompass the full population defined by the NICE scope.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence is required if there is confirmation of the restriction of the population to post-trial of ECM in ICHD. If the above is not possible, further data are required to cover the full population scope.
CKD = chronic kidney disease; EAG = Evidence Assessment Group; ECM = established clinical management; ICHD = in-centre haemodialysis; NICE = National Institute for Health and Care Excellence	

Table 1.3: Key issue 2: Comparison in evidence base is different to comparison in the NICE scope

Report Sections	**2.2 and 2.3**
Description of issue and why the EAG has identified it as important	The comparison in the trials is difelikefalin plus established clinical management (ECM) versus placebo plus ECM, whereas the National Institute for Health and Care Excellence (NICE) scope comparison is difelikefalin versus ECM. The different comparison may lead to a more optimistic impression of the study drug's benefits. It is also unclear how ECM is related to anti-itch medication. In the cost effectiveness analysis (CEA) the only ECM that is listed is anti-itch medication, but not all patients are assumed to receive it: about 40% to 55% depending on severity of pruritus. In the overall KALM-1 and KALM-2 populations, 35-40% were on anti-itch medication at baseline. Although not significantly different, sub-group analysis seemed to show a small increase in treatment effect (odds ratio) in those who had received baseline anti-itch medication. Of course, lack

Report Sections	2.2 and 2.3
	of anti-itch medication in both arms does not provide a comparison of difelikefalin versus anti-itch medication or ECM.
What alternative approach has the EAG suggested?	The EAG would like it noted that the evidence presented in the company submission (CS) is not suitable for recommendation regarding difelikefalin alone i.e., without ECM. If this is required then the Evidence Assessment Group (EAG) would suggest methods to allow comparison between difelikefalin versus ECM, such as an indirect treatment comparison (ITC).
What is the expected effect on the cost effectiveness estimates?	The cost effectiveness of difelikefalin versus ECM is unknown.
What additional evidence or analyses might help to resolve this key issue?	If a recommendation regarding difelikefalin alone is required then incorporation of additional randomised evidence on ECM versus placebo, to allow creation of a network meta-analysis (NMA) yielding difelikefalin versus ECM estimates is required.
CS = company submission; CEA = cost effectiveness analysis; EAG = Evidence Assessment Group; ECM = established clinical management; ITC = indirect treatment comparison; NICE = National Institute for Health and Care Excellence; NMA = network meta-analysis	

1.4 The clinical effectiveness evidence: summary of the EAG's key issues

The EAG identified five major concerns with the evidence presented on the clinical effectiveness. These included: the omission of a clinical effectiveness systematic literature review (SLR) (see Table 1.4); potential limitations to the external validity of the included trials in terms of the use of anti-itch medications other than difelikefalin (Table 1.5) and ethnicity (Table 1.6); and unclear explanations of the statistical methodology used within the included trials (Table 1.7) and for pooling data from the trials (Table 1.8).

Table 1.4: Key issue 3: Inadequate SLR carried out for clinical effectiveness

Report Section	3.1.3
Description of issue and why the EAG has identified it as important	A systematic literature review (SLR) was not carried out for clinical effectiveness. This makes it difficult to know if all appropriate studies have been included in the submission. The Evidence Assessment Group (EAG) has found two additional trials that have been added to the report.
What alternative approach has the EAG suggested?	The company needs to carry out a full SLR for clinical effectiveness.
What is the expected effect on the cost effectiveness estimates?	The cost effectiveness (CE) might have been spuriously increased or decreased.
What additional evidence or analyses might help to resolve this key issue?	Additional evidence yielded by a full SLR.
CE = cost effectiveness; EAG = Evidence Assessment Group; SLR = systematic literature review	

Table 1.5: Key issue 4: Applicability of trial findings unclear because of lack of anti-itch medication sub-grouping

Report Section	**3.2.1.1**
Description of issue and why the EAG has identified it as important	Any differences between established clinical management (ECM) in the trials and in the United Kingdom (UK) target population may limit the applicability of the trials. No sub-grouping for specific anti-itch medication were carried out by the company to facilitate evaluation of the implications of any such differences.
What alternative approach has the EAG suggested?	Sub-grouping for specific anti-itch medications.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Sub-grouping for specific anti-itch medications, together with data on the anti-itch medications used in the UK target population.
EAG = Evidence Assessment Group; ECM = established clinical management; UK = United Kingdom	

Table 1.6: Key issue 5: Applicability of findings may be reduced by differences in ethnicity between the trials and the UK target population

Report Section	**3.2.1.1**
Description of issue and why the EAG has identified it as important	The trials had a larger proportion of Black participants than the United Kingdom (UK) target population. As sub-grouping showed that Black participants tend to have better difelikefalin outcomes, this may further affect the applicability of the trials.
What alternative approach has the EAG suggested?	There is little that can be done as the data have been collected. However, adjustments of the data are possible (see below).
What is the expected effect on the cost effectiveness estimates?	The cost effectiveness (CE) is likely to have been spuriously increased.
What additional evidence or analyses might help to resolve this key issue?	Possible adjustments of the overall trial effects, taking into account ethnicity effects.
CE = cost effectiveness; EAG = Evidence Assessment Group; UK = United Kingdom	

Table 1.7: Key issue 6: Rationale for statistical analysis unclear

Report Section	**3.2.2**
Description of issue and why the EAG has identified it as important	The rationale for the statistical analysis is not justified regarding both multiple imputation (MI) analysis and logistic regression.
What alternative approach has the EAG suggested?	The appropriateness of the statistical methods and the choices in the models must be rooted in the data and the characteristics of the specific studies. The variables used in the MI and the logistic regression models must be justified conceptually and tested statistically. The results of the analysis should be reported in detail.

Report Section	3.2.2
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Justification on the methodology used. Use of a selection process to determine the variables included in the logistic regression models. Present in detail all the result of the analysis.
EAG = Evidence Assessment Group; MI = multiple imputation	

Table 1.8: Key issue 7: Methods used to pool trials were not appropriate

Report Section	3.2.5.1.1
Description of issue and why the EAG has identified it as important	Clinical effectiveness data from the KALM-1 and KALM-2 trials were pooled without adjusting for differences between the trials. This may have resulted in biased estimates of treatment effectiveness.
What alternative approach has the EAG suggested?	Reanalysis of the pooled data from KALM-1 and KALM-2, adjusting for differences between trials, e.g., by including trial name as a covariate.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Reanalysis of the pooled data from KALM-1 and KALM-2, adjusting for differences between trials, e.g., by including trial name as a covariate.
EAG = Evidence Assessment Group	

1.5 The cost effectiveness evidence: summary of the EAG's key issues

A full summary of the CE evidence review conclusions can be found in Section 6.4 of this report. The company's CE results are presented in Section 5, the EAG's summary and detailed critique in Section 4, and the EAG's amendments to the company's model and results are presented in Section 6. The key issues in the CE evidence are discussed in Tables 1.8 to 1.10

Table 1.9: Key issue 8: Approach estimating transition probabilities

Report Section	4.2.6.1 and 4.2.6.2
Description of issue and why the EAG has identified it as important	The company assumed in the base case and an alternative scenario that transitions can be modelled by only looking at the probability of shifting one, two or three health states up or down, regardless of the current health state. This assumption did not seem to be supported by the directly estimated transition probabilities.
What alternative approach has the EAG suggested?	The Evidence Assessment Group (EAG) prefers to use the directly estimated transition probabilities. Though this increases the uncertainty around each probability (power decreases), no further simplifying assumptions are required.
What is the expected effect on the cost effectiveness estimates?	Compared to the company base case (in which transitions are estimated based on aggregated data), the EAG preferred approach increases the incremental cost-effectiveness ratio (ICER) slightly. The alternative scenario presented by the company where the change of state probabilities is estimated from the observed data increases the ICER significantly.

Report Section	4.2.6.1 and 4.2.6.2
What additional evidence or analyses might help to resolve this key issue?	No additional evidence is required, though external data might increase the precision of the estimated transition probabilities.
EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio	

Table 1.10: Key issue 9: Lack of clarity how multiple imputation was used

Report Section	4.2.6.2
Description of issue and why the EAG has identified it as important	In the estimation of the transition probabilities, use has been made of data that was imputed multiple times to account for missing data. It is unclear to the Evidence Assessment Group (EAG) how all transition matrices were derived in light of the multiple imputation. For example, when looking at the directly estimated transition probabilities as presented in response to the clarification letter, it is unclear if these probabilities are based on averages over 20 different probabilities, each from a different complete dataset.
What alternative approach has the EAG suggested?	None, the EAG would like to see more details of who the 20 imputed data sets were combined to find the estimated transition probabilities.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Information is required how analyses per complete dataset were combined in order to find the final estimates, and how uncertainty was estimated (which should be a function of within dataset variation and between dataset variation). In addition, sensitivity analysis should be done to see how different approaches to deal with the missing data impact the results.
EAG = Evidence Assessment Group	

Table 1.11: Key issue 10: Insufficient evidence regarding transitions after 64 weeks

Report Section	4.2.6.3
Description of issue and why the EAG has identified it as important	In the base case, a treatment waning was modelled for the established clinical management (ECM) arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In absence of further real-world evidence to support the waning impact of the ECM and/or the lack of waning throughout the years with difelikefalin, the EAG considers this assumption uncertain.
What alternative approach has the EAG suggested?	The EAG has assumed that both with and without difelikefalin, patients remain in the health state they were in at 64 week. However, different waning patterns were explored in the EAG's scenario analyses for both the difelikefalin and ECM arms.
What is the expected effect on the cost effectiveness estimates?	If waning is included for both treatment groups, the impact on the incremental cost-effectiveness ratio (ICER) is very small. If 5% waning is assumed for ECM patient but not for difelikefalin patients the ICER decreases by about 14%, and with 10% waning the ICER decreases 24%.
What additional evidence or analyses might help to resolve this key issue?	Real world data could provide information about the long-term disease development in patients with pruritus who receive ECM.

Report Section	4.2.6.3
EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio	

Table 1.12: Key issue 11: Insufficient evidence that decreasing level pruritus improved mortality

Report Section	4.2.6.3
Description of issue and why the EAG has identified it as important	The Evidenced Assessment Group (EAG) does not agree with the company's approach to use an increased risk of death for patients in the moderate, severe and very severe health states of the chronic kidney disease-associated pruritis (CKD-aP) population as the EAG considers the evidence presented not substantial enough to establish a causal relationship between pruritus and mortality of these patients.
What alternative approach has the EAG suggested?	The EAG removed this elevated risk of death for these patients from the model.
What is the expected effect on the cost effectiveness estimates?	In the company base case, which does not include dialysis costs, removing the increased mortality risk decreases the incremental cost-effectiveness ratio (ICER). In contrast, for the EAG base case, which does include dialysis costs, this change leads to a higher ICER.
What additional evidence or analyses might help to resolve this key issue?	Ideally a randomised controlled trial (RCT) with very long follow-up, where one group receives treatment to lessen the itching and the other group does not, to see if this impacts survival. However, this is unlikely to be feasible. Alternatively, attempts may be made to design a long-term observational study in such a way that potential biases are avoided.
CKD-ap = chronic kidney disease-associated pruritis; EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; RCT = randomised controlled trial	

1.6 Summary of the EAG's view

1.6.1 Clinical effectiveness

The omission of a clinical effectiveness SLR means that relevant evidence could have been missed from the submission and the impact of this on clinical and CE estimates is uncertain. The clinical effectiveness evidence in the submission was based primarily on two randomised controlled trials (RCTs) (the KALM-1 and KALM-2 trials) comparing difelikefalin 0.5 mcg/kg with placebo (with some concomitant interventions permitted in both groups before and during the trials) in patients receiving in-centre haemodialysis for end-stage renal disease (ESRD) over a period of 12 weeks. Both RCTs had non-comparative extension phases and the submission also referred to another single-arm study (CLIN3105) which focused on safety aspects. The EAG identified two additional relevant RCTs that were not mentioned in the submission that reported a similar treatment comparison and population as the KALM trials but with a follow-up period of 8 weeks.

The evidence overall suggested that difelikefalin 0.5 mcg/kg was more effective than placebo for reducing itching intensity over an 8 or 12-week course of treatment. For QoL, results were more equivocal. Adverse events (AEs) were generally non-serious, and the rate of serious adverse events (SAEs) was similar across study arms.

The results of the included RCTs may have limited generalisability in terms of the use of anti-itch medication other than difelikefalin and the distribution of different ethnic groups in the target UK population therefore findings should be viewed with caution.

1.6.2 Cost effectiveness

Table 1.13 summarises the ICERs of both the company's and EAG's preferred base cases, as well as the impact of each EAG preferred assumption applied separately to the company base case.

Each of the changes by themselves does increase the ICER slightly, except for the inclusion of the dialysis costs in the model, which leads to a substantial increase. However, when the inclusion of these costs is combined with a removal of the assumption that mortality is higher in patients with more severe pruritus, the inclusion does not impact the ICER at all, since the number of life years in each arm will be the same and hence the total costs of dialysis per group.

Combining all changes in the model lead to a EAG preferred base case incremental cost effectiveness results of £35,048 per QALY gained, which is higher than the company ICER of £23,277 per QALY gained.

The probabilistic ICER, £41,157 per QALY gained, is higher than the EAG deterministic base case. This is due to the skewness in the distribution around the transition probabilities whenever these are very close to zero, i.e., a 0% transition in the deterministic analysis will become a small but non-0% transition in the probabilistic sensitivity analysis (PSA). The PSA shows that the probability that difelikefalin combined with established clinical management (ECM) is cost effective at thresholds of £20,000 and £30,000 per QALY gained are 0% and 13%, respectively, using the EAG base case assumptions.

Several scenarios were explored, and most of these led to only small changes in the ICER. The most substantial change occurred when transition probabilities were derived using the observed data to estimate to probability of a change of state, independent on the current state. This scenario yielded an ICER of £51,521 per QALY gained.

Table 1.13: Individual impact of EAG preferred assumptions

Preferred assumption	DFK plus ECM		ECM		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Company base case (original)	[redacted]	[redacted]	£30,442	2.75	[redacted]	[redacted]	£24,293
Company base case (after clarification)	[redacted]	[redacted]	**£23,644**	**2.75**	[redacted]	[redacted]	**£23,277**
EAG change on transition probabilities	[redacted]	[redacted]	£23,590	2.76	[redacted]	[redacted]	£25,792
EAG change on waning effect for the ECM arm	[redacted]	[redacted]	£23,626	2.78	[redacted]	[redacted]	£26,320
EAG change on elevated risk of death	[redacted]	[redacted]	£24,476	2.84	[redacted]	[redacted]	£27,566

Preferred assumption	DFK plus ECM		ECM		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
for patients in moderate, severe and very severe health states							
EAG change on cost of haemodialysis	[redacted]	[redacted]	£92,732	2.75	[redacted]	[redacted]	£33,723
EAG's preferred base case	[redacted]	[redacted]	**£97,611**	**2.88**	[redacted]	[redacted]	**£35,048**

Based on the EAG preferred version of the electronic model
DFK = difelikefalin; EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; Incr. = incremental; QALY = quality-adjusted life year

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
Population	Adults with moderate-to-severe pruritus receiving haemodialysis.	For the treatment of moderate-to-severe pruritus associated with CKD in adult patients receiving ICHD, including where established clinical management (ECM) is insufficient in reducing pruritus.	An update was made as difelikefalin is restricted for ICHD use only.	The company's decision to narrow the population to those having ICHD is in line with the Summary of Product Characteristics (SmPC) which states that "*Kapruvia should be restricted for in-centre haemodialysis use only.*" (p.2).[2] However, there is some ambiguity about the scope of the decision problem, and whether it is narrower than the National Institute for Health and Care Excellence (NICE) scope. This is a result of the clause relating to the sufficiency of ECM in the decision problem. The term 'including' suggests that the decision problem population is not restricted to those where ECM is insufficient, comprising people for whom ECM is both sufficient and insufficient. In this case there would be no conflict with the NICE scope (which would include all adults with severe pruritis receiving haemodialysis, regardless of ECM effectiveness). However, clarification from the company was sought, which demonstrated that the company definitively regards difelikefalin as a second (or later) line drug, implying that 'including' really means 'wholly comprising'. Therefore, the

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
				decision problem defined by the company is narrower than the NICE scope.
Intervention	Difelikefalin	Difelikefalin	No change from scope.	The intervention is given as 'difelikefalin 0.5 mcg/kg' in Table 3 of the CS[1]. There are some suggestions elsewhere in the CS[1] that the actual intervention given in the trials was difelikefalin 0.5 mcg/kg plus ECM, but it is very unclear. If the actual intervention given was difelikefalin 0.5 mcg/kg plus ECM, the placebo arm must also have been given ECM because they were double blinded trials. This means that the trials comprised the comparison: *difelikefalin + ECM versus placebo + ECM*, which is not the same as the NICE scope comparison of *difelikefalin versus ECM.* This is a major departure from the NICE scope. There could also be effects on external validity if the ECM given to both arms in the trials differs from that given to the United Kingdom (UK) target population. In such a case the trials may not be applicable to the target population.
Comparator(s)	Established clinical management without difelikefalin, including gabapentin and pregabalin.	Established clinical management without difelikefalin, including gabapentin and pregabalin.	No change from scope.	Despite the company's claim that they have addressed the NICE scope comparator in their decision problem, thus covering the requested comparison of difelikefalin versus ECM, they have not. The KALM trials ostensibly compare difelikefalin (intervention) versus placebo (comparator),

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
				although in reality this is more likely to be *difelikefalin + ECM versus placebo + ECM*. This is a major problem as the effect from the difelikefalin plus ECM versus placebo plus ECM comparison is likely to be more optimistic than the desired difelikefalin versus ECM comparison, and therefore cannot be used as a valid substitute.
Outcomes	The outcome measures to be considered include: • Itching intensity • Adverse effects of treatment • Health-related quality of life (HRQoL)	As per NICE final scope.	No change from scope.	No EAG comments.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year (QALY). The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the	As per NICE final scope.	No change from scope.	No EAG comments.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
	technologies being compared. Costs will be considered from a National Health Service (NHS) and Personal Social Services (PSS) perspective.			
Subgroups to be considered	None specified.	People with anti-pruritic medication use at baseline. People without anti-pruritic medication use at baseline. People with severe or very severe CKD-aP at baseline.	Currently, there are no approved treatments for CKD-aP. The KALM trials did not directly include any comparator treatments, although patients using anti-itch medication at baseline were allowed to continue doing so. It was deemed relevant to analyse subgroups based on use of anti-pruritic medication at baseline. The third subgroup was included to examine the impact of difelikefalin in the most severe CKD-aP category.	The company implied that ethnicity, gender and age were potential outcome modifiers (see page 22 in Document B, CS[1]). However no sub-grouping was performed for these potential covariates. The company has been asked to perform sub-grouping for these variables, and their analysis shows that 'race' may be an effect modifier. Given the discrepancies between the UK target population and the KALM trials in the proportions of people in different 'race' categories, this finding has implications for the applicability of the trial findings.
Special considerations including issues related to equity or equality	People in lower socio-economic groups are more likely to develop chronic kidney disease (CKD), progress towards kidney failure, and die earlier with CKD. People from Black, Asian, and minority ethnic populations are more	As per NICE final scope.	No change from scope.	No EAG comments.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
	likely to progress to kidney failure faster and less likely to receive a transplant. Women are more likely to be diagnosed with CKD, but less likely to start dialysis. Older people with CKD are less likely to receive a kidney transplant than their younger counterparts. These populations are at greater risk of developing CKD-associated pruritus (CKD-aP) and experiencing symptoms for longer while on dialysis. Therefore, guidance on the use of difelikefalin could have a different impact on people with protected characteristics than on the wider population.[3] Difelikefalin is restricted for in-centre haemodialysis (ICHD) use only. This may be considered to represent a barrier to some patients			

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
	for whom ICHD is not accessible.			

Based on Table 1 of the CS[1]

CKD = chronic kidney disease; CKD-aP = CKD-associated pruritis; CS = company submission; EAG = Evidence Assessment Group; ECM = established clinical management; HRQoL = health-related quality of life; ICHD = in-centre haemodialysis; ITC = indirect treatment comparison; KALM = KALM-1 and KALM-2 were randomised trials to study the safety and efficacy of difelikefalin in haemodialysis patients with moderate-to-severe pruritus; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; QALY = quality-adjusted life year; SmPC = Summary of Product Characteristics; UK = United Kingdom

2.1 Population

The population in the National Institute for Health and Care Excellence (NICE) final scope was: '*Adults with moderate-to-severe pruritus receiving haemodialysis*.'[4] However, the decision problem was slightly different: '*For the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis, including where established clinical management is insufficient in reducing pruritus*.' The company justified the restriction to in-centre haemodialysis (ICHD) use in the decision problem, on the grounds that difelikefalin is restricted for ICHD use only.[1]

EAG comment:

- The company's decision to narrow the population to those having ICHD use appears to make sense, if difelikefalin is 'restricted to in-centre haemodialysis use only'. This is also consistent with the KALM-1 and KALM-2 trial evidence which restricted patients to ICHD for the duration of the studies. There is some further ambiguity about the scope of the decision problem, and whether it is narrower than the NICE scope. This is a result of the clause relating to the sufficiency of established clinical management (ECM) in the decision problem. The term 'including' (*including where established clinical management is insufficient in reducing pruritus*) suggests inclusivity; that is, that the decision problem population is not restricted to those where ECM '*is insufficient*', and comprises people for whom ECM is *either* sufficient, insufficient or not tried. In this case there would be no conflict with the NICE scope, which would include all adults with severe pruritis receiving haemodialysis, regardless of ECM effectiveness. However, if this interpretation of the word 'including' is incorrect, and it indicates that all eligible patients must respond insufficiently to ECM, the decision problem will be restricted to a second line or later population. This possibility is suggested by the company's statement in Section B1.3 that "*If a patient has failed on best supportive care this is when difelikefalin will be offered for the duration of dialysis, as long as a sufficient reduction in itch score has been achieved within the first 12 weeks of treatment.*". In addition, in Figure 3 of Document B, it is stated that "*difelikefalin be used as an adjunct to established clinical management where established clinical management is insufficient in reducing pruritis*".[1]
- The company were asked to clarify: 1) that the decision problem wording implies a restriction to a later line of therapy than first line; 2) which treatments as part of ECM need to have been tried before determining insufficiency in reducing pruritus; and 3) the criteria which would need to be applied in clinical practice for this determination. The company responded by stating that: "*Guidelines recommend ensuring adequate dialysis, normalising the calcium-phosphate balance, controlling parathyroid hormones (PTH) to acceptable levels, correcting any anaemia, and using simple emollients before employing other treatment strategies. If a patient is still suffering from pruritus the next stage is to use best supportive care, including creams and emollients, antihistamines, gabapentin and in some cases ultraviolet therapy or antidepressants. Those on no interventions are also deemed to be on best supportive care. If a patient has failed on first line treatment (best supportive care), difelikefalin will be offered for the duration of dialysis, as long as a sufficient reduction in itch score has been achieved within the first 12 weeks of treatment….. Difelikefalin is to be prescribed after best supportive care has failed.*".[5] This response suggests that difelikefalin is intended for use as a later line of therapy. Therefore, the Evidence Assessment Group (EAG) would conclude that the company decision problem is narrowed relative to the NICE final scope. The acceptability of this discrepancy rests on the premise that difelikefalin should never be a first line treatment. If this premise is true, the NICE final scope may have been defined too

broadly. However, if difelikefalin was considered by NICE in scoping as a suitable first line drug (as the NICE final scope population definition suggests), and that it should be tested in that context as well, then this may have implications for the applicability of findings. This is because results relevant to a specific group given later line treatment may not be relevant to all adults with moderate-to-severe pruritus receiving haemodialysis (as per the NICE final scope).[4]

- Furthermore, it was unclear if the submission population included only adults with Stage 5 chronic kidney disease/end-stage renal disease (CKD/ESRD), and the company were asked to clarify if the submission population of ESRD patients with moderate-to-severe pruritus is narrower than the NICE final scope, which is "*adults with moderate-to-severe pruritus receiving haemodialysis*".[4] The company responded by stating that, "*The submission population is the full population covered by the marketing authorisation for difelikefalin. Difelikefalin is indicated for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients on haemodialysis. Chronic kidney disease patients on haemodialysis have, by definition, end-stage renal disease (ESRD) and are, by definition, Stage 5 CKD…. Moderate to severe pruritus is observed in Stage 4 CKD and potentially in acute renal failure patients on haemodialysis. However, such patients are not included in the authorised indication for difelikefalin nor, therefore, the submission population*".[5] This response shows that there is indeed apparent narrowing (relative to the NICE final scope) of the decision problem to ESRD patients, although the NICE final scope in the remit/appraisal objective states: *"To appraise the clinical and cost effectiveness of difelikefalin within its marketing authorisation for treating pruritus associated with chronic kidney disease in adults having haemodialysis."*. Therefore, the EAG conclude that the decision problem is consistent with the scope in this respect, even though CKD stage was not mentioned.
- To summarise, considering the population in the decision problem and the included trials is narrower than that in the NICE final scope, the EAG has highlighted this as a key issue.

2.2 Intervention

The intervention in the NICE final scope was difelikefalin,[4] which is the same as that reported by the company in the decision problem.[1]

EAG comment:

- The intervention is given as "*difelikefalin 0.5 mcg/kg*" in Table 3 of the company submission (CS).[1] There are some suggestions elsewhere in the CS[1] that the actual intervention given was difelikefalin 0.5 mcg/kg plus ECM. For example, it is stated that "*It is proposed that difelikefalin be used as an adjunct to established clinical management where established clinical management is insufficient in reducing pruritus.*" (page 100, CS[1]). This is also how it is described in the cost effectiveness analysis (CEA). However, it is unclear if this relates to the use of the drug during the KALM randomised controlled trials (RCTs). Document B[1] and the clinical study reports (CSRs)[6, 7] do not explicitly state that ECM was or was not used alongside difelikefalin, which adds to the lack of clarity. Therefore, the company were asked to comment on this and confirmed that: "*the submission intervention is Difelikefalin + established clinical management…".*[5] As the actual intervention given was difelikefalin 0.5 mcg/kg plus ECM, the EAG assumes that patients in the placebo arm were given placebo plus ECM since KALM-1 and KALM-2 were double-blind RCTs. It seems reasonable to assume that adequate randomisation would lead to similar ECM provision across arms, making the actual comparison *difelikefalin 0.5 mcg/kg plus ECM versus placebo plus ECM*. It is tempting to conclude that this comparison effectively simplifies, by a process of cancellation of comparable ECM effects in each arm, to the much simpler *difelikefalin 0.5 mcg/kg versus placebo*, which is clearly different from the NICE scope comparison of *difelikefalin 0.5 mcg/kg*

versus ECM. However, this may be too simplistic, because it ignores the possibility of differential interaction effects between difelikefalin and ECM. For example, in the difelikefalin group, the presence of ECM may have a greater additive effect on any benefits from difelikefalin than might be observed from the presence of the same type and level of ECM on any benefits from placebo in the comparator arm. This could arise because ECM increases the potency of difelikefalin (or *vice versa*). This could result in incomplete cancellation of ECM effects across arms, and therefore the comparison would not simplify to difelikefalin versus ECM. Nevertheless, this is not particularly important, since *difelikefalin 0.5 mcg/kg plus ECM versus placebo plus ECM* is clearly not the same as *difelikefalin 0.5 mcg/kg versus ECM*. The only situation in which *difelikefalin 0.5 mcg/kg plus ECM versus placebo plus ECM* could be comparable to *difelikefalin versus ECM* would be if ECM were rendered completely inert by difelikefalin in the (expected) presence of no interaction between placebo and ECM. Established clinical management being rendered completely inert by difelikefalin is extremely unlikely. Thus, in all conceivable cases, it is likely that the NICE final scope comparison has not been achieved and the EAG has highlighted this possibility as a key issue.

- In addition to the nature of the intervention in terms of whether it is in addition to ECM, is the ambiguity of what constitutes ECM. In the CEA, ECM is costed by only anti-itch medications, but not all patients were assumed to take them (about 40% to 55% depending on severity of pruritus). This use is based on the KALM-1 and KALM-2 trials: indeed, only about 35% to 40% of patients were taking them at baseline (and permitted to continue use during the trials). The company were asked to clarify if the submission is indeed a departure from the NICE final scope, and, if so, how they plan to provide evidence that meets the NICE final scope. The company responded by stating that for the trials both arms received ECM, in line with their current management: "*For the KALM trials there were no changes made to current established management. Where a patient was on no prior medication, this was still considered to be established clinical management*." The company then stated that "*the submission intervention is Difelikefalin + established clinical management, with a comparator of established clinical management.*" and went on to claim that "*the submission is in line with the licenced indication and reflects usage in the KALM-1 and KALM-2 trials".*[5]
- Subgroup analysis did reveal little difference by baseline use, but the question still remains as to the nature of ECM and, of course, baseline use or no baseline use of anti-itch medication in both arms is not the same as difelikefalin alone versus anti-itch medication or difelikefalin versus ECM.
- There may be an impact on external validity if the ECM provided in the trials differs from that in the 'real-world' United Kingdom (UK) target population. Background (ECM) treatments can affect external validity via a theoretical 'swamping' effect. Consider the following *reductio ad absurdum* argument. If the background treatments are themselves so effective that they lead to maximal effectiveness in *both* arms of the trial, the interventions themselves cannot manifest any treatment effect (for all the 'work' has been done by the background treatments, and there is nothing more to be done). On the other hand, if the background treatments are so ineffectual that they have no effect on the outcome at all, then any treatment difference between the evaluated interventions can be fully realised. This argument should demonstrate that the nature of the background treatments has a material effect on the magnitude of the final treatment effect that is observed, and so the potential impact of ECM type on external validity is evident. Therefore, it is important to know the exact nature of the ECMs used in the trial so that judgements can be made about the applicability of trial findings to the UK target population.
- The additional sub-grouping carried out by the company in response to the clarification questions[5] (see EAG comments in Section 3.2.1.1) suggested a trend for the benefits of difelikefalin over placebo (in terms of improvement in the Worst Itching Intensity Numerical Rating Scale [WI-NRS]) to be increased if anti-itch medication, antihistamines, opioids or steroids are used with

difelikefalin. The opposite effect is seen with gabapentin/pregabalin, where the presence of this ECM reduced the benefits of difelikefalin over placebo. All these effects were most noticeable in the KALM-1 trial. This information should be used by the committee, in conjunction with clinical knowledge of the ECM used in the UK, to evaluate the applicability of overall trial findings to the UK target population.

2.3 Comparators

The comparator in the NICE final scope was "*Established clinical management (ECM) without difelikefalin, including gabapentin and pregabalin*".[4] The company stated that their decision problem was the same as the NICE final scope.[1]

EAG comment:

- The issue regarding the nature of ECM has already been discussed in Section 2.2 as it relates to its use in addition to difelikefalin. In addition, the precise nature of ECM in terms of the extent of inclusion of anti-itch medication of various kinds raises an issue of generalisability of the KALM-1 and KALM-2 trials (see Section 3.2.1.1), and the cost effectiveness (CE) evidence informed by those trials.

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- Itching intensity
- Adverse effects of treatment
- Health-related quality of life (HRQoL)

These were all assessed in the KALM trials.[1]

EAG comment: Quality of life is assessed through two separate outcomes, namely the 5-D itch score and the Skindex-10 scale.[1] Rationale for using both is not provided by the company. The EAG noted the possibility of increased probability of a type I error with use of more than one outcome measure.

2.5 Other relevant factors

The marketing authorisation for difelikefalin (Kapruvia®) is for the treatment of moderate-to-severe pruritus associated with CKD in adult patients on haemodialysis. The CS (Section B.1.1) stated that the submission covered the full marketing authorisation.[1]

The Food and Drug Administration (FDA) has approved difelikefalin for treatment of moderate-to-severe pruritus associated with CKD in adults undergoing haemodialysis, the first agent approved from a novel class of kappa opioid receptor agonists.[8]

The company highlighted that people in lower socio-economic groups are more likely to develop CKD, progress towards kidney failure, and die earlier with CKD. People from Black, Asian and minority ethnic populations are more likely to progress to kidney failure faster and less likely to receive a transplant. Women are more likely to be diagnosed with CKD, but less likely to start dialysis. Older people with CKD are less likely to receive a kidney transplant than their younger counterparts. These populations are at greater risk of developing CKD-associated pruritis (CKD-aP) and experiencing symptoms for longer while on dialysis. Therefore, guidance on the use of difelikefalin could have a different impact on people with protected characteristics compared to the wider population.[3] Difelikefalin is restricted for ICHD use only, which may be considered as a barrier to some patients for whom ICHD is less accessible. (CS, Section B.1.4[1]).

EAG comment: Despite the above concerns, sub-group analyses according to such potentially important outcome modifiers was not conducted by the company. The company was asked to provide details of outcomes sub-grouped for race, gender and age. The company's response is detailed in Section 3.2.1.1.[5]

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The CS did not report a systematic literature review (SLR) of clinical effectiveness evidence and instead went straight into the reporting of two RCTs (KALM-1 and KALM-2) which had single-arm extension phases plus an additional single-arm study (CLIN3105).[1] An accompanying set of appendices reported methods and results for four other SLRs: treatment pathway/options (to identify standard of care of adults with CKD-aP); evidence on utilities; costs and resource use; and economic modelling studies.[9] This Section (3.1) focuses on aspects of the clinical effectiveness evidence whilst Sections 4.1 discusses the other reviews.

3.1.1 Searches

The EAG queried the omission of any reference to a SLR to inform the clinical effectiveness section and asked if any searches were undertaken to identify RCTs, observational studies or adverse events (AEs). In their response to clarification the company explained that "*A full systematic literature review was not performed for clinical effectiveness searches to identify randomised control trials, observational studies, or adverse events, as it is known that the number of RCTs in this disease area is limited.*"[5] As part of their justification the company provided results of two searches performed in ClinicalTrials.gov. One contained terms for 'CKD-aP' and 'Chronic Kidney Disease associated pruritis' (n=13) and a second used the terms 'Difelikefalin' and 'CR845' (n=23). The EAG does not agree with this approach or accept the results of the two searches of ClinicalTrials.gov as being sufficient to ensure that no relevant data were missed. NICE clearly state in the manual for health technology evaluations that "*Whatever the sources of evidence available on a particular technology and patient group, a systematic review of the relevant evidence relating to a technology should be done using a pre-defined protocol. This protocol should allow evidence to be included from all sources likely to inform the decision about using the technologies by the NHS. A systematic review attempts to assemble all the available relevant evidence using explicit, valid and replicable methods in a way that minimises the risk of biased selection of studies.*"[10] Further to this, guidance by the Centre for Reviews and Dissemination (CRD) recommends that if searches have been limited by a study design filter (in this case RCTs), additional searches should be undertaken to ensure that AEs that are long-term, rare or unanticipated are not missed.[11] The lack of any appropriate searches means that it is likely that some relevant sources will not have been included in the company's report. Whilst the EAG was unable to undertake a full independent SLR and review the results within the Single Technology Appraisal (STA) timeline, they did conduct two focused searches on the Embase database (shown in Appendix 1). The first for difelikefalin identified two additional papers not reported in the original CS.[12, 13] The second search combined terms for (pruritus plus haemodialysis plus RCTs/observational studies) and retrieved several trials of potentially relevant comparator interventions as well as a relevant network meta-analysis (NMA) which is discussed further in Section 3.2.[14]

3.1.2 Inclusion criteria

The CS did not report an SLR of clinical effectiveness evidence. Although clinical effectiveness evidence from individual studies was discussed, there was no information about how these studies were selected.

EAG comment: The possibility of study selection bias cannot be discounted.

3.1.3 Critique of data extraction

Although details of individual studies were tabulated and discussed in the narrative, the CS did not provide any information about the data extraction approach or process.

EAG comment: With no evidence of a pre-specified plan for data extraction, it is possible that there are inaccuracies in the recorded data.

3.1.4 Quality assessment

The CS did not mention methods to assess the methodological quality of the studies described and did not present any information in relation to study quality/validity.

EAG comment: With no evidence of a pre-specific plan for methodological quality assessment, it is possible that the potential impact of methodological flaws on study results were not adequately considered.

3.1.5 Conclusions on the clinical effectiveness systematic review methods

The CS did not include a SLR of clinical effectiveness evidence. Section B.2.1 of the CS states that 'Appendix D, G, H and I – SLR results' provides "*full details of the process and methods used to identify and select the clinical evidence relevant to the technology being evaluated.*"[9] However, scrutiny of this appendix suggests that whilst certain individual studies were described, there were no details of the methods used to identify and select clinical effectiveness evidence.

EAG comment:

- In the clarification letter, the EAG asked the company to provide full details of the methods of the clinical effectiveness SLR including: the review question; study eligibility criteria; search strategy; data extraction approach; critical appraisal; and methods of pooling data. In their response, the company stated that, "*A full systematic literature review was not performed for clinical effectiveness searches to identify randomised control trials, observational studies, or adverse events, as it is known that the number of RCTs in this disease area is limited*." The EAG considers that an anticipated low volume of relevant evidence does not justify the omission of a systematic approach to identify, appraise and synthesise all available relevant data.
- To explore the potential impact of omitting the SLR of clinical effectiveness evidence, the EAG carried out an informal search from which a relevant SLR was identified.[15] The SLR included three RCTs comparing difelikefalin with placebo in haemodialysis-treated patients with uraemic pruritis. One RCT was a report of the KALM-1 trial included in the CS[16] whilst the other two were also apparently relevant but not considered within the submission.[12, 13] All three RCTs are discussed further in Section 3.2.1.
- The EAG considers that the absence of a clinical effectiveness SLR is a major omission which may have led to the CS failing to consider all relevant studies. The EAG is concerned that the company has deviated from guidance in the 'NICE user guide for company evidence submission template' which clearly states that a SLR of clinical effectiveness evidence is required as part of the CS.[17] The EAG is aware that the NICE user guide mentions an option for companies to refrain from undertaking a clinical effectiveness SLR in exceptional circumstances however, cannot see that any such circumstances apply in this instance. In summary, the omission of a SLR of clinical effectiveness evidence means that the clinical- and cost effectiveness estimates reported in the CS may not be derived from a complete assembly of relevant evidence and the risk of study selection bias cannot be discounted. Considering this, the EAG has highlighted this matter as a key issue.

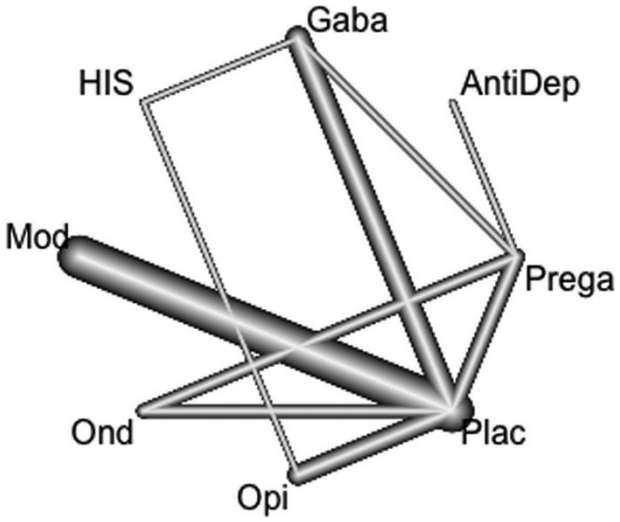
3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

In the CS,[1] the company considered three studies, KALM-1[6, 18] and KALM-2,[7, 19] each consisting of a double-blinded, placebo-controlled phase and an open label extension (OLE) phase, and CLIN3105,[20] an open label single arm study. The results of these studies were used to inform the efficacy and safety outcomes of the CS[1] for difelikefalin for the treatment of moderate-to-severe pruritus associated with CKD in adult patients receiving ICHD.

EAG comment:

- Given that the SLRs may not have been adequate to survey the literature relevant to this submission (please see comments in previous Section), it is possible that the current submission may not be based on the full set of available relevant data.
- Further searching by the EAG showed the above suspicion to be true. Two RCTs (Fishbane 2020[12] and Narita 2022[13]) have also been added to the EAG report.
- A recent NMA by Feng 2020[14] focused on the efficacy of uraemic pruritus treatments in patients undergoing haemodialysis. Twenty-one studies were included with evidence on seven different treatments (medication/class of medication). Three opioid pathway related treatments were included: naltrexone and nalbuphine (both mentioned as disallowed concomitant medication in KALM-1 and KALM-2 unless needed for the treatment of AEs or emergent medical conditions) and the kappa-opioid receptor agonist nalfurafine (not mentioned in relation to concomitant medication during the KALM trials). In addition, gabapentin, pregabalin, antihistamines and antidepressants, which were all allowed in the KALM trials were also included in the network, as shown in Figure 3.1. The results of the NMA showed that opioid pathway related treatments, gabapentin, pregabalin, antihistamines and haemodialysis prescription modification had statistically significant improvement on uremic pruritus compared to placebo. Only two treatments (the serotonin receptor antagonist ondansetron and the antidepressant doxepin) did not have significantly different results compared to placebo. The outcome appeared to be a change in uraemic pruritis measured by visual analogue scale (VAS). The meta-analysis was not accompanied by a systematic review but was reasonably well conducted. A risk of bias assessment of the individual studies was not executed. In addition, only papers in English were considered and there was limited information about the nature of the outcome. Notwithstanding these limitations, the study illustrates the existence of further scientific evidence on available treatments as well as on the feasibility of indirect treatment comparison (ITC) analysis which was not included in the CS.

Figure 3.1: Network diagram of included treatments in the meta-analysis by Feng 2021



Based on Figure 2, Feng 2021[14]
AntiDep = antidepressant; Gaba = gabapentin; HIS = antihistamine; Mod = haemodialysis prescription modification; Ond = ondansetron; Opi = opioid pathway related treatment; Plac = placebo; Prega = pregabalin

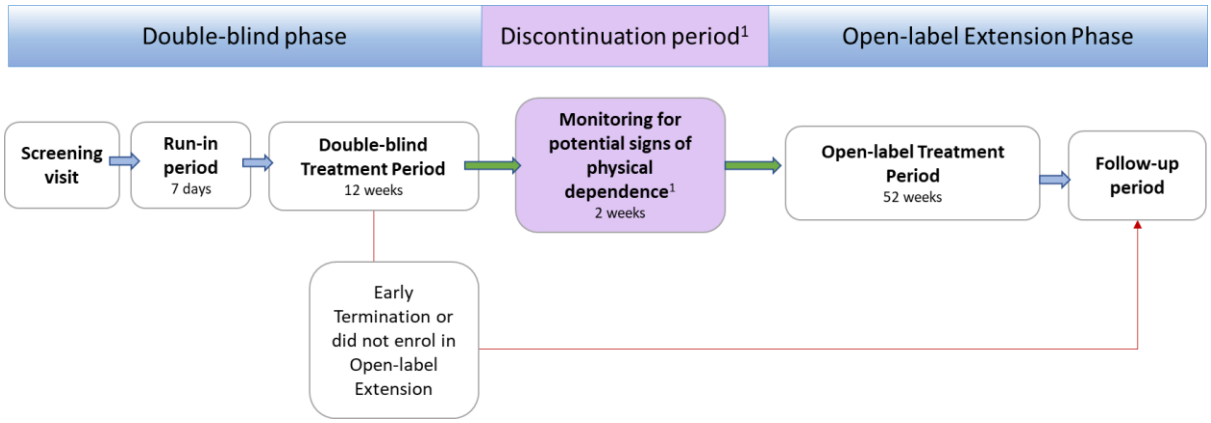
3.2.1 Details of the included trials

3.2.1.1 KALM-1 and KALM-2

Each study comprised a double-blind and an OLE phase. The double-blinded phase evaluated difelikefalin versus placebo at a dose of 0.5 mcg/kg administered after each haemodialysis session (3 times per week) in people with moderate-to-severe pruritus, for a total of up to 36 doses. Both studies were phase 3, randomised, multicentred, placebo-controlled trials. The trial design is illustrated in Figure 3.2. They consisted of a 7-day run-in period during the week prior to randomisation and a 12-week double-blind treatment period where difelikefalin was evaluated relative to placebo. During the run-in period the baseline itch intensity as well as the moderate-to-severe pruritus status was established, while treatment with anti-itch medications and presence of other medical conditions was recorded. The two latter were also used for randomisation stratification. The run-in period was followed by a 12-week double-blind treatment period and a 52-week OLE period. In KALM-1 alone there was a 2-week discontinuation period between the two phases, during which the patients were evaluated for signs of physical dependence. The primary efficacy outcome was based on the WI-NRS score, more specifically, the proportion of patients achieving ≥3-points reduction from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at week 12. An overview of the studies including details of eligibility criteria, locations, concomitant medications and outcomes are reported in Table 3.1.

The OLE phases of KALM-1 and KALM-2 had the same objectives as the double-blinded phases with a focus on safety of the drug administered at the same dose and frequency for up to a 52-week period. The maintenance of the treatment effect of difelikefalin regarding long-term use was also evaluated. A follow-up visit took place 7-10 days either after the end of treatment or the early termination visit. The overview of the OLE phases of KALM-1 and KALM-2 is summarised in Table 3.3.

Figure 3.2: Trial design of KALM-1, KALM-2 double-blind and OLE phases



Based on Figure 4 of Document B of the CS[1]
CS = company submission; OLE = open label extension
[1]The discontinuation period is only applicable to KALM-1 and not KALM-2

Table 3.1: KALM-1 and KALM-2 (double-blinded phase) study overview and summary of methodology

Study	CLIN3102 (KALM-1)	CLIN3103 (KALM-2)
Study design	Phase 3 randomised, 12-week, double-blind, placebo-controlled study	
Population/ Eligibility criteria	Adults (≥18 years of age) with end-stage renal disease (ESRD) who had been on haemodialysis (HD) at least 3 times per week for at least 3 months and who had moderate-to-severe chronic kidney disease-associated pruritis (CKD-aP) defined as a weekly mean score of >4 points on the 24-hour Worst Itching Intensity Numerical Rating Scale (WI-NRS).	Eligible patients were adults (18 to 85 years of age) with ESRD who had been on HD at least 3 times per week for at least 3 months, and who had moderate-to-severe CKD-aP (defined as a weekly mean score >4 on the 24-hour WI-NRS).
N	A total of 378 patients were enrolled between February 2018 and December 2018.	A total of 474 patients were enrolled between July 2018 and February 2020.
Location	57 centres in the United States (US).	93 centres in the US, Australia, Canada, Czech Republic, Germany, Hungary, South Korea, New Zealand, Poland, Taiwan, and the United Kingdom (UK)
Intervention/ Comparator (N)	Patients were randomised 1:1 to receive either • intravenous (IV) difelikefalin (0.5 mcg/kg) (N=189) • placebo (N=189)	Patients were randomised 1:1 to receive either • IV difelikefalin (0.5 mcg/kg) (N=237) • placebo (N=236)
Administration of treatment	The study drug was dispensed by qualified staff members who had received training on study drug handling and administration. Patients received difelikefalin at a dose of 0.5 mcg/kg or placebo after each HD session, generally 3 times per week for up to 12 weeks. Treatment was administered as an IV bolus into the venous line of the HD circuit either during or after rinse back at the end of each HD session.	
Permitted and disallowed concomitant medication	Concomitant medication during the treatment period was restricted as follows: • Investigational drug (other than the study drug) – not allowed • Ultraviolet light-B treatments – not allowed • Naloxone, naltrexone, or mixed agonist-antagonists (e.g., buprenorphine and nalbuphine) - not allowed from the start of dosing of the double-blind treatment period to the end of the open-label treatment period (or from screening to the end of the treatment period for CLIN3105), unless needed for acute treatment of an adverse event or emergent medical condition. • Antihistamines (oral, IV, or topical), corticosteroids (oral, IV, or topical), opioids, gabapentin, or pregabalin - changes to current prescription were to be avoided from screening to the end of the treatment period, unless for the acute treatment of an adverse event or emergent medical condition (in this case, the study Medical Monitor was to be notified and, as appropriate, the adverse event(s) were to be reported). • No new medication to treat itch was to be initiated.	

Study	CLIN3102 (KALM-1)	CLIN3103 (KALM-2)
Indicate if study supports application for marketing authorisation	Yes	
Indicate if study used in the economic model	Yes	
Reported outcomes specified in the decision problem	Primary efficacy outcome: • Proportion of patients achieving ≥3-point reduction from baseline in weekly mean WI-NRS score (week 12)	Primary efficacy outcome: • Proportion of patients achieving at least a 3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at week 12
	Secondary efficacy outcomes: • Change from baseline in health-related quality of life (HRQoL) measured using the Skindex-10 scale total score (week 12) • Change from baseline in HRQoL measured using the 5-D Itch scale total score (week 12) • Proportion of patients achieving ≥4-point reduction from baseline in weekly mean WI-NRS score (week 12)	Secondary efficacy outcomes: • Proportion of patients achieving ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at week 4 and week 8 • Proportion of patients achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at week 4 and week 8 and week 12 • Change from baseline in itch-related HRQoL at the end of week 12, as assessed by the Skindex-10 scale total score • Change from baseline in itch-related HRQoL at the end of Week 12, as assessed by the 5-D Itch scale total score
	Safety: • Severity and seriousness of adverse events and their relationship to study drug	
Outcomes used in the economic model	Primary outcome: For the model, the 5-D Itch score was used. The company's rationale for this was that as 5-D Itch scale total scores provide estimates of treatment efficacy for up to 64-weeks compared with only 12-weeks using WI-NRS, they were used to inform efficacy estimates within the model base case.[1]	
	Secondary outcomes:	

Study	CLIN3102 (KALM-1)	CLIN3103 (KALM-2)
	• WI-NRS total score at baseline, week 4, week 8, and week 12 • Proportion of patients achieving ≥3-point reduction from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score at week 4, week 8, and week 12 • 5-D Itch total score at baseline, week 4, week 8 and week 12	
Pre-planned subgroups	Interim analysis patients and post-interim analysis. By stratification factor: • Use of anti-itch medication at baseline • Presence of specific medical conditions at baseline	
		• By region • By dialysis type (HD or haemodiafiltration)

Based on Tables 3, 5 and 12 of Document B of the CS[1]
CKD-aP = chronic kidney disease-associated pruritis; CS = company submission; ESRD = end-stage renal disease; HD = haemodialysis; HRQoL = health-related quality of life; IV = intravenous; kg = kilogramme; mcg = micrograms; UK = United Kingdom; US = United States; WI-NRS = Worst Itching Intensity Numerical Rating Scale

Table 3.2: KALM-1, KALM-2 (OLE phase) and CLIN3105 study overview and summary of methodology

Study	KALM-1 OLE	KALM-2 OLE	CLIN3105
Study design	Phase 3, open-label, multicentre, long-term (52-week) extension safety study		Phase 3, open-label, global, multicentre, safety and efficacy study
Population/ Eligibility criteria	Adults (≥18 years of age) with end-stage renal disease (ESRD) who had been on haemodialysis (HD) at least 3 times per week for at least 3 months, who had moderate-to-severe chronic kidney disease-associated pruritis (CKD-aP), and who had received at least 30 doses of difelikefalin in the double-blind phase of KALM-1.		Adults (18-85 years of age) with ESRD who had been on HD at least three times per week for at least three months and who had moderate-to-severe CKD-aP.
N	313	399	222
Location	57 centres in the United States (US)	93 centres in the US, Australia, Canada, Czech Republic, Germany, Hungary, South Korea, New Zealand, Poland, Taiwan, and the United Kingdom (UK)	43 centres across the US, Czech Republic, Hungary, and Poland
Intervention	Difelikefalin 0.5 mcg/kg	Difelikefalin 0.5 mcg/kg	Difelikefalin 0.5 mcg/kg
Administration of treatment	Patients received difelikefalin at a dose of 0.5 mcg/kg after each HD session, generally 3 times per week for up to 52 weeks. This was in addition to the treatments received during the double-blind phase (0.5 mcg/kg after each HD session, generally 3 times per week for up to 12 weeks). Treatment was administered as an intravenous (IV) bolus into the venous line of the HD circuit either during or after rinse back at the end of each HD session.		Patients received difelikefalin 3 times per week for up to 12 weeks, for a total of up to 36 doses. Difelikefalin was administered as a 0.5 mcg/kg IV bolus into the venous line at the end of haemodialysis, either during rinse back or after rinse back.
Permitted and disallowed concomitant medication	Concomitant medication during the treatment period was restricted as follows: • Investigational drug (other than the study drug) – not allowed • Ultraviolet light-B treatments – not allowed • Naloxone, naltrexone, or mixed agonist-antagonists (e.g., buprenorphine and nalbuphine) - not allowed from the start of dosing of the double-blind treatment period to the end of the open-label treatment period (or from screening to the end of the treatment period for CLIN3105), unless needed for acute treatment of an adverse event or emergent medical condition. • Antihistamines (oral, IV, or topical), corticosteroids (oral, IV, or topical), opioids, gabapentin, or pregabalin - changes to current prescription were to be avoided from screening to the end of the treatment period, unless for the acute treatment of an adverse event or emergent medical condition (in this case, the study Medical Monitor was to be notified and, as appropriate, the adverse event(s) were to be reported). • No new medication to treat itch was to be initiated.		

Study	KALM-1 OLE	KALM-2 OLE	CLIN3105
Indicate if study supports application for marketing authorisation	Yes		-
Indicate if study used in the economic model	Yes		-
Outcomes	The following assessments were used to evaluate the safety of difelikefalin in patients undergoing haemodialysis and experiencing moderate-to-severe pruritus: • Adverse events • Vital signs • Electrocardiograms • Clinical laboratory values		
Used in the economic model	• 5-D Itch total score at baseline (week 12 of double-blind phase) and at week 52 • Adverse events		No outcomes used in the economic model.
Pre-planned subgroups	No pre-planned subgroups		

Based on Tables 4, 6 and 12 of Document B of the CS[1]
CKD-aP = chronic kidney disease-associated pruritis; CS = company submission; ESRD = end-stage renal disease; HD = haemodialysis; IV = intravenous; mcg = micrograms; OLE = open label extension; WI-NRS = Worst Itching Intensity Numerical Rating Scale, UK = United Kingdom; US = United States

EAG comment:

Discontinuation period

- The double-blinded period of KALM-1 was followed by a 2-week discontinuation period before the OLE phase started, during which the patients were evaluated for signs of physical dependence. The company were asked to discuss why the same design was not also followed by KALM-2, and to discuss whether the 2-week discontinuation period influenced the efficacy and safety results. The company responded by stating that, "*Results from KALM-1 were considered robust enough not to warrant interrupting patient treatment in KALM-2… Comparing the results from KALM-1 with KALM-2 suggests there is no obvious effect of the 2-week discontinuation period on the efficacy and safety of difelikefalin over the 64 weeks of treatment.*"[5] The EAG is satisfied with this response.

Concomitant treatments

- It was initially unclear from the trial documentation whether background treatments (ECM) were given to the patients, and, if so, which ones were given. For example, it is stated in the CS that, "*It is proposed that difelikefalin be used as an adjunct to established clinical management where established clinical management is insufficient in reducing pruritus*." (page 100, CS[1]). Furthermore, Table 12 of Document B of the CS reports that the permitted concomitant medications include antihistamines, corticosteroids and opioids, as well as gabapentin, or pregabalin.[1] On the other hand, the CSRs for both KALM-1 and KALM-2 state that patients who "*had received new or changed prescription for opioids, gabapentin, or pregabalin within 14 days prior to screening*" would be excluded for the double-blinded phases (Section 9.3.2),[6, 7] implying that if a patient was already receiving any of them at a certain dose 14 days prior to screening, they would be allowed to enrol in the trial and continue using them at the same dose. During the clarification phase, the EAG asked the company to confirm if patients could be treated with opioids, gabapentin or pregabalin during the trial, and if opioid antagonists were permitted (as opposed to agonists). The company responded that, "*Subjects were allowed to be treated with the comparators. The only restriction was that changes to current prescription should be avoid from screening to the end of the double-blind treatment period, unless needed for the acute treatment of AEs or emergent medical conditions (as per study protocols)….. The opioid antagonists naloxone and naltrexone were not permitted to be used from the start of dosing of the double-blind treatment period to the end of the open-label treatment period, unless needed to for the acute treatment of an adverse event or emergent medical condition. The same was true of mixed agonist-antagonists such as buprenorphine and nalbuphine.*"[5] The EAG appreciates the detail and clarity of this response.
- In the clarification letter, the EAG asked the company to tabulate the concomitant medications. The company provided the following table of the pooled concomitant medications from KALM-1 and KALM-2.[5]

Table 3.3: A summary of the pooled concomitant medications from KALM-1 and KALM-2

Medication	Placebo (n=425)	Difelikefalin (n=426)
Any baseline use of an anti-itch medication	163 (38.4%)	159 (37.3%)
Most commonly used anti-itch medications at baseline (>2%)		
Diphenhydramine	100 (23.5%)	104 (24.4%)
Hydroxyzine	52 (12.2%)	42 (9.9%)
Hydrocortisone	16 (3.8%)	11 (2.6%)
Cetirizine	10 (2.4%)	7 (1.6%)
Clemastine	10 (2.4%)	7 (1.6%)
Based on Company's response to clarification[5]		

- Since the array of concomitant treatments used might alter during the trial, the company was asked to report the number of patients with concomitant medication prescription adjustments due to AEs or medical conditions. In relation to prescription adjustments, the company stated that, "*A review of Listing 16.2.10.1 (Prior and Concomitant Medications) indicates that 11 patients in each group had adjustments to their concomitant anti-itch medication during the double -blind period in the KALM-1 study and 6 patients in each group had adjustments to their concomitant anti-itch medication during the double -blind period in the KALM-2 study*."[5] The EAG appreciate these additional data and do not see clear evidence of any between-arm differences likely to affect internal validity.
- Document B contains several references to the use of antidepressants as part of ECM, but these are not mentioned in Table 12 (summary of methodology of KALM-1, KALM-2 and CLIN3105) in the CS as being allowed or disallowed from the clinical effectiveness studies.[21] During the clarification phase, the EAG asked the company to confirm whether antidepressants were permitted as a concomitant intervention in the clinical effectiveness studies. The company stated that "*Antidepressants were permitted if they were part of established (>2 weeks) clinical management for a patient*."[5] The EAG is satisfied with this response.
- The anti-asthma leukotriene receptor antagonist (LTRA) montelukast, was listed in Table 55 of the CS as "*established clinical management for CKD-associated pruritus*".[1] The company was asked to provide evidence of the use of montelukast as an established intervention for CKD-aP and to clarify whether it was a permitted concomitant trial medication. The company directed the EAG to the treatment pathway SLR of CKD-aP reported in 'Appendix D, G, H and I – SLR results'[9] where evidence to support the use of montelukast as a treatment option for CKD-aP was cited.[22, 23] The company also confirmed that montelukast was a permitted concomitant medication in the KALM trials.[5] The EAG is satisfied with this response.

Applicability

- The EAG had concerns about the applicability of the trial results to the UK target population. This is based upon two sets of characteristics that were identified by the EAG as possibly differing between trials and the target population: 1) the concomitant anti-itch medications used and 2) race, gender and age. This concern will now be discussed in detail, in relation to each of these two sets of characteristics.

1. Applicability - concomitant anti-itch medications used.

The EAG requested information on the specific concomitant anti-itch medication used in each trial, together with the mode of action and whether they are part of ECM. The following tables (Table 3.5 and Table 3.6) were provided by the company in response.[5]

Table 3.4: Details of concomitant medication in KALM-1

Medication generic name (drug class)	Mode of action	Included in ECM?	Placebo (N=188) n (%)	DFK (N=189) n (%)	All patients (N=377) n (%)
DIPHENHYDRAMINE (antihistamine)	Inhibits the effects of histamines in the body, providing symptomatic relief of itching	Yes	71 (37.8%)	63 (33.3%)	134 (35.5%)
HYDROXYZINE (antihistamine)		Yes	19 (10.1%)	18 (9.5%)	37 (9.8%)
HYDROCORTISONE (corticosteroid)	Inhibits immune response by modifying the function of dermal cells, epidermal cells and leucocytes, reducing itch and inflammation	Yes	8 (4.3%)	6 (3.2%)	14 (3.7%)
TRIAMCINOLONE (corticosteroid)		Yes	3 (1.6%)	5 (2.6%)	8 (2.1%)
AMMONIUM LACTATE (topical emollient)	Promotes moisturisation and hydration of skin and provides symptomatic relief of itching	Yes	4 (2.1%)	2 (1.1%)	6 (1.6%)

Based on Response to clarification[5]
DFK = difelikefalin; ECM = established clinical management

Table 3.5: Details of concomitant medication in KALM-2

Medication (drug class)	Mode of action	Included in ECM?	Placebo (N=188) n (%)	DFK (N=189) n (%)	All patients (N=377) n (%)
DIPHENHYDRAMINE (antihistamine)	Reduces the effects of histamines in the body, providing symptomatic relief of itching	Yes	26 (11.0%)	45 (19.1%)	71 (15.1%)
HYDROXYZINE (antihistamine)		Yes	27 (11.4%)	22 (9.4%)	49 (10.4%)
CLEMASTINE (antihistamine)		Yes	10 (4.2%)	8 (3.4%)	18 (3.8%)
CETIRIZINE (antihistamine)		Yes	7 (3.0%)	4 (1.7%)	11 (2.3%)
LORATADINE (antihistamine)		Yes	4 (1.7%)	5 (2.1%)	9 (1.9%)
CHLORPHENAMINE (antihistamine)		Yes	5 (2.1%)	2 (0.9%)	7 (1.5%)
HYDROCORTISONE (corticosteroid)	Inhibits immune response by modifying the function of dermal cells, epidermal cells and leucocytes, reducing itch and inflammation	Yes	8 (3.4%)	4 (1.7%)	(2.5%)

Based on Response to clarification[5]
DFK = difelikefalin; ECM = established clinical management

- In terms of applicability, it is important to consider whether the overall array of anti-itch medications allowed in the KALM trials is comparable to those received in standard UK care (see Section 2.2). The CS did not include information on the use of anti-itch medications in the UK population with CKD-aP.
- If any differences in anti-itch medications used between trial and target population do exist, then it is vital that data on the outcomes for trial participants using the various anti-itch medications are examined. This will allow any effect on outcome arising from differences between trial and target population in anti-itch medication to be assessed. A sub-group analysis by anti-itch medications had not been originally performed by the company, so the company was asked to provide one. The company has responded with data from KALM-1 and KALM-2 studies, sub-grouped by the existence or not of five key ECMs.[5] This has been summarised below for the 12-week (final) data only in Table 3.7.
- These data show that, in general, there is a trend for the benefits of difelikefalin over placebo (in terms of WI-NRS improvement) to be increased if anti-itch medication, antihistamines, opioids or steroids are used with difelikefalin. The opposite effect is seen with gabapentin/pregabalin, where the presence of this ECM reduced the benefits of difelikefalin over placebo. All these effects were more noticeable in the KALM-1 study.
- Comparing the data presented in Tables 3.5, 3.6 and 3.7 there are inconsistencies between the number of patients receiving concomitant medication during the double-blind period. For example, in Table 3.5, regarding KALM-1, 81 patients in the treatment arm and 90 in the placebo arm are reported to be receiving concomitant antihistamines. On the other hand, in Table 3.7, 82 patients in the treatment arm and 96 in the placebo arm are reported to use concomitant antihistamines. Similar discrepancies are noticed in all the anti-itch medication categories. The origin of these inconsistencies and their effect in the CS is not clear.

  Furthermore, it should be noted that the EAG requested a further level of detail to be presented in the subgroup analysis. This would include the same resolution of information reported in Tables 3.5 and 3.6. For example, in Table 3.6, antihistamines have been sub-grouped into six separate drugs, whilst in Table 3.7 the sub-group categories have been limited to a lower resolution, stopping at 'antihistamines'. A greater resolution of sub-grouping would allow for more direct comparison and interpretation of the results, in terms of treatments administered in the UK. For example, if a certain kind of antihistamine tends to be used predominantly in the UK, the results of the sub-group category that accord with that particular drug would be very informative.
- In addition to the concerns about external validity, relatively large between-arm differences are observed within each separate trial (see Tables 3.4 and 3.5), which could represent a threat to internal validity.
- Since differences between ECM in the included trials and the UK target population may limit the generalisability of clinical effectiveness evidence from the trials, the EAG has highlighted this as a key issue.

Table 3.6: Sub-group analyses for KALM-1 and KALM-2 by ECM types; >3 point improvement in WI-NRS from baseline to 12 weeks

Type of concomitant ECM	KALM-1						KALM-2					
	ECM type – Yes			ECM type – No			ECM type – Yes			ECM type – No		
	n		OR (95% CI)	N		OR (95% CI)	n		OR (95% CI)	n		OR (95% CI)
	DIF	Plac		DIF	Plac		DIF	Plac		DIF	Plac	
Anti-itch medication	85	97	3.16 (1.64,6.09)	104	92	2.20 (1.19, 4.06)	94	91	1.76 (0.91,3.39)	143	145	1.51 (0.92,2.48)
Antihistamines	82	96	3.43 (1.77,6.66)	107	93	2.06 (1.11,3.80)	86	88	1.67 (0.86,3.25)	151	148	1.50 (0.92,2.44)
Opioids	48	60	3.69 (1.52,8.98)	141	129	2.36 (1.40,3.97)	55	69	1.74 (0.80,3.79)	182	167	1.47 (0.94,2.30)
Gabapentin/ pregabalin	53	47	1.39 (0.60,3.22)	136	142	3.32 (1.96,5.61)	43	36	1.43 (0.53,3.86)	194	200	1.57 (1.03,2.39)
Steroids	33	31	3.54 (0.60,20.83)	156	158	1.67 (0.83,3.33)	29	37	NA (NA, NA)	208	199	1.54 (1.01,2.34)

Based on Appendix S in the company's response to clarification[24]
CI = confidence interval; DIF = difelikefalin; ECM = established clinical management; OR = odds ratio; Plac= placebo; WI-NRS = Worst Itching Intensity Numerical Rating Scale
Notes: multiple imputation with missing-at-random assumption used. Odds ratio was based on a logistic regression model, adjusting for baseline WI-NRS score.

2. *Applicability - race, gender and age.*

- As previously discussed, the population characteristics of the trial need to be comparable to those of the target population when considering whether the results of a trial are relevant to a particular health service. This is particularly important for those characteristics that have been suggested by the company to be outcome modifiers, such as ethnicity, gender and age (see page 22 in Document B in the CS[1]). To evaluate comparability in terms of race, gender and age between the trials and the UK target population (all those with CKD and pruritis in the UK) it is necessary to know the race, age and gender characteristics of the UK target population however, these data were not reported in the CS.[1]
- There are two ways to obtain such characteristics for the UK target population, and both have their advantages and disadvantages. One way is to measure the characteristics of UK participants in the trials to gain a (probably non-random) sample estimate. Though efficient, this is not always ideal because trial populations may not always be fully representative of the patient population. A more rigorous method is to obtain whole-population datasets of UK patients, but such datasets are not always available. The EAG therefore requested both sources of information from the company. Both sources of information are discussed in the following Sections.
- Because KALM-2 and KALM-2 OLE appear to have recruited patients from study centres in the UK, the company was asked to provide more details about the UK participants in the trial. The company stated that, '*There were 20 GBR subjects (6 DFK and 14 placebo) recruited from 5 centres in the UK*', and provided baseline characteristics for them as follows:[5]

Table 3.7: Baseline characteristics of UK partipants

Baseline characteristic	All patients (N=20)
Number of participants	**20**
Mean age, years (SD)	**64.9 (11.11)**
Male	9 (45.0%)
Female	11 (55.0%)
Ethnicity – n (%)	
Hispanic or Latino	0 (0.0%)
Not Hispanic or Latino	20 (100%)
Not reported	0 (0.0%)
Unknown	0 (0.0%)
Race – n (%)	
Asian	2 (10.0%)
Black or African American	3 (15.0%)
White	14 (70.0%)
Other	1 (5.0%)
Mean prescription dry body weight, kg (SD)	**79.98 (21.658)**
Baseline Worst Itching numerical rating scale (NRS), mean (SD)	**7.31 (1.624)**
Baseline anti-itch medication use – [1] n (%)	
Yes	7 (35.0%)
No	13 (65.0%)

Baseline characteristic	All patients (N=20)
Specific medical conditions? – [1] n (%)	
Yes	4 (20.0%)
No	16 (80.0%)
Mean duration of pruritus, years (SD)	**2.72 (3.407)**
Mean years since diagnosis of end stage renal disease (ESRD), years (SD)	**5.94 (6.552)**
Years since diagnosis of chronic kidney disease (CKD)	
N	20
Mean (SD)	14.29 (11.918)
Years on chronic haemodialysis (HD), mean (SD)	5.67 (6.444)
Aetiology of CKD [2]	
Diabetes	10 (50.0%)
Hypertension	11 (55.0%)
Large vessel disease	0
Glomerulonephritis	3 (15.0%)
Vasculitis	0
Interstitial nephritis	0
Pyelonephritis	1 (5.0%)
Cystic	0
Hereditary	0
Congenital	0
Neoplasms	0
Tumours	0
Urologic	0
Nephrotic syndrome	4 (20.0%)
Unknown	2 (10.0%)
Other	2 (10.0%)

Based on Response to clarification[5]
CKD = chronic kidney disease; ESRD = end stage renal disease; HD = haemodialysis; kg = kilogram; NRS = numerical rating scale; SD = standard deviation; UK = United Kingdom

- The company was also asked to provide data on the proportions of people in different ethnic sub-groups (for example, Asian, Black, White), the mean age, and the proportions of males and females in the UK population among people with CKD and pruritis. The company responded by providing the following Table:[5]

Table 3.8: Demographic data from the UKRR and UK patiens from the pooled KALM studies

	Ethnicity (Race)			Gender	Age
	White	Black	Asian/other	Male	Median
United Kingdom Renal Register (UKRR) Adults	67.6%	12.8%	19.6%	62.3%	66.5

	Ethnicity (Race)			Gender	Age
	White	Black	Asian/other	Male	Median
in-centre haemodialysis (ICHD)					
KALM pooled dataset	60.8%	29.2%	10.0%	59.6%	60.0

Based on Company's response to clarification[5] which in turn used the UKRR[25] and the pooled KALM dataset[26] as data sources
ICHD = in-centre haemodialysis; KALM = KALM-1 and KALM-2 trials; UK = United Kingdom; UKRR = UK Renal Registry

- The company stated that, "*The UK population is slightly older and consists of slightly more white and fewer black patients than the population of patients participating in the KALM studies. The trial data is considered to be representative of a UK population with CKD-aP, verified by clinicians."* The company went on to say that "*Advisors were clear that the KALM-1 and KALM-2 are high quality studies. The majority (7) of the group agreed that the KALM-1 and KALM-2 studies were broadly generalisable to the UK population."*[5] The EAG appreciate the detail of these new data. The data for the UK Renal Register (UKRR) Adults in Table 3.8 tally reasonably well with the data from the subset of UK participants in the KALM trials given in Table 3.8, which shows that the UK participants in the trial were characteristic of the UK target population.
- Importantly, based on the data in Table 3.8, the EAG conclude that there is a relatively strong signal that the overall population in the KALM trials and the UK target population are not comparable in terms of age and race. This contrasts with the views of the clinicians cited by the company however, the data in Table 3.13 suggest that the populations are not comparable. The EAG has highlighted this potential discrepancy between the research and target populations as a key issue.
- If the trial and target population have different characteristics (as discussed above) then this may influence the validity of inferring any effects from the trial to the target population if those differing characteristics are effect modifiers. Effect modification may be inferred from sub-group analysis. A sub-group analysis including the characteristics of race, age and gender is presented in Figure 6 of the pre-proof of the Topf 2022[26] paper which presents a pooled analysis of all four KALM-1 and KALM-2 studies. However, such pooling may obscure important effects within each individual RCT. The company was therefore asked to provide subgroup analysis for race, gender and age for each RCT individually for the primary efficacy results, accompanied by a discussion. In response, the company provided WI-NRS outcome data sub-grouped into age, sex and race categories, as tabulated below.[5]

Table 3.9: Sub-group analyses for KALM-1 and KALM-2 by age, sex and race

Sub-grouping variable	KALM-1: >3 point improvement in WI-NRS from baseline to 12 weeks OR (95% CI) for difelikefalin versus placebo		KALM-2: >3 point improvement in WI-NRS from baseline to 12 weeks OR (95% CI) for difelikefalin versus placebo	
Age categories	**<65 years**	**>65 years**	**<65 years**	**>65 years**
Age	3.38 (1.97,5.79)	1.24 (0.53,2.92)	1.49 (0.92,2.41)	1.69 (0.86,3.32)
Sex categories	**Male**	**Female**	**Male**	**Female**
Sex	3.16 (1.75,5.71)	2.01 (1.02,3.99)	1.17 (0.71,1.92)	2.29 (1.22,4.33)

Sub-grouping variable	KALM-1: >3 point improvement in WI-NRS from baseline to 12 weeks OR (95% CI) for difelikefalin versus placebo			KALM-2: >3 point improvement in WI-NRS from baseline to 12 weeks OR (95% CI) for difelikefalin versus placebo		
Race categories	**White**	**Black or African American**	**Other**	**White**	**Black or African American**	**Other**
Race	2.67 (1.39,5.12)	3.21 (1.60,6.42)	1.10 (0.23,5.20)	1.56 (0.99,2.47)	2.26 (0.89,5.70)	0.68 (0.19,2.50)

Based on Appendix T, response to clarification[24]
CI = confidence interval; KALM = KALM trials; OR = odds ratio; WI-NRS = Worst Itching Intensity Numerical Rating Scale
Notes: Multiple imputation with missing-at-random assumption used. Odds ratio was based on a logistic regression model, adjusting for baseline WI-NRS score

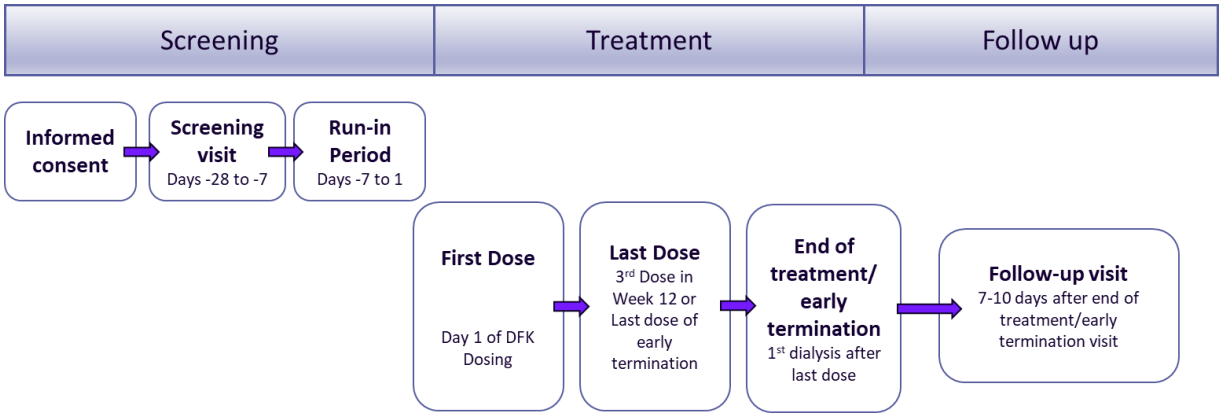
- The sub-grouped results from KALM-1 suggest that being over 65 may reduce the benefits of difelikefalin, but this effect was not observed in KALM-2. This was echoed by the company in their response: "*Although generally response is greater to the <65 subgroup it was numerically higher in older patients in the KALM-2 study*." Similarly, the sub-grouped results from KALM-1 suggest that being female may reduce the benefits of difelikefalin, but an opposite effect was observed in KALM-2.[5] Given this heterogeneity across trials, the possibility that age and sex are effect modifiers is uncertain.
- However, for race, both KALM-1 and KALM-2 trials suggested that being Black or African American improved the efficacy of difelikefalin. Therefore, it is possible that race may be an effect modifier. Given that the KALM trials contained over double the proportion of Black participants (29.2%) as would be found in the UK target population (12.8%) (Table 3.9) this suggests that the overall efficacy observed from the KALM trials may overestimate the efficacy that would be observed in the UK target population.
- Whilst the EAG accept the uncertainty in the sub-group estimates in Table 3.9 and realise that it cannot be definitively concluded that race is an effect modifier, the EAG believe there is enough evidence to suggest consideration of this point by the Committee.
- The EAG also accept that this is a *post-hoc* sub-group analysis. However, it is important to note that the company mentioned age, sex and race as potential effect modifiers in their original CS (see page 22 in Document B, CS[1]) and a sub-group analysis of grouped KALM data has previously been presented in Topf 2022.[26] The company should therefore have chosen these variables, *pre-hoc*, for sub-group analysis in the CS.[1] Thus, the EAG would argue that this finding is not the result of a random *post-hoc* 'fishing exercise', but instead the result of following-up decisions that should have been made *a priori* by the company.
- The company was also asked to provide an illustration of all the subgroup analysis results reported in Appendix E of Document B[1] for ease of comparison. An updated table was provided by the company[5] but has not been reproduced here as it repeats much of the information previously presented. However, the EAG appreciates the increased clarity of presentation.

3.2.1.2 CLIN3105

CLIN3105 was an open-label, multicentre, Phase III study conducted in the United States (US) and Europe. It was designed to evaluate the safety and efficacy of intravenous (IV) difelikefalin at a dose of 0.5 mcg/kg moderate-to-severe CKD-aP in patients undergoing haemodialysis. Patients received difelikefalin as an IV bolus after the end of their dialysis during a treatment period of up to 12 weeks, so that each patient received difelikefalin 3 times per week for a total of up to 36 doses. End of treatment (EOT) was defined as the first day of dialysis following the last dose of the drug. The EOT

procedures were conducted during the dialysis visit following the last dose of the study drug. A final safety follow-up visit was conducted 7 to 10 days after the EOT or early termination visit (Figure 3.3).

Figure 3.3: CLIN3105 study design



Based on Figure 5 of the CS[5]
CS = company submission; DFK = difelikefalin

The screening period to assess eligibility occurred within 28 days prior to treatment and consisted of a screening visit and a run-in period. The purpose of the run-in period was to confirm that each patient had moderate-to-severe pruritus. The screening period was also used to record each patient's use of antipruritic medications.

If patients continued to meet all inclusion criteria and no exclusion criteria at the end of the run-in period, they could start the treatment period and begin treatment with IV difelikefalin 0.5 mcg/kg.

EAG comment:

- It is unclear how the data from the CLIN3105 trial were used in the submission and whether they were used to supplement the data from the KALM-1 and KALM-2 trials. The company was asked to clarify this. The company responded by stating that, "*CLIN3105 was included in the submission, but not included in the economic model because it did not contain a relevant comparator arm. CLIN3105 gathered data on sleep quality using the Sleep Quality Questionnaire. CKD-aP patients often report restless and poor-quality sleep as a result of their itch, causing considerable burden on quality of life; effect on sleep quality is therefore considered an important outcome of difelikefalin. This outcome was not investigated in both KALM-1 and KALM-2, so supplementary data from CLIN3105 is used. Furthermore, CLIN3105 provides real world evidence for difelikefalin with patients in full knowledge of the treatment, as opposed to a blinded trial. For these reasons, CLIN3105 was included in the submission to supplement data provided by KALM-1 and KALM-2.*"[5] This response confirms that the CLIN3105 study is not relevant to this submission because it does not cover the NICE scope outcomes.

3.2.1.3 Narita 2022

Narita 2022 was one of two RCTs not featured in the CS but identified by the EAG as being potentially relevant to the submission as outlined in Section 3.1.5.[13]

Narita 2022[13] was a phase 2 RCT of 247 patients, undergoing treatment with difelikefalin (0.25, 0.5 or 1.0 mcg/kg) or placebo at 94 sites in Japan. Difelikefalin (0.25, 0.5, and 1.0 mcg/kg) and placebo were intravenously administered 3 times a week at the end of each haemodialysis session for 8 weeks.

Like the KALM trials, this study can be regarded as difelikefalin plus ECM versus placebo plus ECM, because anti-pruritis concomitant medication was allowed, provided there was no change to the regimen during the study. However, use of nalfurafine, opioids, and phototherapy were prohibited during the trial.

Participants had ESKD and moderate to severe pruritis (weekly mean score >4 points on numerical rating scale (NRS)), were 20 years or older, and had been having thrice-weekly haemodialysis for at least 12 weeks. They needed to be non-responsive to systemic treatments and/or topical agents, and to have a moderate or severe Shiratori Severity Score for at least two out of the 7 days preceding treatment. Patients with and without a history of using nalfurafine could be included. Exclusion criteria were liver cirrhosis, phototherapy history, previous AEs attributed to nalfurafine and previous hypersensitivity to opioids.

The primary endpoint was the change from baseline in the weekly mean WI-NRS score at week 8. Secondary outcomes included change in itch-related QoL score using the Skindex-16 and 5-D Itch scale. Safety was assessed according to AEs, laboratory tests, vital signs, body weight, and 12-lead electrocardiogram.

3.2.1.4 Fishbane 2020

Fishbane 2020 was the second of two RCTs not featured in the CS but identified by the EAG as being potentially relevant to the submission as outlined in Section 3.1.5.[12]

Fishbane 2020[12] was a randomised, double-blind, placebo-controlled phase 2 trial (ClinicalTrials.gov: NCT02858726) over 8-weeks in haemodialysis patients with moderate-to-severe pruritus. Difelikefalin (0.5, 1.0, and 1.5 mcg/kg) and placebo were intravenously administered 3 times a week at the end of each haemodialysis session for 8 weeks. The study was conducted in the US at 33 sites.

As for the KALM trials and Narita 2022[13] this study can be regarded as comparing difelikefalin plus ECM versus placebo plus ECM, because concomitant medication was allowed to be continued provided that it had been on stable use for 14 days prior to screening. However, patients on opioid antagonists or opioid mixed agonist-antagonists were excluded from the trial.

Inclusion criteria were adults ≥18 years with ESRD who were on haemodialysis 3 times per week for at least 3 months before screening; and persistent pruritus during the month before screening, with weekly mean WI-NRS score over the week before randomisation >4.

The primary endpoint was the change from baseline in the weekly mean WI-NRS score at week 8. Secondary outcomes included change in itch-related QoL score using the Skindex-10 and 5-D Itch scale. Safety was assessed according to AEs, laboratory tests, vital signs, body weight, and 12-lead electrocardiogram.

3.2.2 Statistical analysis of the included studies

3.2.2.1 Studies included in the CS

Different analysis approaches were taken for the outcomes in the CS.[1] The primary outcome was a ≥3-point improvement from baseline at week 12 with respect to the WI-NRS score. In the double-blinded phases of KALM-1 and KALM-2, multiple imputation was used to handle missing data. Subsequently, logistic regression models were implemented to evaluate the primary outcome of efficacy for difelikefalin versus placebo. The variables used in the model were trial group, baseline WI-NRS score, baseline use of antipruritic medication, and history of prespecified medical conditions.

The analysis was divided into two parts: interim and post-interim assessment. The interim assessment was to be executed after 50% of the first 350 patients either completed the 12-week intervention period or discontinued the trial regimen and the post-interim assessment was to include the rest of the patients. The purpose of the interim analysis was to verify that the sample size remaining in the study was sufficient to maintain the power of the calculations. Pooling of the two parts was executed using the Cui, Hung, Wang (CHW) weighted test statistic.[27] The same approach was also taken for the secondary outcome of the proportion of patients achieving ≥4-point reduction from baseline in weekly mean WI-NRS score (week 12).

For the two other secondary outcomes, based on the 5-D Itch scale and the Skindex-10 scale, an analysis of covariance (ANCOVA) model was used. The covariates in the model were the baseline score and stratification factors. Further details on the analysis of the two RCTs, and information on the statistical methodology of the one-arm studies is provided in Table 3.10.

Table 3.10: Statistical analysis summary

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
KALM-1	Primary efficacy outcomes: • The percentage of patients who had an improvement of ≥3 points from baseline at week 12 in the weekly mean score on the daily Worst Itching Intensity Numerical Rating Scale (WI-NRS). Prespecified secondary efficacy outcomes were: • Mean change from baseline at week 12 in the 5-D Itch scale total score • Mean change from baseline at week 12 in the Skindex-10 scale total score • Percentage of patients who had a decrease of at least four points from baseline at week 12 in the weekly mean WI-NRS score.	In the primary analysis, for each imputed data set, the difference between placebo and difelikefalin were analysed using a logistic regression model containing terms for trial group, baseline WI-NRS score, baseline use of antipruritic medication, and history of prespecified medical conditions. The multiple imputation process was implemented separately for patients contributing to the interim assessment and those who underwent randomisation after the interim assessment. The final p-value was calculated with the use of the Cui, Hung, Wang (CHW) weighted test statistic. Testing of the primary outcome was two-sided at an alpha level of 0.05. Secondary outcomes were analysed according to a prespecified hierarchy (first 5-D Itch scale, then Skindex-10 scale, and percentage of patients with a decrease of ≥4 points from baseline to week 12 in the weekly mean WI-NRS score).	A total of 378 patients underwent randomisation. It was calculated that, assuming a response in 30% of the placebo group, a planned sample of 350 patients would result in a 79% to 90% or greater power to detect a difference of 15% to 20% in the primary outcome, on the basis of a two-sided Chi square continuity corrected test at a significance level of 0.05. An interim analysis for sample size re-estimation was conducted by an independent data monitoring committee after 50% of the first 350 patients either completed the 12-week intervention period or discontinued the trial regimen. No change was made to the original enrolment target of 350 patients.	In the primary analysis, missing weekly mean WI-NRS scores were estimated with the use of multiple imputation, under a MAR assumption. WI-NRS scores reported when patients were no longer receiving difelikefalin or placebo after the completion or discontinuation of the trial regimen were censored and treated as missing data.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		The changes in scores on the 5-D Itch and Skindex-10 scales at week 12 were analysed with the use of an analysis of covariance (ANCOVA) model, with trial group as a fixed effect and baseline score and stratification factors as covariates. The percentage of patients who had a decrease of ≥4 points from baseline to week 12 in the weekly mean WI-NRS score was analysed with the use of the method described for the primary outcome. To control the type I error, a gatekeeping strategy was implemented. Testing of the secondary outcomes was to proceed only if the primary efficacy analysis was significant at the 5% level. Testing of the secondary outcomes was two-sided and performed sequentially with an alpha value of 0.05. All the efficacy analyses were conducted in the intention-to-treat (ITT) population, which was defined as all the patients who underwent randomisation.		

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
KALM-1 OLE	Primary efficacy outcome: The change in total 5-D Itch score and change by domain score from baseline. Secondary efficacy outcomes are: • To evaluate the efficacy of intravenous (IV) difelikefalin at a dose of 0.5 mcg/kg compared to placebo in improving itch-related health-related quality of life (HRQoL) measures in haemodialysis patients with moderate-to-severe pruritus. • To evaluate the safety of IV difelikefalin at a dose of 0.5 mcg/kg in haemodialysis patients with moderate-to-severe pruritus.	The 5-D Itch scale was the only measured used to evaluate efficacy in the open label extension (OLE) phase. The 5-D Itch scale scores will be analysed using a mixed model with repeated measures (MMRM). The model will contain treatment sequence, week, and treatment-by-week interaction as fixed effects, and baseline score and the randomisation stratification variables as covariates. Two independent analyses will be presented using different time points for the baseline values and changes from baseline using each of those baselines. In the first analysis, all visits in both the double-blind and the open-label treatment periods will be included; the baseline will be the 5-D Itch scale total score collected on day 1, prior to randomisation. In the second analysis, only the visit in the open-label treatment period will be included; the baseline will be the last 5-D Itch scale total score in the double-blind treatment period.	The sample size for the OLE phase was not defined a priori: all patients who were eligible and willing to continue into the OLE phase were enrolled.	The scoring manual for 5-D Itch scale does not give specific direction regarding scoring when some questions are missing; therefore, each domain and the total score will be set to missing when any of their individual components are missing, with the exception of the disability domain. The maximum of any items present for disability will be used for that domain. Missing data will be handled implicitly in the MMRM model. Assuming that the data are MAR, the estimates calculated from the MMRM described in the statistical analysis.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		An unstructured covariance matrix will be used to model the within-subject errors. Should the model fail to converge, a compound symmetric covariance matrix will be used instead. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Missing scores will not be imputed. Assuming that the data are missing-at-random (MAR), the estimates calculated from the MMRM described above are unbiased. Standard descriptive statistics will be reported for each time point on the values and changes from baseline along with the least squares (LS) means, standard errors, 95% confidence intervals (CIs), and differences from baseline within each treatment sequence reported with LS means, standard errors, and 95% CIs. Plots will also be created. The above analyses for the 5-D Itch scale total score will be repeated for each of its domain scores.		

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
KALM-2	Please see KALM-1 hypothesis objective for primary and secondary efficacy outcomes.	The efficacy of difelikefalin 0.5 mcg/kg compared to placebo in pivotal Phase 3 study KALM-2 will be evaluated based on one primary and seven secondary efficacy endpoints. • The proportion of patients who have an improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS score ≥3 points will be calculated for each imputed dataset. Differences between difelikefalin 0.5 mcg/kg and placebo with respect to the primary endpoint will be compared using a logistic regression model containing terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and presence of specific medical conditions. • The observed number and proportion of patients with ≥ 3-point improvement among the non-imputed data will be reported along	The planned sample size for this study was 350 (175 per treatment group) male and female haemodialysis patients with chronic moderate-to-severe pruritus (mean baseline 24-hour WI-NRS score ≥5), randomised at approximately 95 clinical sites. The sample size calculation was based on results of the completed Phase 2 double-blind, placebo-controlled study CR845-CLIN2101, which evaluated difelikefalin in patients with ESRD and moderate-to-severe pruritus undergoing haemodialysis. Assuming a true response rate of 30% for the placebo group and a true response rate of 50% for the difelikefalin group (defining response as an improvement from baseline ≥3 points with respect to the WI-NRS at week 12), a 2-sided continuity corrected Chi square would have 96% power to detect a treatment difference. The power of this test statistic would be ≥84% for differences from placebo as low as 0.16.	In the primary efficacy analysis, missing NRS data at the end of week 12 will be imputed using a multiple imputation (MI) approach, assuming that patients who discontinue double-blind treatment early would have similar WI-NRS scores as other patients in their respective treatment arm who have complete data: • Intermittent missing NRS scores will first be imputed using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data. • For each stage, MI will be performed within treatment group with covariates for baseline NRS score, both randomisation stratification factors, region and the non-missing NRS scores for each week. Should convergence issues occur due to small cell size for the categorical covariates corresponding to strata (at

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		with the imputed data logistic regression model-based estimates of the proportions of responders, odds ratio, 95% CIs, and p-value.	Based on the results of a planned interim assessment conducted when approximately 50% of the 350 patients had either completed the 12-week double-blind treatment period or had discontinued from treatment early, the size of the study was increased by approximately 20%, to 430 patients.	either stage), those specific covariates will be removed from the model.
KALM-2 OLE		Please see KALM-1 open label extension (OLE) statistical analysis. Additionally: as a separate analysis, the number and percentage of patients who have a 5-point or greater improvement will be reported by visit and treatment sequence. This will be repeated as above for each baseline.	The sample size for the OLE phase was not defined a priori: all patients who were eligible and willing to continue into the OLE phase were enrolled.	The scoring manual does not give specific direction regarding scoring when some questions are missing; therefore, each domain and the total score will be set to missing when any of their individual components are missing, with the exception of the disability domain. The maximum of any items present for disability will be used for that domain. Missing data will be handled implicitly in the MMRM model. Missing scores will not be imputed. Assuming that the data are MAR, the estimates calculated from the MMRM described in the statistical analysis section are unbiased.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
CLIN3105	Primary objective: • To evaluate the safety of difelikefalin at a dose of 0.5 mcg/kg IV in patients undergoing haemodialysis and experiencing moderate-to-severe pruritus. Secondary objectives: • To evaluate the effectiveness of difelikefalin at a dose of 0.5 mcg/kg IV in reducing the intensity of itch in patients undergoing haemodialysis and experiencing moderate-to-severe pruritus • To evaluate the effectiveness of difelikefalin at a dose of 0.5 mcg/kg IV in improving itch-related HRQoL and quality of sleep measures in patients undergoing haemodialysis and experiencing moderate-to-severe pruritus.	This study uses the following five instruments to assess effectiveness: • WI-NRS • Sleep Quality Questionnaire • 5-D Itch scale • Skindex-10 • EQ-5D-5L-P No primary efficacy endpoint was defined. All effectiveness analyses were performed on the safety population. For the WI-NRS, Sleep Quality Questionnaire, 5-D Itch scale, and the Skindex-10 scale, summary statistics (n, mean, standard deviation (SD), minimum, maximum) for the respective baseline and week 12 score were produced, along with the change from baseline. For the WI-NRS and Sleep Quality Questionnaire, the count and percentage of patients with an improvement in WI-NRS from baseline of >0, ≥1, ≥2, ≥3, ≥4, ≥5, and ≥6-points at week 12 were reported. The count and percentage of patients with an improvement from baseline of ≥3	Approximately 200 male and female patients with moderate-to-severe pruritus undergoing haemodialysis were to be enrolled in this study at approximately 50 United States (US) and non-US clinical sites. No sample size calculation was performed to select this sample size.	Missing data will not be imputed. Data from patients who terminated prematurely will be included in any analyses for which their data is available, unless otherwise specified. Please see Section 8.2 of the statistical analysis plan (SAP) for further details.

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		and ≥4-points at week 12 were also reported and stratified by region. Quantitative laboratory parameters were summarised using descriptive statistics for observed values and the changes from baseline to each time point (when applicable), including the designation of last post-baseline treatment visit. Observed measurements of vital signs and the changes from baseline were summarised using descriptive statistics (n, mean, SD, median, minimum, and maximum) for baseline, each post-baseline assessment, and the last post-baseline treatment visit.		

Based on Table 14 of Document B of the CS[1]
5-D Itch scale = 5-dimension Itch scale; ANCOVA = analysis of covariance; CI = confidence interval; CS = company submission; EQ-5D-5L = EuroQol-5 dimension-5 level; ESRD = end-stage renal disease; HRQoL = health-related quality of life; ITT = intention-to-treat; IV = intravenous; kg = kilogramme; LS = least squares; MAR = missing at random; mcg = micrograms; MCMC = Markov Chain Monte Carlo; MI = multiple imputation; MMRM = mixed model with repeated measures; NRS = numerical rating scale; OLE = open label extension; SAP = statistical analysis plan; SD = standard deviation; US = United States; WI-NRS = Worst Itching Intensity Numerical Rating Scale

EAG comment:

- Multiple imputation (MI) was mentioned throughout the CS.[1] The use of MI was built in the statistical analysis plan (SAP) for handling missing NRS data. It is not clear why the SAP presupposed that the proportion and nature of missing data would justify its use. The company was asked to provide the rationale for using MI over other available methods for handling missing data, and to elaborate on the specific methods used within the MI process. The company responded[5] by stating that: '*Thae choice of multiple imputation for the treatment of missing data was suggested by the Food and Drug Administration (FDA) during a meeting held on September 6th 2017 to discuss the Phase 3 clinical development program for IV difelikefalin. Specifically, the FDA stated that "The efficacy analyses should be based on the intent-to-treat (ITT) population (i.e., all randomized subjects)" and added that the protocol "should pre-specify a scientifically sound primary imputation method (e.g. multiple imputation) to handle missing data. Multiple imputation is also one of the analytical methods recommended by the National Research Council Committee on National Statistics in their 2010 report on the prevention and treatment of missing data in clinical trials. Based on these regulatory and technical recommendations, Cara decided to use multiple imputation as the primary method for the treatment of missing data in the pivotal studies KALM-1 and KALM-2.*' The company stated that the specific methods used in the MI process were as follows:
    - *Intermittent missing weekly mean WI-NRS scores were first imputed using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data.*
    - *The monotone missing weekly mean WI-NRS values were then multiply imputed with the SAS MI procedure using the monotone regression method.*
    - *For each stage, MI was performed within treatment group with covariates for baseline WI-NRS score, both randomisation stratification factors, region (in CLIN3103 only), and the non-missing WI-NRS scores for each week. Should convergence issues occur due to small cell size for the categorical covariates corresponding to strata (at either stage), those specific covariates will be removed from the model. For study CLIN3103, the handling of convergence issues related to the region covariate were described in section 8.1.4 of the SAP.*
    - *The proportion of subjects who have an improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score ≥3 points will be calculated for each imputed dataset. Differences between DFK 0.5 mcg/kg and placebo with respect to the primary endpoint were compared using a logistic regression model containing terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and presence of specific medical conditions. For KALM-2, the handling of convergence issues with the region covariate were described in section 8.1.4 of the SAP.*
    - *Twenty imputations were performed.*
    - *Results of the logistic regression on the multiply imputed data sets will be summarised by the SAS MIANALYZE procedure.*
    - *The above MI process was implemented independently among subjects contributing to the interim results and those following the interim analysis. Likewise, the logistic regression and results described above were generated independently for both samples, with the samples combined and adjusted using the methodology proposed by Cui, Hung, Wang (1999).*[27]

  - *Sections 13 and 14 of the SAP for each of the pivotal studies provided sample SAS code in additional to the seeds to be used in the multiple imputation methodology. This pre-specification ensured that the analysis was not data driven.*[5]
- Whilst the above response provides adequate detail on the methodological processes of MI, the EAG does not think that the response provided adequate rationale for the use of MI. The statement that the MI process was used because the FDA told the company to do it this way is insufficient and suggests that the company do not fully understand why this method was used.
- The company was also asked to provide an overview of missingness patterns observed and their frequencies, both for the intermediate missings and the monotone missings. The company responded by stating that: "*The pattern of missing data of the weekly mean WI-NRS in studies KALM-1 and KALM-2 are provided in Tables 3c-1 and 3c-2, respectively in Appendix V as well as the reasoning for this missing data. In both studies, none of the covariates were missing. There were few intermittent missing WINRS scores. In KALM-1, 77.8% of the placebo patients had complete data; this percentage was 74.6% for patients randomised to DFK. A similar result was observed in KALM-2, with 80.9% of the placebo patients having complete data compared to 73.8% of DFK patients. 12.8% of placebo patients and 16.9% of DFK patients in CLIN3102 had a missing weekly WI-NRS score at Week 12, the primary timepoint. Similarly, 12.3% of the patients randomised to placebo had a missing weekly mean WI-NRS at Week 12 compared to 19.4% of the patients randomised to DFK.*"[5]
- The covariates used in the MI analysis for KALM-1 and KALM-2 were not reported in Document B.[1] In clarification question A26, the EAG asked the company to confirm that the covariates for KALM-1 MI analysis were: baseline WI-NRS score, both randomisation stratification factors (use of anti-itch medication during the week prior to randomisation and presence of specific medical conditions), and the non-missing NRS scores for each week, as stated in the SAP. The company responded by stating that: "*We can confirm that the covariates for the KALM-1 (CLIN3102) MI were baseline WI-NRS score, both randomisation stratification factors (use of anti-itch medication during the week prior to randomisation and presence of specific medical conditions), and the non-missing NRS scores for each week*."[5]. Nevertheless, replying to a further question (A27) the company provided a slightly different list of covariates stating that "*The covariates were the baseline mean WI-NRS score, randomization stratification factors, and the prior weeks' mean WI-NRS scores*."[5]
- The company was also asked to confirm that the covariates for KALM-2 MI analysis were: baseline WI-NRS score, both randomisation stratification factors (use of anti-itch medication during the week prior to randomisation and presence of specific medical conditions), region, and the non-missing NRS scores for each week, as stated in the SAP. The company stated that, "*We can confirm that the covariates for the KALM-2 (CLIN3103) MI analysis were baseline WI-NRS score, both randomisation stratification factors (use of anti-itch medication during the week prior to randomisation and presence of specific medical conditions), region, and the non-missing NRS scores for each week.*"[5] The EAG is satisfied with this response.
- The company was asked to provide the rationale for the use of specific covariates over other potential prognostic variables correlated to the outcome of interest. The company stated that, "*The adjustment of the analysis based on the use of anti-itch medication was suggested by the FDA, as there could be potential for differential placebo and DFK response depending on the status of this covariate. In addition, both stratification factors were included as covariates based on recommendations in ICH E9. Region was added to account for possible differences in the patient's responses to treatment based on regional differences. The baseline WI-NRS was included as a covariate because the primary endpoint, based on an improvement from baseline, would be*

correlated with the baseline level. No other prognostic factors were identified as important in the Phase 2 studies that were used to plan the pivotal Phase 3 studies."[5] The EAG is surprised that the prognostic factor of race was not considered, given that it was suggested, *pre-hoc*, as a potential effect modifier by the company, and was later shown by a sub-group analysis to have an effect on outcome.

- The covariates used in the logistic regression analysis were: trial group, baseline WI-NRS score, baseline use of antipruritic medication, and history of prespecified medical conditions in KALM-1 plus region in KALM-2. The company was asked to provide the rationale and validity of using these variables and asked to discuss if other baseline characteristics were considered for use in the logistic regression models, such as gender, race and age. In response the company reiterated the response to the same question on the MI analysis in the previous comment. In a further inquiry on how the variables were selected in the model the company responded that "*The variables to be included in the logistic model were specified a priori in the study protocol and statistical analysis plan. There was no additional selection or de-selection of variables in the model*."[5]. Similar to the MI analysis there is a noteworthy lack of rationale and justification on the conceptualization of the model. The EAG has highlighted this as a key issue.
- The interim analysis was included in the study protocol as a means of sample size re-estimation, should it be required. Regarding KALM-1, the CSR[6] states that "… *there were no changes to the original enrolment target of 350 subjects*" (page 82). Nevertheless, the primary efficacy analysis was conducted separately for interim analysis and post-interim analysis patients. The company were asked to provide the rationale for executing a "*separate*" interim analysis. The company responded by stating that, "*the interim analysis was performed to inform the IDMC recommendation that no change was required to the original enrolment target. This triggered the prespecified requirement for the multiple imputation approach and logistic regression to be implemented independently for subjects contributing data to the interim analysis and subjects contributing data following the interim analysis. The primary analysis was also conducted separately for interim analysis subjects and post-interim analysis subjects to evaluate the potential impact of the interim analysis on the properties of statistical inference at the end of the study.*"[5]
- In addition, the company was asked to provide full results without splitting the data with respect to interim versus post-interim status which are partially reported in Table 21 of Document B.[1] The company responded by stating that, "*Splitting of the data with respect to interim and post-interim status was only applied to the primary efficacy variable i.e. the Worst Itching Intensity Numerical Rating Scale. All other results are presented without splitting the data*."[5]

3.2.2.2 Statistical analysis of Narita 2022

A sample size of 60 participants per group was used based on a common standard deviation (SD) of 2.5 and the following mean differences between the placebo and each of the treatment groups: −0.6 (0.25 μg/kg group), −1.3 (0.5 μg/kg group), and −1.3 (1.0 μg/kg group). This provided 80% power to show superiority of 0.5 μg/kg or 1.0 μg/kg of difelikefalin over placebo, with a 2-sided p value <.05. The intention-to-treat (ITT) approach was followed as the primary analysis for the efficacy outcomes. No imputation was used for missing values. A mixed-effects model for repeated measures (MMRM) was used for the primary analysis. This involved adjustment for baseline values, treatment group, time point, and treatment group by time point interaction. Adverse events were coded using the Medical Dictionary for Regulatory Activities (MedDRA) Japanese Edition, Version 21.1 (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use).[28]

EAG comment: No comments.

3.2.2.3 Statistical analysis of Fishbane 2020

A sample size of 40 patients per arm was calculated based on an assumed SD of 2.4, 80% and a mean difference (MD) of 1.5 points (difelikefalin - placebo) in weekly mean of the 24-hour WI-NRS (alpha = 0.05). A mixed effects model with repeated measures was used, adjusting for treatment, week, the interaction between treatment and week, prior anti-pruritic medication usage and baseline WI-NRS score. The former three variables were fixed effects and the latter two were covariates. Data were assumed to be missing at random and no imputation was used. All efficacy analyses were conducted on the full analysis population (patients who received ≥1 dose, analysed according to planned treatment arm), whereas analyses of safety data were performed on the safety population (patients who received ≥1 dose, analysed according to actual treatment arm).

EAG comment: No ITT analysis was performed, but this does not appear to affect risk of bias as the attrition was very low.

3.2.3 Baseline characteristics

3.2.3.1 KALM RCTs

Baseline characteristics for the KALM-1 and KALM-2 studies are shown in Tables 3.12 and 3.13.

Table 3.11: Baseline characteristics of KALM-1 (double-blind safety population)

Baseline characteristic	**Difelikefalin**	**Placebo**
Number of participants	**(n=189)**	**(n=188)**
Mean age, years (SD)	58.2 (11.16)	56.8 (13.89)
Sex – n (%)		
Male	112 (59.3%)	118 (62.8%)
Female	77 (40.7%)	70 (37.2%)
Ethnicity – n (%)		
Hispanic or Latino	64 (33.9%)	68 (36.2%)
Not Hispanic or Latino	123 (63.8%)	120 (63.8%)
Unknown	2 (1.1%)	0
Race		
American Indian or Alaska Native	6 (3.2%)	5 (2.7%)
Asian	6 (3.2%)	7 (3.7%)
Black or African American	82 (43.4%)	75 (39.9%)
Native Hawaiian or Other Pacific Islander	4 (2.1%)	4 (2.1%)
White	91 (48.1%)	94 (49.5%)
Unknown	1 (0.5%)	2 (1.1%)
Other	1 (0.5%)	2 (1.1%)
Mean prescription dry body weight, kg (SD)	**85.91 (20.264)**	**84.98 (21.084)**
Baseline WI-NRS, mean (SD)	**7.06 (1.439)**	**7.25 (1.606)**
Baseline anti-itch medication use? [1] – n (%)		
Yes	72 (38.1%)	78 (41.5%)
No	117 (61.9%)	110 (58.5%)
Specific medical condition? [1] – n (%)		
Yes	25 (13.2%)	28 (14.9%)

Baseline characteristic	Difelikefalin	Placebo
No	164 (86.8%)	160 (85.1%)
Mean duration of pruritus, years (SD)	3.19 (3.244)	3.45 (3.369)
Mean years since diagnosis of ESRD, years (SD)	4.66 (3.898)	5.66 (5.178)
Years since diagnosis of CKD		
n	187	189
Mean (SD)	6.92 (5.926)	7.03 (5.739)
Years on chronic haemodialysis, mean (SD)	**4.37 (3.982)**	**4.73 (4.219)**
Aetiology of CKD [2]		
Diabetes	107 (56.6%)	94 (50.0%)
Hypertension	129 (68.3%)	139 (73.9%)
Large vessel disease	4 (2.1%)	4 (2.1%)
Glomerulonephritis	7 (3.7%)	8 (4.3%)
Vasculitis	0	0
Interstitial nephritis	1 (0.5%)	0
Pyelonephritis	0	0
Cystic	1 (0.5%)	2 (1.1%)
Hereditary	1 (0.5%)	2 (1.1%)
Congenital	0	0
Neoplasms	1 (0.5%)	1 (0.5%)
Tumours	2 (1.1%)	0
Urologic	0	0
Nephrotic syndrome	2 (1.1%)	4 (2.1%)
Unknown	7 (3.7%)	6 (3.2%)
Other	11 (5.8%)	16 (8.5%)
Based on Table 7 of CS[1]		

Baseline characteristic	Difelikefalin	Placebo
CKD = chronic kidney disease; CS = company submission; SD = standard deviation; ESRD = end stage renal disease; WI-NRS = Worst Itching Intensity Numerical Rating Scale [1] observed stratum values [2] more than one item may have been checked		

Table 3.12: Baseline characteristics KALM-2 (double-blind safety population)

Baseline characteristic	Difelikefalin	Placebo
Number of participants	**236**	**235**
Mean age, years (SD)	**59.7 (13.11)**	**59.6 (13.07)**
Sex – n (%)		
Male	135 (57.4%)	139 (58.9%)
Female	100 (42.6%)	97 (41.1%)
Ethnicity – n (%)		
Hispanic or Latino	68 (28.9%)	68 (28.8%)
Not Hispanic or Latino	163 (69.4%)	166 (70.3%)
Not reported	2 (0.9%)	2 (0.8%)
Unknown	2 (0.9%)	0
Race		
American Indian or Alaska Native	1 (0.4%)	1 (0.4%)
Asian	12 (5.1%)	20 (8.5%)
Black or African American	53 (22.6)	38 (16.1%)
Native Hawaiian or Other Pacific Islander	1 (0.4%)	3 (1.3%)
White	162 (68.9%)	169 (71.6%)
Other	6 (2.6%)	5 (2.1%)
Mean prescription dry body weight, kg (SD)	**81.56 (19.731)**	**79.95 (19.450)**
Baseline WI-NRS, Mean (SD)	**7.27 (1.358)**	**7.12 (1.363)**
Baseline anti-itch medication use? [1] – n (%)		
Yes	87 (37.0%)	85 (36.0%)
No	148 (63.0%)	151 (64.0%)
Specific medical condition? [1] – n (%)		
Yes	41 (17.4%)	37 (14.7%)
No	194 (82.6%)	199 (84.3%)
Mean duration of pruritus, years (SD)	**3.21 (4.567)**	**3.20 (3.184)**
Mean years since diagnosis of ESRD, years (SD)	**5.23 (4.677)**	**5.46 (4.509)**
Years since diagnosis of CKD		
n	234	232
Mean (SD)	9.28 (7.638)	9.76 (7.009)
Years on chronic haemodialysis, mean (SD)	4.83 (4.588)	5.09 (4.327)
Aetiology of CKD [2]		
Diabetes	118 (50.2%)	112 (47.5%)
Hypertension	121 (51.5%)	114 (48.3%)
Large vessel disease	4 (1.7%)	3 (1.3%)

Baseline characteristic	Difelikefalin	Placebo
Glomerulonephritis	14 (6.0%)	17 (7.2%)
Vasculitis	3 (1.3%)	2 (0.8%)
Interstitial nephritis	2 (0.9%)	1 (0.4%)
Pyelonephritis	3 (1.3%)	1 (0.4)
Cystic	18 (7.7%)	16 (6.8%)
Hereditary	13 (5.5%)	6 (2.5%)
Congenital	1 (0.4%)	3 (1.3%)
Neoplasms	0	2 (0.8%)
Tumours	1 (0.4%)	1 (0.4%)
Urologic	6 (2.6%)	9 (3.8%)
Nephrotic syndrome	3 (1.3%)	6 (2.5%)
Unknown	8 (3.4%)	14 (5.9%)
Other	26 (11.1%)	28 (11.0%)

Based on Table 8 of CS[1]
CKD = chronic kidney disease; CS = company submission; SD = standard deviation; ESRD = end stage renal disease; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[1] observed stratum values
[2] more than one item may have been checked

EAG comment:

- In the KALM-1 study there were small differences in the level of baseline severity, with the difelikefalin group having slightly less severe itching and having less usage of anti-itching medication. These small differences are unlikely to have had an important effect on outcome, as the key itching outcome was based on the change from baseline.
- In the KALM-2 study, similarly small differences were observed in baseline itching, but in the opposite direction, with greater severity in the difelikefalin group. Again, these are unlikely to have affected outcome. However, there are large between-arm differences in ethnicity in the KALM-2 study, with more participants in the difelikefalin arm being Black/African American. This may have affected outcome because ethnicity could be an outcome modifier (see Section 2.5). Based on the sub-group analysis results, being Black or African American improves outcome with difelikefalin relative to placebo (see Section 3.2.1.1).
- The number of participants recruited to the double-blind phase of KALM-1 is shown as N=378 in both Table 3 and Figure 21 and N=377 (189 + 188 = 377) in Table 7 of Document B.[1] The company was asked to provide the correct number of participants (overall and per treatment arm) or explain the discrepancy. The company responded by stating that, *"Figure 21 in the document provides an explanation of the discrepancy between the ITT population (378) and those who received the allocated intervention (378). There was one patient in the placebo group who did not meet the entry requirement and therefore did not receive an allocated intervention and was therefore excluded from the double-blind safety population in Table 7. An updated clarification on this has also been included in Table 3 of the amended Document B."*[5, 29] The EAG is satisfied with this response.
- The number of participants recruited to the double-blind phase of KALM-2 is discrepant between Table 5 (N=474), Figure 23 (N=473) and Table 8 (N=471; 236 + 235 = 471) in Document B.[1] The company was asked to provide the correct number of participants (overall and per treatment arm) or explain the discrepancy. The company responded by stating that, "*Figure 23 in the document*

provides an explanation of the discrepancy between the ITT population (473) and those who received the allocated intervention (471). There were two patients in the difelikefalin group who did not receive the allocated intervention and were therefore excluded from the double-blind safety population in Table 8. There was a typo in Table 5. Further clarification on the double-blind safety population in Table 5 has been updated in Document B."[5, 29] The EAG is satisfied with this explanation.

- Specific medical conditions have been lumped together. Given that different conditions may have very different effects on outcome, this is potentially misleading. The company was asked to disaggregate these medical conditions. The company responded by stating that, "*The specific medical conditions are of interest as they are typically associated with the pharmacology of kappa opioid receptor agonists. Stratification was performed to ensure balanced groups for those pre-existing medical conditions so as not to confound the assessment of the safety profile of difelikefalin. However, as they were not identified as potential modifiers of treatment response no analysis of the individual conditions has been performed.*"[5] The EAG considers that this response overlooks the potential for different conditions to have different effects on outcome. In addition, it is unclear from this response if the company stratified using separate "*specific medical conditions*" or stratified using a lumped category of "*specific medical conditions*".[5]

3.2.3.2 OLE and one arm studies

Baseline characteristics for KALM-1 OLE, KALM-2 OLE and CLIN3105 are given in Table 3.13 to Table 3.15.

Table 3.13: Baseline characteristics KALM-1 OLE (open-label safety population)

Baseline characteristic	Pbo/DFK	DFK/DFK
Number of participants	**162**	**151**
Mean age, years (SD)	**57.1 (13.79)**	**58.0 (11.45)**
Sex – n (%)		
Male	102 (63.0%)	87 (57.6%)
Female	60 (37.0%)	64 (42.4%)
Ethnicity – n (%)		
Hispanic or Latino	56 (34.6%)	52 (34.4%)
Not Hispanic or Latino	106 (65.4%)	98 (64.9%)
Unknown	0	1 (0.7%)
Race		
American Indian or Alaska Native	4 (2.5%)	3 (2.0%)
Asian	7 (4.3%)	5 (3.3%)
Black or African American	68 (42.0%)	67 (44.4%)
Native Hawaiian or Other Pacific Islander	3 (1.9%)	1 (0.7%)
White	76 (46.9%)	74 (49.0%)
Unknown	2 (1.2%)	0
Other	2 (1.2%)	1 (0.7%)
Mean prescription dry body weight, kg (SD)	**84.53 (20.885)**	**85.87 (20.905)**
Baseline WI-NRS, Mean (SD)	**7.20 (1.586)**	**7.00 (1.440)**

Baseline characteristic	Pbo/DFK	DFK/DFK
Baseline anti-itch medication use? [1] – n (%)		
Yes	64 (39.5%)	54 (35.8%)
No	98 (60.5%)	97 (64.2%)
Specific medical condition? [1] – n (%)		
Yes	23 (14.2%)	22 (14.6%)
No	139 (85.8%)	129 (85.4%)
Mean duration of pruritus, years (SD)	3.53 (3.439)	3.29 (3.492)
Mean years since diagnosis of ESRD, years (SD)	5.77 (5.272)	4.67 (4.011)
Years since diagnosis of CKD		
N	161	151
Mean (SD)	7.0 (5.829)	6.97 (5.995)
Years on chronic haemodialysis, mean (SD)	4.85 (4.404)	4.44 (4.131)
Aetiology of CKD [2]		
Hypertension	120 (74.1%)	107 (70.9%)
Diabetes	82 (50.6%)	82 (54.3%)
Other	13 (8.0%)	8 (5.3%)
Glomerulonephritis	8 (4.9%)	4 (2.6%)
Unknown	5 (3.1%)	5 (3.3%)
Large Vessel Disease	3 (1.9%)	4 (2.6%)
Nephrotic Syndrome	2 (1.2%)	2 (1.3%)
Cystic	2 (1.2%)	1 (0.7%)
Hereditary	2 (1.2%)	1 (0.7%)
Neoplasms	1 (0.6%)	1 (0.7%)
Tumours	0	2 (1.3%)
Interstitial Nephritis	1 (0.6%)	0
Congenital	0	0
Pyelonephritis	0	0
Urologic	0	0
Vasculitis	0	0

Based on Table 9 of CS[1]
CKD = chronic kidney disease; CS = company submission; DFK = difelikefalin; OLE = open label extension; SD = standard deviation; ESRD = end stage renal disease; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[1] observed stratum values
[2] more than one item may have been checked
Note: Baseline characteristics were recorded during the screening visit for the double-blind treatment phase

Table 3.14: Baseline characteristics KALM-2 OLE (open-label safety population)

Baseline characteristic OLE	Pbo/DFK	DFK/DFK
Number of participants	**210**	**189**

Baseline characteristic OLE	Pbo/DFK	DFK/DFK
Mean age, years (SD)	**59.4 (13.13)**	**59.7 (12.88)**
Sex – n (%)		
Male	124 (59.0%)	110 (58.2%)
Female	86 (41.0%)	79 (41.8%)
Ethnicity – n (%)		
Hispanic or Latino	56 (26.7%)	58 (30.7%)
Not Hispanic or Latino	152 (72.4%)	127 (67.2%)
Unknown	0	2 (0.5%)
Not reported	2 (1.0%)	4 (1.0%)
Race		
American Indian or Alaska Native	1 (0.5%)	1 (0.5%)
Asian	16 (7.6%)	9 (4.8%)
Black or African American	33 (15.7%)	39 (20.6%)
Native Hawaiian or Other Pacific Islander	3 (1.4%)	1 (0.5%)
White	153 (72.9%)	135 (71.4%)
Other	4 (1.9%)	4 (2.1%)
Mean prescription dry body weight, kg (SD)	**79.67 (19.227)**	**81.75 (20.326)**
Baseline WI-NRS, Mean (SD)	**7.07 (1.352)**	**7.24 (1.396)**
Baseline anti-itch medication use? [1] – n (%)		
Yes	75 (35.7%)	65 (34.4%)
No	135 (64.3%)	124 (65.6%)
Specific medical condition? [1] – n (%)		
Yes	35 (16.7%)	30 (15.9%)
No	175 (83.3%)	159 (84.1%)
Mean duration of pruritus, years (SD)	3.31 (3.258)	2.92 (2.837)
Mean years since diagnosis of ESRD, years (SD)	5.61 (4.668)	5.19 (4.848)
Years since diagnosis of CKD		
n	206	188
Mean (SD)	10.04 (7.254)	9.29 (7.949)
Years on chronic haemodialysis, mean (SD)	5.23 (4.488)	4.82 (4.797)
Aetiology of CKD [2]		
Hypertension	99 (47.1%)	100 (52.9%)
Diabetes	96 (45.7%)	93 (49.2%)
Other	26 (12.4%)	21 (11.1%)
Cystic	15 (7.1%)	14 (7.4%)
Glomerulonephritis	17 (8.1%)	12 (6.3%)

Baseline characteristic OLE	Pbo/DFK	DFK/DFK
Unknown	13 (6.2%)	7 (3.7%)
Hereditary	5 (2.4%)	12 (6.3%)
Urologic	8 (3.8%)	5 (2.6%)
Nephrotic Syndrome	6 (2.9%)	3 (1.6%)
Large Vessel Disease	3 (1.4%)	4 (2.1%)
Vasculitis	2 (1.0%)	2 (1.1%)
Pyelonephritis	1 (0.5%)	2 (1.1%)
Congenital	2 (1.0%)	0
Interstitial Nephritis	1 (0.5%)	1 (0.5%)
Neoplasms	2 (1.0%)	0
Tumours	1 (0.5%)	1 (0.5%)

Based on Table 10 of CS[1]

CKD = chronic kidney disease; CS = company submission; DFK = difelikefalin; ESRD = end stage renal disease; OLE = open label extension; Pbo = placebo; SD = standard deviation; WI-NRS = Worst Itching Intensity Numerical Rating Scale.

[1] observed stratum values

[2] more than one item may have been checked

Note: Baseline characteristics were recorded during the screening visit for the double-blind treatment phase

Table 3.15: Baseline characteristics CLIN3105 (safety population)

Baseline characteristic	Difelikefalin
Number of participants	**222**
Mean age, years (SD)	**58.1 (12.81)**
Sex – n (%)	
Male	121 (54.5%)
Female	101 (45.5%)
Ethnicity – n (%)	
Hispanic or Latino	48 (21.6%)
Not Hispanic or Latino	173 (77.9%)
Not reported	1 (0.5%)
Race	
American Indian or Alaska Native	2 (0.9%)
Asian	7 (3.2%)
Black or African American	110 (49.5%)
Native Hawaiian or Other Pacific Islander	3 (1.4%)
White	96 (43.2%)
Other	4 (1.89%)
Mean target dry body weight at baseline, kg (SD)	**86.64 (23.548)**
Baseline WI-NRS, mean (SD)	**7.57 (1.331)**
Baseline anti-itch medication use? – n (%)	
Yes	70 (31.5%)

Baseline characteristic	Difelikefalin
No	152 (68.5%)
Mean duration of pruritus, years (SD)	**3.89 (3.312)**
Mean years since diagnosis of ESRD (SD)	**5.87 (4.690)**
Mean years since diagnosis of CKD (SD)	**8.51 (6.878)**
Mean years on chronic haemodialysis (SD)	**5.42 (4.413)**
Aetiology of CKD [1]	
Hypertension	135 (60.8%)
Diabetes	110 (49.5%)
Other	25 (11.3%)
Glomerulonephritis	11 (5.0%)
Large vessel disease	4 (1.8%)
Urologic	3 (1.4%)
Pyelonephritis	2 (0.9%)
Cystic	2 (0.9%)
Unknown	2 (0.9%)
Interstitial nephritis	1 (0.5%)
Nephrotic syndrome	1 (0.5%)
Tumours	1 (0.5%)
Vasculitis	1 (0.5%)

Based on Table 11 of CS[1]
CKD = chronic kidney disease; CS = company submission; ESRD = end stage renal disease; SD = standard deviation; WI-NRS = Worst Itching Intensity Numerical Rating Scale
Notes: Percentages were based on the number of patients in the safety population and noted parenthetically. Vital signs baseline was defined as the last measurement taken on or prior to the first day of dosing
[1] more than one item may have been checked

EAG comment: None.

3.2.3.3 Baseline characteristics of Narita 2022

Baseline characteristics for Narita 2022[13] are given in Table 3.16

Table 3.16: Baseline characteristics of Narita 2022

	Difelikefalin 0.25 mcg/kg (n=61)	Difelikefalin 0.5 mcg/kg (n=61)	Difelikefalin 1.0 mcg/kg (n=61)	Placebo (n=63)
Male	50 (82)	45 (74)	47 (77)	43 (68)
Female	11 (18)	16 (26)	14 (23)	20 (32)
Age, mean (SD)	64.2 (11.2)	65.6 (11.4)	64.4 (11.7)	64.1 (12.7)
Dry weight, mean (SD), kg	61.25 (13.86)	59.98 (11.22)	62.85 (13.39)	60.63 (12.71)
Primary disease caused ESKD (overlapping), No				
Diabetic nephropathy	28	32	32	27

Glomerulonephritis	12	11	13	10
Nephrosclerosis	8	14	7	10
Polycystic kidney	2	0	4	3
Other	5	4	3	8
Unspecified	8	1	3	8
Type of dialysis, No. (%)				
Haemodialysis	23 (38)	27 (44)	30 (49)	24 (38)
Off-line hemodiafiltration	1 (2)	2 (3)	1 (2)	1 (2)
Online hemodiafiltration	29 (48)	25 (41)	22 (36)	30 (48)
Intermittent infusion hemodiafiltration	8 (13)	7 (11)	8 (13)	8 (13)
Concomitant antipruritic agents, No. (%)				
Corticosteroids	24 (39)	27 (44)	23 (38)	21 (33)
Antihistamines	47 (77)	46 (75)	46 (75)	51 (81)
Moisturisers	40 (66)	38 (62)	38 (62)	33 (52)
Others	19 (31)	13 (21)	19 (31)	19 (30)
Other				
Specific signs or symptoms at screening, No. (%)	5 (8)	5 (8)	6 (10)	7 (11)
Weekly NRS score, mean (SD)	6.35 (1.24)	6.83 (1.40)	6.47 (1.29)	6.53 (1.31)
Duration of dialysis, mean (SD), y	7.0 (6.5)	6.7 (7.2)	7.7 (6.5)	6.8 (6.1)
Single-pool Kt/V, mean (SD)	1.435 (0.267)	1.511 (0.309)	1.516 (0.415)	1.498 (0.343)
Urea reduction ratio, mean (SD), %	68.8 (7.2)	70.2 (6.6)	70.6 (8.4)	70.2 (7.6)
Disease duration of itch, mean (SD), y	3.7 (3.5)	4.5 (4.4)	4.8 (4.9)	4.3 (4.4)
Prior treatment with nalfurafine, No. (%)	30 (49)	30 (49)	33 (54)	34 (54)

Based on Narita 2022[13]
ESKD = end stage kidney disease; kg = kilogram; kt/V = clearance of urea multiplied by dialysis duration and normalized for urea distribution volume; microg = micrograms; No. = number; NRS = numerical rating scale; SD = standard deviation; y = years

EAG comment: The between-arm comparability in Narita 2022[13] appears to be adequate.

3.2.3.4 Baseline characteristics of Fishbane 2020

Baseline characteristics for Fishbane 2020[12] are given in Table 3.17.

Table 3.17: Baseline characteristics of Fishbane 2020

	Difelikefalin 0.5 mcg/kg (n=44)	Difelikefalin 1.0 mcg/kg (n=41)	Difelikefalin 1.5 mcg/kg (n=44)	Placebo (n=45)
Male n(%)	**29 (80)**	**26 (84)**	**29 (74)**	**27 (84)**
Age, median (y)	**57**	**59**	**56.5**	**60**

	Difelikefalin 0.5 mcg/kg (n=44)	Difelikefalin 1.0 mcg/kg (n=41)	Difelikefalin 1.5 mcg/kg (n=44)	Placebo (n=45)
Dry weight, mean (SD), kg	**83.5 (20.9)**	**85.4 (25.1)**	**82.8 (20.3)**	**81.0 (19.8)**
Primary disease caused ESKD (overlapping), No				
Diabetes	24	20	19	21
Hypertension and large vessel disease	21	20	24	21
Glomerulonephritis/nephritis	6	4	2	5
Other	2	3	2	1
Interstitial nephritis/pyelonephritis	0	0	0	1
Cystic/hereditary/congenital disease	2	2	1	0
Urologic	0	1	0	0
Unknown	0	0	1	0
Race, No. (%)				
Black/African American	24 (54.5)	22 (53.7)	31 (70.5)	25(55.6)
White	17 (38.6)	19 (46.3)	10 (22.7)	16 (35.6)
Other (Asian, American Indian, Hawaiian or Other Pacific Islander)	3 (6.8)	0	3 (6.8)	3 (6.7)
Not reported	0	0	0	1(2.3)
Concomitant anti-pruritus agents, No. (%)				
Any prior anti-pruritic medication	20 (45.5)	17 (41.5)	18 (40.9)	18 (40)
Diphenhydramine hydrochloride	11 (25)	11 (26.8)	11 (25)	11 (24.4)
Hydroxyzine hydrochloride	6 (13.6)	2 (4.9)	3.0 (6.8)	2 (4.4)
Topical hydrocortisone	1 (2.3)	2 (4.9)	1 (2.3)	5 (11.1)
Other				
Patient-assessed disease severity, category C[a], n (%)	18 (40.9)	14 (34.1)	9 (20.5)	10 (22.2)
Duration of CKD-aP, y (mean [SD])	4.7 (3.9)	4.6 (4.3)	3.9 (3.4)	4.4 (4.7)
Years on chronic haemodialysis (mean [SD])	5.4 (4.9)	6.3 (4.7)	5.5 (4.4)	5.9 (4.9)
Most recent single-pool Kt/V, mean (SD)	1.6 (0.2)	1.6 (0.3)	1.6 (0.3)	1.6 (0.3)
Most recent urea reduction ratio, mean (SD), %	72 (3.9)	78.2 (6.6)	71.6 (3)	73.3 (4.6)

Based on Fishbane 2020[12]
CKD = chronic kidney disease; CKD-aP = CKD-associated pruritis; ESRD = end stage renal disease; Kt/V = clearance of urea multiplied by dialysis duration and normalised for urea distribution volume; No. = number; SD = standard deviation; y = years

	Difelikefalin 0.5 mcg/kg (n=44)	Difelikefalin 1.0 mcg/kg (n=41)	Difelikefalin 1.5 mcg/kg (n=44)	Placebo (n=45)
[a]Disease severity category C: '*I often have scratch marks on my skin that may or may not bleed or get infected; I often have a problem sleeping because of itching; my itching often makes me feel agitated or sad*.'				

EAG comment: The between-arm comparability in Fishbane 2020[12] appears to be adequate.

3.2.4 Risk of bias assessment

3.2.4.1 KALM RCTs

Tables 3.18 and 3.19 provide the risk of bias assessment for KALM-1 and KALM-2 provided by the company, indicating no serious risk of bias for each study.

Table 3.18: Quality assessment of KALM-1

Question	Response
Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes
Were there any unexpected imbalances in dropouts between groups?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes
Based on Table 15 of CS[1] CS = company submission; ITT = intention-to-treat	

Table 3.19: Quality assessment of KALM-2

Question	Response
Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes
Were there any unexpected imbalances in dropouts between groups?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes
Based on Table 16 of CS[1] CS = company submission; ITT = intention-to-treat	

EAG comment:

- The EAG carried out its own evaluation of each RCTs risk of bias. For both studies, the existence of allocation concealment was not clear. An interactive web response system was described, which

implies that allocation concealment was carried out, but does not confirm it. Attrition bias was low risk for both studies. This was because an ITT analysis was used for all outcomes. In addition, for data lost to follow up the arm differential was small, and the missing rate was less than the event rate for the primary outcome. Blinding appears to have been adequately carried out for participants, clinicians and assessors in both studies, so performance and detection bias also appear to be low risk. Overall, both studies were designated by the EAG to be at serious risk of bias due to the uncertainty resulting from the poor reporting of allocation concealment.

3.2.4.2 OLE and one arm studies

Table 3.20, Table 3.21 and

Table 3.22 provide the risk of bias assessment for KALM-1 OLE, KALM-2 OLE and CLIN3105 provided by the company, indicating no serious risk of bias for each study, apart from a lack of complete follow up for KALM-2 OLE.

Table 3.20: Quality assessment of KALM-1 OLE

Question	Response
Was the cohort recruited in an acceptable way?	Yes
Was the exposure accurately measured to minimise bias?	Yes
Was the outcome accurately measured to minimise bias?	Yes
Have the authors identified all important confounding factors?	Yes
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes
Was the follow-up of patients complete?	Yes
How precise (for example, in terms of confidence interval and p values) are the results?	Yes*

Based on Table 17 and Table 70 of CS[1]
*Further information provided in Table 70 of the CS: "*For each time point, standard descriptive statistics were used to report observed scores and the changes from baseline, along with the LS means, standard errors, 95% CIs, and differences from baseline within each treatment sequence (reported with LS means, standard errors, and 95% CIs). See results section for KALM-1 OLE.*"[1]
CS = company submission; OLE = open label extension

Table 3.21: Quality assessment of KALM-2 OLE

Question	Response
Was the cohort recruited in an acceptable way?	Yes
Was the exposure accurately measured to minimise bias?	Yes
Was the outcome accurately measured to minimise bias?	Yes
Have the authors identified all important confounding factors?	Yes
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes
Was the follow-up of patients complete?	No
How precise (for example, in terms of confidence interval and p values) are the results?	Yes*

Based on Table 18 and Table 72 of CS[1]
*Further information provided in Table 72 of the CS: "*For each time point, standard descriptive statistics were used to report observed scores and the changes from baseline, along with the LS means, standard errors, 95% CIs, and differences from baseline within each treatment sequence.*"[1]
CS = company submission; OLE = open label extension

Table 3.22: Quality assessment of CLIN3105

Question	Response
Was the cohort recruited in an acceptable way?	Yes
Was the exposure accurately measured to minimise bias?	Yes
Was the outcome accurately measured to minimise bias?	Yes
Have the authors identified all important confounding factors?	Yes
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes
Was the follow-up of patients complete?	Yes
How precise (for example, in terms of confidence interval and p values) are the results?	Yes*

Based on Table 19 and Table 73 of CS[1]
*Further information provided in Table 73 of the CS: "*For each time point, standard descriptive statistics were used to report observed scores and the changes from baseline, along with the LS means, standard errors, 95% CIs, and differences from baseline within each treatment sequence.*"[1]
CS = company submission

EAG comment: The company's conclusion that the one-arm study designs were at low risk of bias ignores the problems common to all one-arm trials: that the lack of a comparator means that it is impossible to extricate intervening effects from true treatment effects. Such trial designs should, by default, be regarded as at serious risk of bias.

3.2.4.3 Narita 2022

The risk of bias was evaluated by the EAG to be serious. This was because allocation concealment was not adequately described. However, blinding was well-reported, and attrition would be unlikely to be a source of bias for the primary analyses.

3.2.4.4 Fishbane 2020

The risk of bias was evaluated by the EAG to be serious. Allocation concealment was only partially described, but blinding was well-reported, and loss of data was too low and comparable between groups to create any significant attrition bias.

3.2.5 Efficacy results of the included studies

3.2.5.1 Itching intensity

3.2.5.1.1 Proportion of patients achieving a ≥3-point improvement in WI-NRS from baseline at week 12

KALM-1

Table 3.23 summarises results for the ITT population, based on the combined data from interim and post-interim analysis patients. At week 12, the least squares (LS) mean percentage of patients with at least a 3-point improvement from baseline in the WI-NRS was 51.0% in the difelikefalin group, compared with 27.6% in the placebo group. The odds ratio (OR) for a ≥3-point improvement from baseline with difelikefalin versus placebo was 2.72 (95% confidence interval (CI), 1.72 to 4.30; p<0.001).

Table 3.23: Primary analysis: patients with a ≥3-point improvement from baseline at week 12 with respect to the WI-NRS score – MI with MAR assumption (population: ITT)

Combined estimates at week 12	Placebo (N=189)	DFK (N=189)
Observed ≥3-point NRS improvement [1] - n (%)		
Yes	51 (30.9%)	82 (52.2%)
No	114 (69.1%)	75 (47.8%)
Missing	24	32
LS means estimate of percent with improvement [2]		
Percent (95% CI)	27.6% (20.2%, 36.6%)	51.0% (42.9%, 58.9%)
LH odds ratio (95% CI)	-	2.72 (1.72, 4.30)
CHW p-value	-	<.001

Based on Table 20 in CS[1]
CHW = Cui, Hung, Wang; CI = confidence interval; CS = company submission; ITT = intention-to-treat; LH = Lawrence, Hung; LS = least squares; MAR = missing at random; MI = multiple imputation; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[1] counts and percentages were based on non-missing data
[2] estimated percent, odds ratio, and p-value used a logistic regression model with terms for treatment group, baseline WI-NRS score, use of anti-itch medication during the week prior to randomisation, and the presence

Combined estimates at week 12	Placebo (N=189)	DFK (N=189)
of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim patients and post-interim patients separately Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the LH/CHW methodology		

Sensitivity analysis on the ITT data was also performed without using the CHW adjustment procedure and without splitting of the data with respect to interim versus post-interim status. This analysis also showed results favouring difelikefalin at week 12, with an odds ratio (difelikefalin versus placebo) for a ≥3-point improvement from baseline in WI-NRS score of 2.62 (95% CI, 1.68 to 4.09). Three further sensitivity analyses were conducted, one treating patients who discontinued study drug early as non-responders, one using MI based on a missing not at random (MNAR) assumption, and a tipping point analysis which used MI with MNAR for treated patients and missing-at-random (MAR) for placebo patients.[5]. The results of the sensitivity analysis are presented in Table 3.25.

Table 3.24: Supportive and sensitivity analyses – percentage of patients with a ≥3-point improvement in WI-NRS at week 12 (population: ITT and per protocol)

Analysis statistic	Placebo	DFK
Sensitivity analyses		
Patients who discontinued early as non-responders[a]	-	-
N	189	189
LS mean percent with improvement (95% CI)	26.0% (19.0%, 34.5%)	44.6% (35.4%, 54.2%)
LH odds ratio (95% CI)	-	2.29 (1.46, 3.60)
CHW p-value	-	<.001
MI with MNAR assumption[a]	-	-
N	189	189
LS mean % with improvement (95% CI)	27.6% (20.2%, 36.4%)	44.6% (35.4%, 54.2%)
LH odds ratio (95% CI)	-	2.33 (1.47, 3.71)
CHW p-value	-	<.001
Tipping point [a]	-	-
N	189	189
Highest shift parameter without tipping	6.50	6.50
Percent with improvement (95% CI)	29.1% (21.5%, 38.1%)	42.8% (33.7%, 52.4%)
LH odds ratio	-	1.82 (1.16, 2.86)
CHW p-value	-	.009
Additional analysis		
Per protocol population[a]		
N	169	163

Analysis statistic	Placebo	DFK
LS mean percent with improvement (95% CI)	27.0% (19.1%, 36.6%)	50.4% (47.1%, 53.6%)
LH odds ratio (95% CI)		2.74 (1.71, 4.41)
CHW p-value		<.001
No CHW adjustment for interim analysis[a]		
N	189	189
LS mean percent with improvement (95% CI)	28.3% (21.0%, 37.1%)	50.9% (41.6%, 60.2%)
LH odds ratio (95% CI)		2.62 (1.68, 4.09)
P-value		<.001

Based on Table 21 of CS Document B[29] and the company's response to clarification[5]
CHW = Cui, Hung, Wang; CI = confidence interval; CS = company submission; ITT = intention-to-treat; LH = Lawrence, Hung; LS = least squares; MI = multiple imputation; MNAR = missing-not-at-random; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[a]Analysis based on interim and post-interim patients combined

According to the company conclusions on all sensitivity analyses confirmed the MI analysis results. The odds ratio results remained statistically significant for all three sensitivity analyses.

KALM-2

Table 3.25 summarises results for the ITT population, based on the combined data from interim and post-interim analysis patients. At week 12, the LS mean percentage of patients with at least a 3-point improvement from baseline in the WI-NRS was 54.0% in the difelikefalin group, compared with 42.2% in the placebo group. The estimated OR for a ≥3-point improvement from baseline with difelikefalin versus placebo was in favour of difelikefalin 1.61 (95% CI, 1.08 to 2.41).

Table 3.25: Analysis: Patients with ≥3-point improvement from baseline at week 12 with respect to the WI-NRS score – multiple imputations with MAR assumption (population: ITT)

Combined estimates (week 12)	Placebo (n=236)	DFK (n=237)
Observed ≥3-point NRS improvement [1] - n (%)		
Yes	77 (33.2%)	95 (49.7%)
No	130 (62.8%)	96 (50.3%)
Missing	29	46
LS means estimate of percent with improvement [2]		
Percent (95% CI)	42.2% (32.5%, 52.5%)	54.0% (43.9%, 63.9%)
LH odds ratio (95% CI)		1.61 (1.08, 2.41)
CHW p-value		0.020

Combined estimates (week 12)	Placebo (n=236)	DFK (n=237)
Based on Table 26 in CS[1] CHW = Cui, Hung, Wang; CI = confidence interval; CS = company submission; ITT = intention-to-treat; LH = Lawrence, Hung; LS = least squares; MAR = missing-at-random; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale [1] counts and percentages were based on non-missing data [2] estimated percentage, odds ratio and p-value used a logistic regression model with terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim patients and post-interim patients separately Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate		

Sensitivity analysis on the ITT data was also performed without using the CHW adjustment procedure and without splitting of the data with respect to interim versus post-interim status. This analysis also suggested results favouring difelikefalin at week 12, with an OR (difelikefalin versus placebo) for a ≥3-point improvement from baseline in WI-NRS score of 1.54 (95% CI, 1.05 to 2.27). Similarly to KALM-1, three sensitivity analyses were run and the results are presented in Table 3.26.

Table 3.26: Analysis: Patients with ≥3-point improvement from baseline at week 12 with respect to the WI-NRS score – multiple imputations with MAR assumption (population: ITT)

Analysis statistic	Placebo	DFK
Sensitivity analyses		
Patients who discontinued early as non-responders[a]		
N	236	237
LS mean percent with improvement (95% CI)	37.2% (27.8%, 47.6%)	47.7% (33.4%, 54.7%)
LH odds ratio (95% CI)	-	1.31 (0.89, 1.94)
CHW p-value	-	0.168
MI with MNAR assumption [a]		
N	236	237
LS mean percent with improvement (95% CI)	39.9% (30.6%, 50.1%)	50.7% (41.2, 60.1%)
LH odds ratio (95% CI)	-	1.55 (1.05, 2.28)
CHW p-value	-	0.029
Tipping point [a]		
N	236	237
Highest shift parameter without tipping	0.75	0.75
% with improvement (95% CI)	41.9% (32.0%, 52.4%)	52.1% (42.5%, 61.5%)
LH odds ratio	-	1.51 (1.01, 2.35)
CHW p-value	-	0.044

Analysis statistic	Placebo	DFK
Additional analysis		
Per protocol population[a]		
N	213	205
LS mean percent with improvement (95% CI)	39.7 (29.7%, 50.7%)	52.0% (43.8%, 60.2%)
LH odds ratio (95% CI)	-	1.65 (1.08, 2.51)
CHW p-value	-	0.019
No CHW adjustment for interim analysis[a]		
N	236	237
LS mean percent with improvement (95% CI)	42.6% (33.4%, 52.3%)	53.4% (43.7%, 62.8%)
LH odds ratio (95% CI)	-	1.54 (1.05, 2.27)
P-value	-	0.027

Based on Table 27 of the revised CS Document B[29] and the company's response to clarification[5]
CHW = Cui, Hung, Wang; CI = confidence interval; CS = company submission; DFK = difelikefalin; ITT = intention-to-treat; LH = Lawrence, Hung; LS = least squares; MI = multiple imputation; MAR = missing-at-random; MNAR = missing-not-at-random; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[a]Analysis based on interim and post-interim patients combined

As part of their response to the clarification questions, the company stated that "*In conclusion, the several sensitivity analyses confirm the conclusions of the primary analyses, which were conducted according to the intent-to-treat principle*."[5]. Nevertheless, it should be noted that the results of the sensitivity analysis which treated patients who discontinued as non-responders did not have statistically significant results, contrary to the results of the same sensitivity analysis conducted for KALM-1. The company did not offer any discussion on the matter.

Pooling of KALM-1 and KALM-2 main data

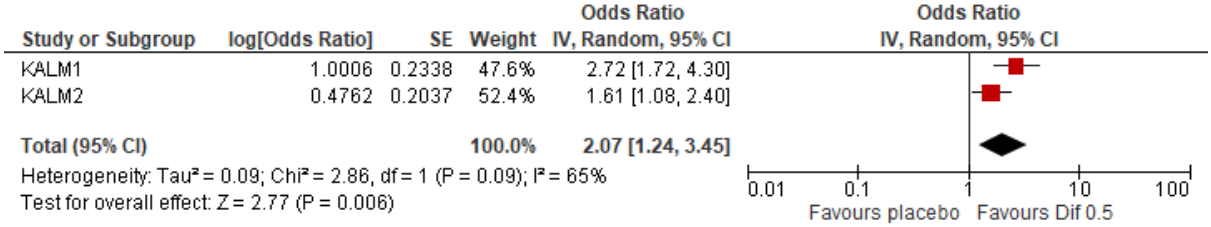
In the pooled analysis, the odds of achieving a ≥3-point reduction in WI-NRS score at week 12 were almost twice as large with difelikefalin versus placebo (OR [95% CI]: 1.93 [1.44, 2.57]). This result was not reported fully in the CS, and has been derived from Topf 2022.[26]

EAG comment:

- In Topf 2022[26], it is stated that *"Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants".* This implies that a meta-analysis was not used for pooling, and that the individual patient data from the two studies were simply added together. This approach was probably inappropriate because *"such an analysis is against the main principle of meta-analysis…a pooled analysis like this will yield over-precise results (CIs too narrow) and it may well lead to bias if any of the trials has unequal numbers in the two arms"*.[30]. The EAG highlighted this as a key issue. The EAG has carried out an inverse variance meta-analysis of the results from KALM-1 and KALM-2 on Review Manager 5.3. Heterogeneity was serious (I^2=65%) and so a random effects model was used.

- The pooled OR (95% CI) was 2.07 (1.24 to 3.45) (Figure 3.4). This may be a more valid pooled result than that submitted by Topf 2022.[26]

Figure 3.4: Meta-analysis of KALM-1 and KALM-2



CI = confidence interval; IV = intravenous; SE = standard error

3.2.5.1.2 Proportion of patients achieving a ≥4-Point improvement in weekly 24-hour WI-NRS at 12 weeks

KALM-1

Table 3.27 summarises the analysis of the proportion of patients achieving a ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour WI-NRS at week 12 of the double-blind treatment period. At week 12, the LS mean percentage of patients with a ≥4-point improvement in WI-NRS from baseline was 38.9% in the difelikefalin group and 18.0% in the placebo group; the odds ratio with difelikefalin was 2.89 (95% CI, 1.75 to 4.76).

Table 3.27*:* Patients with a ≥4-point improvement from baseline at week 12 in WI-NRS score – MI with MAR assumption (population: ITT)

Combined assessments (week 12)	Placebo (n=189)	DFK (n=189)
Observed ≥4-point NRS improvement [1] – n (%)		
Yes	35 (21.2%)	64 (40.8%)
No	130 (78.8%)	93 (59.2%)
Missing	24	32
LS means estimate of percent with improvement [2]		
Percent (95% CI)	18.0% (21.1%, 26.0%)	38.9% (29.8%, 48.7%)
LH odds ratio (95% CI)		2.89 (1.75, 4.76)
CHW p-value		<.001

Based on Table 24, CS[1]
CHW = Cui, Hung, Wang; CI = confidence interval; CS = company submission; DFK = difelikefalin; ITT = intention-to-treat; LH = Lawrence, Hung; LS = least squares; MI = multiple imputation; MAR = missing-at-random; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[1] counts and percentages were based on non-missing data
[2] estimated percent, odds ratio, and p-value use a logistic regression model with terms for treatment group, baseline WI-NRS score, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim patients and post-interim patients separately

KALM-2

Table 3.28 summarises the analysis of this endpoint. At week 12, the LS mean percentage of patients with a ≥4-point improvement in WI-NRS from baseline was 41.2% in the difelikefalin group and 28.4% in the placebo group; the OR was 1.77 (95% CI, 1.14 to 2.74).

Table 3.28: Patients with a ≥4-point improvement from baseline at week 12 in WI-NRS Score – MI with MAR assumption (population: ITT)

Combined estimates (week 12)	Placebo (n=236)	DFK (n=237)
Observed ≥4-point NRS improvement [1] - n(%)		
Yes	52 (25.1%)	72 (37.7%)
No	155 (74.9%)	119 (62.3%)
Missing	29	46
LS means estimate of percent with improvement [2]		
Percent (95% CI)	28.4% (21.3%, 37.7%)	41.2% (33.0%, 50.0%)
LH odds ratio (95% CI)	-	1.77 (1.14, 2.74)
CHW p-value	-	0.010

Based on Table 29, CS[1]
CHW = Cui, Hung, Wang; CI = confidence interval; CS = company submission; DFK = difelikefalin; ITT = intention-to-treat; LH = Lawrence, Hung; LS = least squares; MI = multiple imputation; MAR = missing-at-random; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale
[1] counts and percentages were based on non-missing data
[2] estimated percent, odds ratio, and p-value use a logistic regression model with terms for treatment group, baseline W-NRS score, region, use of anti-itch medication during the week prior to randomisation, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim patients and post-interim patients separately
Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and p-value using the LH/CHW methodology

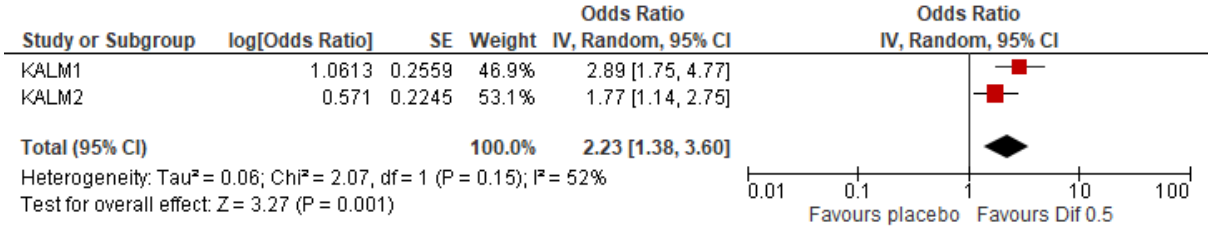
Pooling of KALM-1 and KALM-2 main data

In the pooled population, achievement of ≥4-point reduction in weekly mean of daily WI-NRS score was significantly greater with difelikefalin versus placebo at all time points from week 3 to week 12 (week 12 least-squares (LS) mean estimate [95% CI]: 38.7% [32.8%, 45.0%] versus 23.4% [18.7%, 28.8%], P<0.001)). Unfortunately, no between-arm results were provided. This result was not reported in the CS, and has been derived from Topf 2022.[26]

EAG comment:

- In Topf 2022[26], it is stated that **"***Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants".* This implies that a meta-analysis was not used for pooling, and that the individual patient data from the two studies were simply added together. This approach was probably inappropriate because "*such an analysis is against the main principle of meta-analysis…a pooled analysis like this will yield over-precise results (CIs too narrow) and it may well lead to bias if any of the trials has unequal numbers in the two arms*".[30] The EAG has therefore carried out an inverse variance meta-analysis of the results from KALM-1 and KALM-2 on Review Manager 5.3. Heterogeneity was serious (I^2=52%) and so a random effects model was used.
- The pooled OR (95% CI) was 2.23 (1.38 to 3.60) (Figure 3.5). This provides the informative between-arm effect that was lacking in Topf 2022.[26]

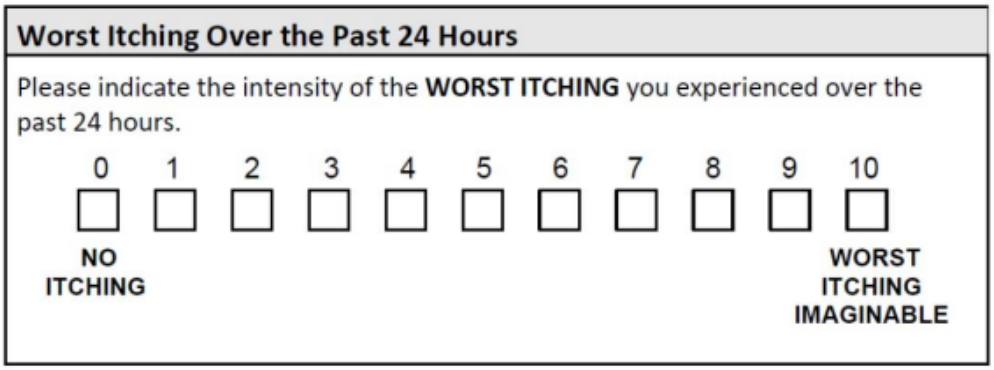
Figure 3.5: Meta-analysis of KALM-1 and KALM-2



CI = confidence interval; Dif = difelikefalin; SE = standard error; IV = intravenous

- In the effectiveness conclusions of the CS[1] the company states that "*At Week 12, a majority of the subjects reported at least a 3-point (73.7%) or 4-point (59.3%) improvement from baseline in the weekly mean of the daily 24-hour WI-NRS score, which was previously established as a clinically meaningful threshold for this patient population.*[31]" The company was asked to report how these data were calculated, as the referenced abstract does not contain the reported results. The company responded by stating that, "*At Week 12, a majority of the subjects reported at least a 3-point (73.7%) or 4-point (59.3%) improvement from baseline in the weekly mean of the daily 24-hour WI-NRS score' is referring to the results taken from the KALM trial. The reference to a 3 (or more) point improvement being a clinically meaningful threshold is taken from Vernon et al., 2021, page 1133:*[32] *'These analyses demonstrated that a reduction of ≤3 points on the WI-NRS marks an appropriate threshold for defining a clinically meaningful change in pruritus in patients with CKD-aP.'*"[5] The EAG thanks the company for this response.
- It is unclear if the patient-reported, single item WI-NRS was used to determine the severity of CKD-aP across all relevant trials. The company was asked to clarify if this was so or to explain the measure used to determine the severity of CKD-aP. The company stated that, "*All trials measured itch severity using both the WI-NRS and 5-D itch scores*."[5] The EAG appreciates the clarification.
- The company was also asked to provide information on the methodology of the WI-NRS score. The company stated that, "*The Worst Itching Intensity Numerical Rating Scale (WI-NRS) is a simple, single-item patient-reported outcome measure to assess the intensity of the worst itching a patient has experienced over the past 24 hours, as described in 'Clinically Meaningful Change in Itch Intensity Scores: an Evaluation in Patients with Chronic Kidney Disease associated Pruritus' (Figure 3).'*"[5]

Figure 3.6: Worst Itching Intensity Numerical Rating Scale (WI-NRS)



Based on Company response to clarification document[5]
WI-NRS = Worst Itching Intensity Numerical Rating Scale

- The company was also asked to discuss the validity of this scale in capturing the severity of CKD-aP. The company explained that, "*Various pieces of literature*[31, 33, 34] *have found WI-NRS to be a*

reliable, reproducible, and valid measure of itch intensity in moderate-to-severe CKD-aP patients, and therefore a reasonable choice….".[5] The EAG checked the references and found that the data presented in Vernon 2021[31] support the use of the WI-NRS scale, although it should be noted that the main author has a conflict of interest. Mathur 2010[33] and Storck 2021[34] do not directly mention the WI-NRS. However, it is assumed that the "*numerical rating scale*" that Storck 2021[34] refers to, is equivalent to the WI-NRS. If so, Storck 2021[34] supports the company's statement. If the 'Brief Itching Inventory' referred to by Mathur 2010[33] is synonymous with the WI-NRS, then Mathur 2010[33] also supports the company's statement. However, the EAG considers that there is a lack of clarity.

- The company was asked to justify the choice of >4 points weekly mean, as a benchmark for moderate-to-severe pruritus. The company stated that, "*The paper 'A Randomized Controlled Trial of Difelikefalin for Chronic Pruritus in Haemodialysis Patients' states 'Itching severity scores collected via the WI-NRS have been categorized in the literature as mild (<4), moderate (≥4 to <7), or severe (≥7).[12, 35]*"[5] The EAG checked the references and the primary reference Reich 2012[35] confirms the company's statement.
- The SAP (page 28) of KALM-1 and KALM-2 states that, "*In the primary efficacy analysis, missing NRS data at the end of Week 12 will be imputed using a multiple imputation (MI) approach, assuming that subjects who discontinue double-blind treatment early would have similar Worst Itching Intensity NRS scores as other subjects in their respective treatment arm who have complete data*".[36] The company was asked to provide a formal presentation of the MI analysis. In their response they provided further clarification stating that "*The weekly mean WI-NRS score was computed as the average of daily e-diary entries (scored on an integer scale from 0 to 10). If there were more than 3 missed e-diary entries in a week, the weekly mean WI-NRS score was set to missing*.", while "*The reason for a missed e-diary entry was not collected.*"[5]
- Regarding the pattern of missing data, according to the company most of them were recorded on the 12th week but occurred in other weeks as well, as shown in Table 3.29. Nevertheless, as shown in Table 3.29 in KALM-1, 32 (16.9%) patients in the treatment arm and 23 (12.2%) patients in the placebo arm had more than one weekly mean WI-NRS score missing; in KALM-2, 44 (18.6%) patients in the treatment arm and 12 (10.6%) patients in the placebo arm had more than one weekly mean WI-NRS score missing. In addition, a noteworthy proportion of the patients (6.8% to 9.5% in both arms of both studies) had missing data other that week 12 and continued with the trial.

Table 3.29: Weekly missing mean WI-NRS scores in KALM-1 and KALM-2.

		Weekly mean WI-NRS scores (n, %)					
	N	**Complete**	**Missing any**	**Missing one only**	**Missing week 12 only**	**Missing week 12 and others**	**Missing weeks other than 12th**
KALM-1							
Difelikefalin	189	141 (74.6%)	48 (25.4%)	16 (8.5%)	4 (2.1%)	28 (14.8%)	16 (8.5%)
Placebo	189	147 (77.8%)	42 (22.2%)	19 (10.1%)	7 (3.7%)	17 (8.9%)	18 (9.5%)
KALM-2							
Difelikefalin	237	175 (73.8%)	62 (26.2%)	18 (7.6%)	11 (4.6%)	35 (14.8%)	16 (6.8%)
Placebo	236	191 (80.9%)	45 (19.1%)	20 (8.5%)	10 (4.2%)	19 (8.1%)	16 (6.8%)

Based on the company's response to clarification[5], including Table 3c-1 and Table 3c-2 of Appendix V[24]
WI-NRS = Worst Itching Intensity Numerical Rating Scale

- The reasons for early discontinuation are provided in two tables both named 'Table: 14.1.1.1' in two files included in Appendix V.[24] These tables do not report the reasons for intermediate missing data. The company stated that "*The primary efficacy analysis used imputed data based on a missing at random (MAR) assumption, that is, that subjects who discontinued double-blind treatment early would have similar WI-NRS scores as subjects in the same treatment group who had complete data*" (p. 32).[5] The EAG would like to point out that the reason of missing data, whether recorded or hypothesised, provides the conceptual basis of determining if the data are missing at random, completely at random or not at random, which in turn determines whether MI is appropriate or not. Table 14.1.1.1 for KALM-1 reports as reasons for discontinuation: adverse effects, eligibility criteria, non-compliance, withdrawn consent and other; while Table 14.1.1.1 for KALM-2 reports as reasons for discontinuation: AEs, lack of therapeutic efficacy, lost to follow-up, eligibility criteria, non-compliance, withdrawn consent and other. The EAG considers that the company has not provided adequate commentary on how the reasons for discontinuation might affect the MI analysis.
- When presenting an MI analysis, it is considered good practice to present the results of analysis restricted to observed cases along with the MI results so that a comparison can be made. In the response to clarification the company reported the observed data of patients with no missing data (Table 3.30) without the corresponding odds ratio estimates.[5] Nevertheless, the EAG calculated the odds ratios and found that they are comparable to the MI results.

Table 3.30: Patients with ≥3-point improvement from baseline at week 12 with respect to the WI-NRS score; counts and percentages based on non-missing data alone

Study		Placebo	Difelikefalin
KALM-1	N	165	157
	≥3-point improvement	51 (30.9%)	82 (52.2%)
KALM-2	N	207	191
	≥3-point improvement	77 (37.2%)	95 (49.7%)

Based on the company's response to clarification question A27.[5]
WI-NRS = Worst Itching Intensity Numerical Rating Scale

- In the table presenting the sensitivity analysis for KALM-1 in the company's response to clarification (page 34)[5] the OR for the primary MI is reported to be 3.31 (95% CI; 1.67, 6.57) which is different to the OR presented in the CS (2.72 [95% CI; 1.72, 4.30]).
- The EAG requested that the company formally report the design and results of the logistic regression executed for the primary endpoint as key aspects were missing from the CS. The company has only in part provided the requested information. They state that all the covariates were determined a priori in the protocol and SAP and no covariate selection method was employed.[5] The EAG understands that a selection process (e.g. stepwise, forward-backward elimination) is used as standard process in logistic regression analysis to examine the effects of the covariates on the fit of the model. In addition, as standard practice, the company was asked to:
    - present the results of the logistic regression in a table of statistics showing model output (regression coefficient [β], standard error, p-value and associated statistics such as z-score) for the intercept and each predictor variable, with definition of the reference value for each predictor variable
    - odds ratio estimates for each predictor variable
    - and interpretation of the relationship between each predictor variable and the response variable assuming other variables held constant (e.g., quantity of increase/decrease of estimated value for a 1-unit increase in the predictor variable)

The above requested data and discussion were not provided. Instead, the company stated that "*The objective of the primary analysis was to estimate the treatment effect, rather than to develop a predictive model. To that end, the regression coefficients from the logistic regression on each imputed dataset were not combined*."[5] The EAG would like to point out that the relationship between covariates (predictor variables) and the response variable in the logistic model is by definition a predictive one. It is also not clear what the company means by stating that the β on each imputed dataset were not combined. The company was also asked to present the logistic analysis results adjusted (for all relevant predictors) and unadjusted overall OR estimates, to which they referred us to their response to another question in the request for clarification, "*Please see the tables 'Subjects with ≥3-point improvement from baseline to Week 12 in WI-NRS, ITT population (CR845-CLIN3102)' and 'Subjects with ≥3-point improvement from baseline to Week 12 in WI-NRS, ITT population (CR845-CLIN3103)' in the response to A27 for this information.*"[5]. Unfortunately, the requested information is not reported there.

3.2.5.1.3. Change from baseline in the weekly mean NRS score at 8 weeks

NARITA 2022

The mean NRS change from baseline scores for each of the three difelikefalin and the placebo group are given in Table 3.31. Adjusted mean differences (difelikefalin – placebo) were given for the two significant results as follows:

- 0.5 µg/kg of difelikefalin (adjusted mean difference −0.80; 95% CI, −1.55 to −0.04)
- 1. 0 µg/kg of difelikefalin (adjusted mean difference −0.78; 95% CI, −1.54 to −0.03)

Table 3.31: Mean NRS score (adjusted mean [SE] change) at week 8

	Placebo	0.25 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
Mean NRS score (adjusted mean [SE] change) at week 8	−2.86 (0.29)	−2.97 (0.29)	−3.65 (0.30)*	−3.64 (0.30)*

Based on Narita 2022[13]
NRS = numerical rating scale; SE = standard error
*Significant difference (p=0.04) compared to placebo

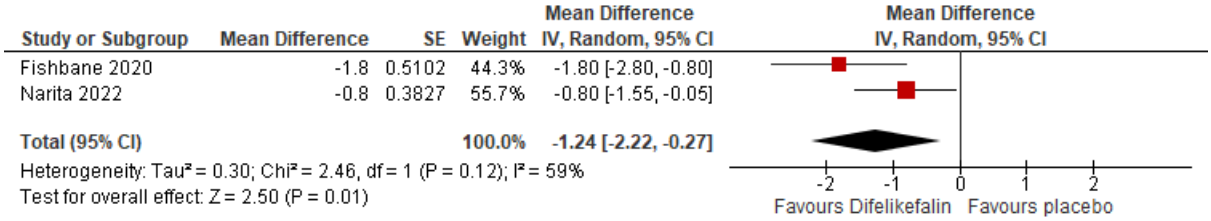
FISHBANE 2020

Adjusted mean differences (difelikefalin – placebo) for the mean NRS change from baseline scores were given for the as follows:

- 0.5 µg/kg of difelikefalin (adjusted mean difference −1.80; 95% CI, −2.8 to −0.8)
- 1. 0 µg/kg of difelikefalin (adjusted mean difference −0.80; 95% CI, −1.9 to 0.2).
- 1. 5 µg/kg of difelikefalin (adjusted mean difference −1.2; 95% CI, −2.3 to −0.2).

EAG comment: A meta-analysis was carried out by the EAG to synthesise the results of Narita 2022[13] and Fishbane 2020.[12] Heterogeneity was serious (I^2=59%) and so a random-effects method was used. The pooled result (Figure 3.7) suggested a more favourable outcome for difelikefalin 50 mcg/kg over placebo.

Figure 3.7: Meta-analysis of Fishbane 2020 and Narita 2022



CI = confidence interval; SE = standard error; IV = intravenous

3.2.5.1.4. Proportions of participants with ≥3-point improvement in the NRS score at week 8

NARITA 2022

Table 3.32 provides data on the proportions of an at least 3-point improvement in the NRS score at week 8 in each of the four groups.

- The RR for 0.25 µg/kg of difelikefalin versus placebo was calculated to be 1.05 (0.65 to 1.70)
- The RR for 0.5 µg/kg of difelikefalin versus placebo was calculated to be 1.20 (0.74 to 1.97)
- The RR for 1.0 µg/kg of difelikefalin versus placebo was calculated to be 1.13 (0.69 to 1.86)

Table 3.32: Proportions of an at least 3-point improvement in the NRS score at week 8

	Placebo	0.25 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
No. (%) with three points or higher improvement in NRS	29/58 (50%)	31/59 (53%)	32/53 (60%)	30/53 (57%)
Based on Narita 2022[13] NRS = numerical rating scale				

EAG comment:

- The data from Narita 2022[13] have not been pooled with the data from the KALM RCTs because of the relatively large difference in follow-up time (8 weeks and 12 weeks respectively) that might have an impact on outcome.

3.2.5.1.5. Proportions of participants with ≥4-point improvement in the NRS score at week 8

NARITA 2022

Table 3.32 provides data on the proportions of an at least 4-point improvement in the NRS score at week 8 in each of the four groups.

- The RR for 0.25 µg/kg of difelikefalin versus placebo was calculated to be 0.94 (0.52 to 1.68)
- The RR for 0.5 µg/kg of difelikefalin versus placebo was calculated to be 1.40 (0.80 to 2.46)
- The RR for 1.0 µg/kg of difelikefalin versus placebo was calculated to be 1.20 (0.67 to 2.14)

Table 3.33: Proportions of an at least 4-point improvement in the NRS score at week 8

	Placebo	0.25 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
No. (%) with four points or higher improvement in NRS	21/58 (36%)	20/59 (34%)	27/53 (51%)	23/53 (43%)

Based on Narita (2022) [13]
NRS = numerical rating scale

EAG comment: The data from Narita 2022[13] have not been pooled with the data from the KALM RCTs because of the relatively large difference in follow-up time (8 weeks and 12 weeks respectively) that might have an impact on outcome.

3.2.5.1.6 The Patient Global Impression of Change (PGIC) at 12 weeks

The Patient Global Impression of Change (PGIC) is a global patient-reported outcome (PRO) measure which assesses the overall change in itch (no change, improvement, or worsening) relative to the start of the study.[37] The scale has only one item: each patient was asked to mark the category that best described the change in itch, ranging from "Very Much Improved" to "Very Much Worse".

KALM-1

[redacted]

KALM-2

[redacted]

Pooling of KALM-1 and KALM-2 main data

The proportion of participants who achieved complete response on the WI-NRS was significantly greater with difelikefalin versus placebo at week 12 (12.0% versus 6.7%, P=0.006) with significant differences between difelikefalin and placebo starting at week 3 and sustained at all time points up to week 12). Unfortunately, no between-arm results were provided. This result was not reported in the CS and has been derived from Topf 2022.[26]

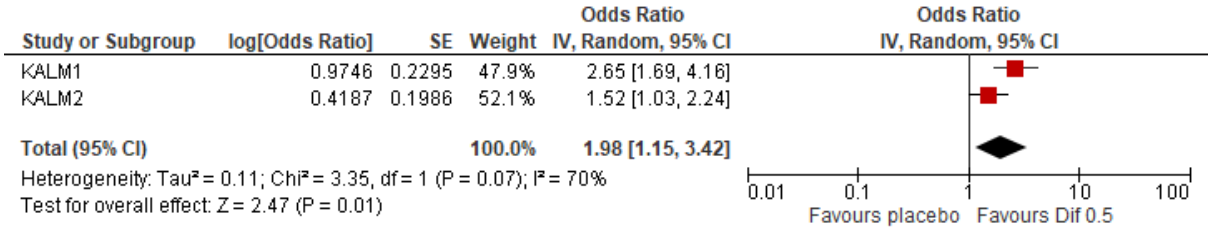
EAG comment:

- In Topf 2022[26], it is stated that **"***Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants".* This implies that a meta-analysis was not used for pooling, and that the individual patient data from the two studies were simply added together. This approach was probably inappropriate because "*such an analysis is against the main principle of meta-analysis…a pooled analysis like this will yield over-precise results (CIs too narrow) and it may well lead to bias if any of the trials has*

unequal numbers in the two arms".[30] The EAG has therefore carried out an inverse variance meta-analysis of the results from KALM-1 and KALM-2 on Review Manager 5.3. Heterogeneity was serious (I^2=70%) and so a random effects model was used.

- The pooled OR (95% CI) was 1.98 (1.15 to 3.42) (Figure 3.8). This provides an informative between-arm effect that was lacking in Topf 2022.[26]

Figure 3.8: Meta-analysis of KALM-1 and KALM-2



CI = confidence interval; Dif = difelikefalin; SE = standard error; IV = intravenous

3.2.5.1.7 The Patient Global Impression of Change (PGIC) at 8 weeks

NARITA 2022

The PGIC outcomes in the difelikefalin 0.5 µg/k and 1.0 µg/kg groups were reported by the trial authors as showing an improvement over the placebo group (Table 3.34). However, only descriptive statistics were presented.

Table 3.34: The Patient Global Impression of Change (PGIC) at 8 weeks

No. (%)	Placebo (n=61)	0.25 µg/kg of difelikefalin (n=61)	0.5 µg/kg of difelikefalin (n=59)	1.0µg/kg of difelikefalin (n=60)
Very much improved	11 (17.7)	12 (19.7)	22 (37.3)	19 (31.7)
Much improved	15 (24.2)	21 (34.4)	17 (28.8)	23 (38.3)
Minimally improved	23 (37.1)	21 (34.4)	13 (22.0)	11 (18.3)
No change	11 (17.7)	7 (11.5)	6 (10.2)	5 (8.3)
Minimally worse	0	0	1 (1.7)	1 (1.7)
Much worse	1 (1.6)	0	0	1 (1.7)
Very much worse	1 (1.6)	0	0	0

Based on Narita 2022[13]
kg = kilogram; µg = microgram

FISHBANE 2020

The 50 mcg dose led to a greater proportion of responders than placebo at 8 weeks, but similar benefits were not seen for the 1.0 or 1.5 µg doses (Table 3.35).

Table 3.35: The Patient Global Impression of Change (PGIC) at 8 weeks

No. (%)	Placebo (n=45)	0.5 µg/kg of difelikefalin (n=44)	1.0 µg/kg of difelikefalin (n=41)	1.5 µg/kg of difelikefalin (n=44)
Responder rate at week 8 n (%)	18(41.9)	32(78)*	25(62.5)	22(56.4)
Based on Fishbane 2020[12] kg = kilogram *Significant difference (p=<0.01) compared to placebo				

EAG comment: The 8-week data from Narita 2022[13] and Fishbane 2020[12] have not been pooled with the 12 week data from the KALM RCTs because of the relatively large difference in follow-up time that might have an impact on outcome.

3.2.5.2 Health-related quality of life

3.2.5.2.1 Itch-related quality of life – change from baseline in total 5-D Itch scale score at end of Week 12

The five dimensions of itch assessed are degree, duration, direction, disability, and distribution.[38] Each domain is scored 1-5 with a total score range of 5-25 (5=no pruritus, 25=most severe pruritus). A 5-point change is considered clinically significant.

KALM-1

Compared with the placebo group, the difelikefalin group again showed a statistically significant (p<.001) reduction in total 5-D Itch scale score at the end of week 12, with a LS mean treatment group difference of -1.3 (95% CI, -2.0 to -0.5) (Table 3.36). The findings for the ANCOVA analysis in the per protocol population were also in favour of difelikefalin and statistically significant (p<.001).

Table 3.36: ANCOVA analysis of change from baseline in total 5-D Itch score at week 12 – MI (population: ITT)

	Placebo (N=189)	DFK (N=189)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-3.7	-5.0	-1.3	<.001
(SE)	(0.33)	(0.33)	(0.38)	-
95% CI	(-4.4, -3.1)	(-5.7, -4.4)	(-2.0, -0.5)	-
Based on Table 22, CS[1] ANCOVA = analysis of covariance; CI = confidence interval; CS = company submission; ITT = intention-to-treat; LS = least squares; MAR = missing-at-random; MI = multiple imputation; SE = standard error. Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption.				

A mixed-effects model for repeated measures ((MMRM) sensitivity analysis was performed on the change from baseline in the total 5-D Itch scale score by time point with no data imputation. At the end of week 12, the LS mean change from baseline in total 5-D Itch scale score was -4.9 (95% CI, -5.6 to -4.3) in the difelikefalin group and -3.6 (95% CI, -4.2 to -2.9) in the placebo group. The LS mean treatment group difference (difelikefalin minus placebo) of -1.3 (95% CI, -2.2 to -0.5) was statistically significant (p=.002) in favour of difelikefalin.

EAG comment: The company claimed that difelikefalin had achieved a clinically significant effect because it had reached a change of -5 in that single group. However, the MID of 5 should really be applied to the MD between difelikefalin and placebo, because this represents the true treatment effect (with intervening effects like the placebo effect eliminated). Using this criterion, the effect was not clinically significant.

KALM-2

Table 3.37 summarises the change from baseline in total 5-D Itch scale score at the end of week 12, using ANCOVA with MI of missing data under a MAR assumption. Compared with the placebo group, the difelikefalin group showed a greater reduction in total 5-D Itch scale score at the end of week 12, with a LS mean treatment group difference of -1.1 (95% CI, -1.7 to -0.4). Although the nominal p-value was 0.002, this difference could not be declared to be statistically significant based on the hierarchical testing order, as the prior secondary endpoint (Skindex-10 at week 12) was not statistically significant.

Table 3.37: ANCOVA analysis of change from baseline in total 5-D Itch score at week 12 - MI (population: ITT)

End of week 12 change from baseline	Placebo (n=236)	Difelikefalin (n=237)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-3.8	-4.9	-1.1	0.002
(SE)	(0.36)	(0.36)	(0.35)	-
95% CI	(-4.5, -3.1)	(-5.6, -4.2)	(-1.7, -0.4)	-

Based on Table 33, CS[1]
ANCOVA = analysis of covariance; CI = confidence interval; CS = company submission; DFK = difelikefalin; ITT = intention-to-treat; LS = least squares; MAR = missing-at-random; MI = multiple imputation; SE = standard error
Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score, region, and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption.

Sensitivity analyses were performed for the 5-D Itch scale secondary efficacy outcome. One such analysis was an MMRM sensitivity analysis of the change from baseline in the total 5-D Itch scale score by time point with no imputation for missing data. At the end of week 12, the LS mean treatment group difference (difelikefalin minus placebo) of -1.2 (95% CI, -1.9 to –0.5) was in favour of difelikefalin.

Pooling of KALM-1 and KALM-2 main data

In the pooled analysis, LS mean (95% CI) change from baseline to week 12 in 5-D Itch total score was −4.9 (−5.4, −4.5) in the difelikefalin group and −3.7 (−4.1, −3.3) in the placebo group (P<0.001). Unfortunately, no between-arm results were provided. This result was not reported in the CS, and has been derived from Topf 2022.[26]

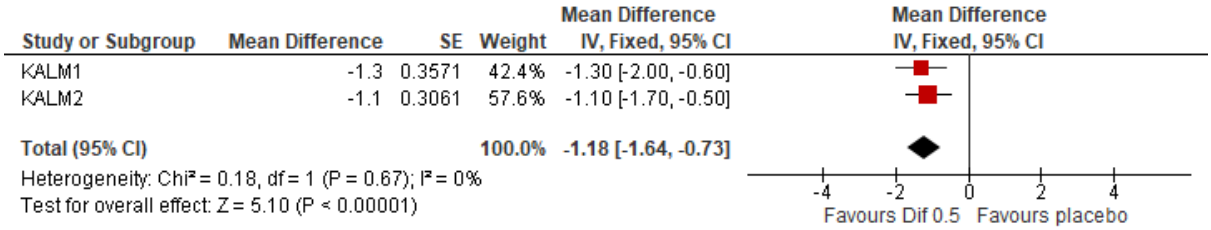
EAG comment:

- In Topf 2022[26], it is stated that **"***Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants".* This implies that a meta-analysis was not used for pooling, and that the individual patient data from the two studies were simply added together. This approach was probably inappropriate because "*such an analysis is against the main principle of meta-analysis…a pooled analysis like this will yield over-precise results (CIs too narrow) and it may well lead to bias if any of the trials has*

unequal numbers in the two arms".[30] The EAG has therefore carried out a Mantel-Haenszel meta-analysis of the results from KALM-1 and KALM-2 on Review Manager 5.3. Heterogeneity was not serious (I^2=0%) and so a fixed effects model was used.

- The pooled MD (95% CI) was -1.18 (-1.64 to -0.73) (Figure 3.9**)**. This provides an informative between-arm effect that was lacking in Topf 2022.[26]

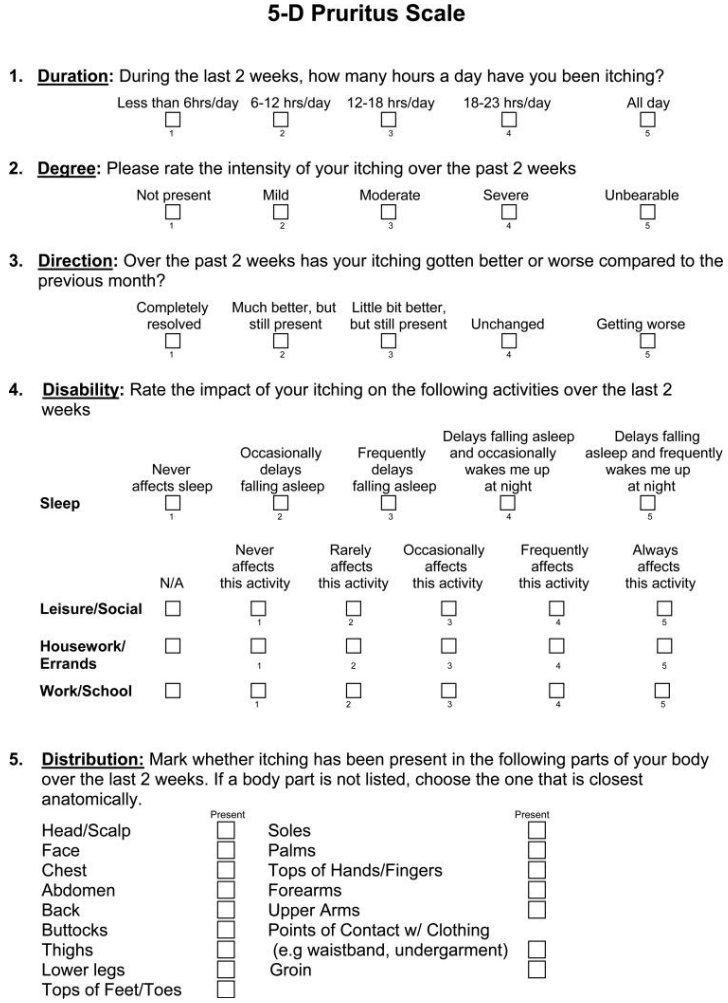
Figure 3.9: Meta-analysis of KALM-1 and KALM-2



CI = confidence interval; Dif = difelikefalin; SE = standard error; IV = intravenous

- The rationale for the 5-D Itch score was not apparent. The company was asked to provide evidence of the methodology and validation of the 5-D Itch score outcome used in the CS.[1] The company responded by stating that, "*The 5-D Itch scale is a multidimensional questionnaire which assesses itch severity and itch-related quality of life over the previous 2 weeks. The questionnaire covers 5 dimensions of itch, including the degree, duration of itch/day, direction (improvement/worsening), disability (impact on activities such as work), and body distribution of itch. The total 5-D Itch scale score ranges from 5 to 25, with higher scores indicating worse responses. The scale has been validated in patients with chronic pruritus, including haemodialysis patients, and has been shown to be sensitive to changes in pruritus over time.*[38] *Additionally, with limited options for itching scales (Appendix N: Clinical Opinion and consensus report), 5D-Itch is both commonly used and produces valid and reproducible results. It is therefore an appropriate choice for measuring itch in CKD-aP patients for this submission……".*[5] Figure 3.10 represents the 5-D Itch scale tool.

Figure 3.10: 5-D Itch scale



Based on Figure 2 of the company's response to clarification[5] which in turn used Elman 2010[38] as the source material.
5-D = 5 dimensions

- The EAG checked the Elman 2010[38] reference and considers that it supports the company's statement. The EAG is satisfied with this response.

3.2.5.2.2 Itch-related QoL – change from baseline in 5-D itch scale total score at 8 weeks

NARITA (2022)

The 0.25 and 0.5 mg doses of difelikefalin were reported by the authors to improve 5-D Itch scale total score relative to placebo at 8 weeks, but no statistical analysis was performed (Table 3.38).

Table 3.38: 5-D itch total score at 8 weeks

	Placebo	0.25 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
5-D itch change in total score at 8 weeks (adjusted weekly mean ± standard error)	-5.8 (0.3)	-6.6 (0.3)	-6.5 (0.4)	-6.8 (0.3)
Based on Narita 2022[13] 5-D = five dimensions; kg = kilogramme; µg = microgram				

FISHBANE 2020

Doses of 0.5, 1.0 and 1.5 µg of difelikefalin were reported by the authors to improve 5-D Itch scale total score relative to placebo at 8 weeks (Table 3.39).

Adjusted LS mean differences (difelikefalin – placebo) were as follows:

- 0.5 µg/kg of difelikefalin (adjusted mean difference −2.9; 95% CI, −4.4 to −1.5)
- 1. 0 µg/kg of difelikefalin (adjusted mean difference −2.7; 95% CI, −4.2 to -1.1).
- 1. 5 µg/kg of difelikefalin (adjusted mean difference −1.9; 95% CI, −3.5 to −0.4).

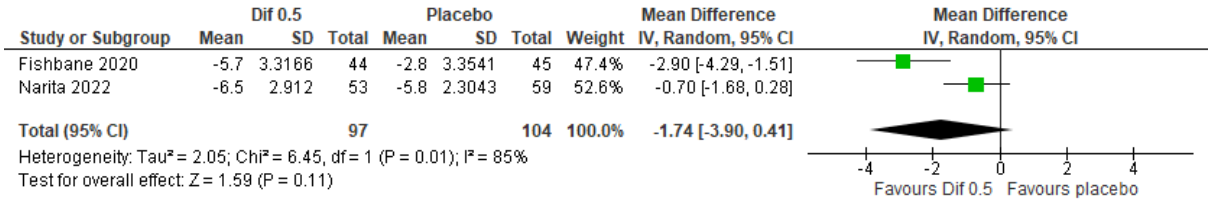
Table 3.39: 5-D itch total score at 8 weeks

	Placebo	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin	1.5 µg/kg of difelikefalin
5-D itch mean change from baseline score at 8 weeks (adjusted weekly mean [SE])	-2.8 (0.5)	-5.7 (0.5)*	-5.4 (0.6)*	-4.7 (0.6)**
Based on Narita 2022[13] *Significant difference (p=<0.01) compared to placebo **Significant difference (p=0.016) compared to placebo SE = standard error; kg = kilogramme; µg = microgram				

EAG comment:

- The 8-week data from Narita 2022[13] and Fishbane 2020[12] were not been pooled with the 12 week data from the KALM RCTs because of the relatively large difference in follow up time and the possible impact this might have on outcome.
- A meta-analysis was carried out by the EAG to synthesise the results of Narita 2022[13] and Fishbane 2020[12]. Heterogeneity was very serious (I^2=85%) and so a random-effects method was used. The pooled result suggested a trend for a benefit of difelikefalin 50 mcg/kg over placebo (Figure 3.11), but there was some uncertainty about the true direction of effect in the population.

Figure 3.11: Meta-analysis of Fishbane 2020 and Narita 2022



CI = confidence interval; Dif = difelikefalin; SD = standard deviation; IV = intravenous

3.2.5.2.3 Itch-related QoL – change from baseline in Skindex-10 scale score at end of Week 12

The Skindex-10 scale is a patient-reported measurement of itch and its impact on QoL in the last week and has been specifically developed for patients with CKD-aP. It consists of 10 questions across three domains (disease, mood/emotional stress, and social functioning). Each of the 10 questions is scored from 0 to 6 (0=never bothered; 6=always bothered), meaning the total score varies from 0 to 60. A 15-point change in score is regarded as clinically significant.

KALM-1

At the end of week 12, the LS mean change in total Skindex-10 scale score was greater in the difelikefalin group than in the placebo group (-17.2 versus -12.0). A statistically significant LS mean difference was noted: -5.1 (95% CI, -8.0 to -2.3); p<.001 (Table 3.40).

Table 3.40: ANCOVA Analysis of change from baseline in total Skindex-10 scale at week 12 - MI (Population: ITT)

	Placebo (N=189)	DFK (N=189)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-12.0	-17.2	-5.1	<.001
(SE)	(1.24)	(1.26)	(1.44)	-
95% CI	(-14.5, -9.6)	(-19.6, -14.7)	(-8.0, -2.3)	-

Based on Table 23, CS[1]
ANCOVA = analysis of covariance; CI = confidence interval, CS = company submission; DFK = difelikefalin; LS = least squares; MAR = missing-at-random; MI = multiple imputation; SE = standard error; ITT = intention-to-treat
Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption

The results of the MMRM sensitivity analysis (no imputation) of total Skindex-10 scale at week 12 for the ITT population were similar to ANCOVA with MI. At the end of week 12, the treatment group difference difelikefalin minus placebo) of -5.2 (95% CI -8.3 to -2.1) was in favour of difelikefalin and statistically significant (p<0.001).

EAG comment: The company claimed that difelikefalin had achieved a clinically significant effect because it had reached a change of >-15 in that single group. However, the MID of 15 should really be applied to the MD between DFK and placebo, because this represents the true treatment effect (with intervening effects like the placebo effect eliminated). Using this criterion, the effect was not clinically significant.

KALM-2

Table 3.41 summarises the change from baseline in total Skindex-10 scale score at the end of week 12 for the ITT population, using ANCOVA with MI under the MAR assumption. Compared with the placebo group, the difelikefalin group showed a numerically greater reduction in LS mean total Skindex-10 scale score (-16.6 versus -14.8) at the end of week 12, with a LS mean treatment group difference of -1.8 (95% CI, -4.3 to 0.8), which was not statistically significant (p=0.171).

Table 3.41: ANCOVA analysis of change from baseline in total Skindex-10 scale at week 12 - MI under MAR assumption (population: ITT)

End of week 12 change from baseline	Placebo (N=236)	Difelikefalin (N=237)	Difference in LS means (DFK minus placebo)	P-value
LS mean	-14.8	-16.6	-1.8	0.171
(SE)	(1.32)	(1.35)	(1.29)	-
95% CI	(-17.4, -12.2)	(-19.3, -14.0)	(-4.3, 0.8)	-

Based on Table 32, CS[1]

End of week 12 change from baseline	Placebo (N=236)	Difelikefalin (N=237)	Difference in LS means (DFK minus placebo)	P-value
ANCOVA = analysis of covariance; CI = confidence interval, CS = company submission; DFK = difelikefalin; LS = least squares; MAR = missing-at-random; MI = multiple imputation; SE = standard error; ITT = intention-to-treat Note: LS means, SEs, and 95% CIs were based on ANCOVA with fixed effects for treatment, with baseline score and the randomisation stratification variables as covariates. Missing values were imputed using MI under MAR missing data assumption				

The MMRM and ANCOVA sensitivity analysis were performed for the Skindex-10 scale. All analyses were reported to be consistent with the key analysis, but no data were shown.

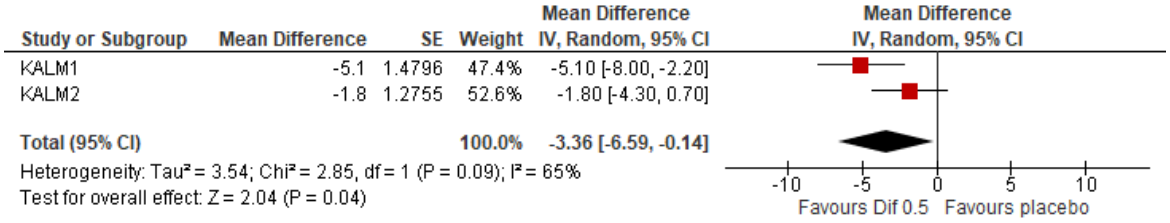
Pooling of KALM-1 and KALM-2 main data

In the pooled analysis, LS mean (95% CI) change from baseline to week 12 in Skindex-10 total score was −16.9 (−18.6, −15.2) in the difelikefalin group and −13.5 (−15.1, −11.8) in the placebo group (P=0.001). Unfortunately, no between-arm results were provided. This result was not reported in the CS and has been derived from Topf 2022.[26]

EAG comment:

- In Topf 2022[26] it is stated that **"***Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants".* This implies that a meta-analysis was not used for pooling, and that the individual patient data from the two studies were simply added together. This approach was probably inappropriate because "*such an analysis is against the main principle of meta-analysis…a pooled analysis like this will yield over-precise results (CIs too narrow) and it may well lead to bias if any of the trials has unequal numbers in the two arms*".[30] The EAG has therefore carried out a Mantel-Haenszel meta-analysis of the results from KALM-1 and KALM-2 on Review Manager 5.3. Heterogeneity was serious (I^2=65%) and so a random effects model was used.
- The pooled MD (95% CI) was -3.36 (-6.59 to -0.14) (Figure 3.12). This provides an informative between-arm effect that was lacking in Topf 2022.[26]

Figure 3.12: Meta-analysis of KALM-1 and KALM-2



CI = confidence interval; Dif = difelikefalin; SE = standard error; IV = intravenous

3.2.5.2.4 Itch-related QoL – change from baseline in Skindex-10 scale score at end of Week 8

FISHBANE 2020

Doses at 0.5, 1.0 and 1.5 mcg of difelikefalin were reported by the authors to improve Skindex-10 scale score relative to placebo at 8 weeks (Table 3.42).

Adjusted LS mean differences (difelikefalin – placebo) were as follows:

- 0.5 µg/kg of difelikefalin (adjusted mean difference −10.4; 95% CI, −16 to −4.8)

- 1. 0 µg/kg of difelikefalin (adjusted mean difference −7.2; 95% CI, −13.1 to -1.4).
- 1. 5 µg/kg of difelikefalin (adjusted mean difference −6.8; 95% CI, −12.8 to −0.8).

Table 3.42: Skindex-10 score at 8 weeks

	Placebo	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin	1.5 µg/kg of difelikefalin
5-D Itch mean change from baseline score at 8 weeks (adjusted weekly mean [(SE])	-8.2 (2)	-18.7 (2)*	-15.5 (2.2)**	-15.1 (2.3)***
Based on Fishbane 2020[12] SE = standard error *Significant difference (p<0.001) compared to placebo **Significant difference (p=0.016) compared to placebo ***Significant difference (p=0.026) compared to placebo				

3.2.5.2.5 Itch-related QoL – Skindex-16 overall score at 8 weeks

NARITA 2022

The 0.25 and 0.5 mcg doses of difelikefalin were reported by the authors to improve Skindex-16 overall score relative to placebo at 8 weeks, but no statistical analysis was performed (Table 3.43).

Table 3.43: Skindex-16 overall score at 8 weeks

	Placebo	0.35 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
Skindex-16 overall score at 8 weeks (adjusted weekly mean [SE])	-24.04 (1.94)	-24.25 (1.96)	-27.79 (2.05)	-22.69 (2.04)
Based on Narita 2022[13] SE = standard error				

3.2.5.2.6 Longer term outcome: 5D itch score at week 52

KALM-1 OLE

[redacted]

When patients randomised to placebo during the double-blind treatment period transitioned to active treatment (between double-blind week 12 and open-label week 4), [redacted]

[redacted] Over the course of long-term treatment through week 52, the total number of patients [redacted] Table 3.44

Figure 3.13: [redacted]

[redacted]

Table 3.44: [redacted]

	[redacted]		[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted][1] [redacted]				

KALM-2 OLE

In the first analysis of treatment effect, all visits in both the double-blind treatment period and open-label treatment period were included; the baseline was the 5-D Itch scale total score collected on day 1 of the double-blind treatment period prior to randomisation (i.e., double-blind baseline). In the second analysis, only the visits in the open-label treatment period were included; the baseline was the last 5-D Itch scale total score collected in the double-blind treatment period (i.e., open-label baseline). The study was stopped early by the sponsor due to reasons unrelated to safety or lack of drug effect; only limited meaningful conclusions could be drawn from the small number of patients that completed 52 weeks of treatment (n=5). Thus, results at week 36 are discussed (n=52).

The mean (SD) baseline 5-D Itch score was [redacted] At the first assessment post-baseline (after the double-blind week 4, on d y 29), [redacted]

[redacted]

Table 3.45, [redacted]

When patients randomised to placebo during the double-blind treatment period transitioned to open-label treatment (between double-blind week 12 and open-label week 4), [redacted] at the end of open-label week 4 for double-blind placebo (n=200) and difelikefalin (n=167) patients respectively). Over the course of long-term treatment through week 36, the total number of patients decreased at each time point, due to discontinuation (with two patients discontinuing due to lack of therapeutic efficacy). The change from baseline was [redacted] at the end of open-label week 36 for placebo/difelikefalin [n=30] and difelikefalin/difelikefalin [n=22] patients respectively).

Figure 3.14: [redacted]

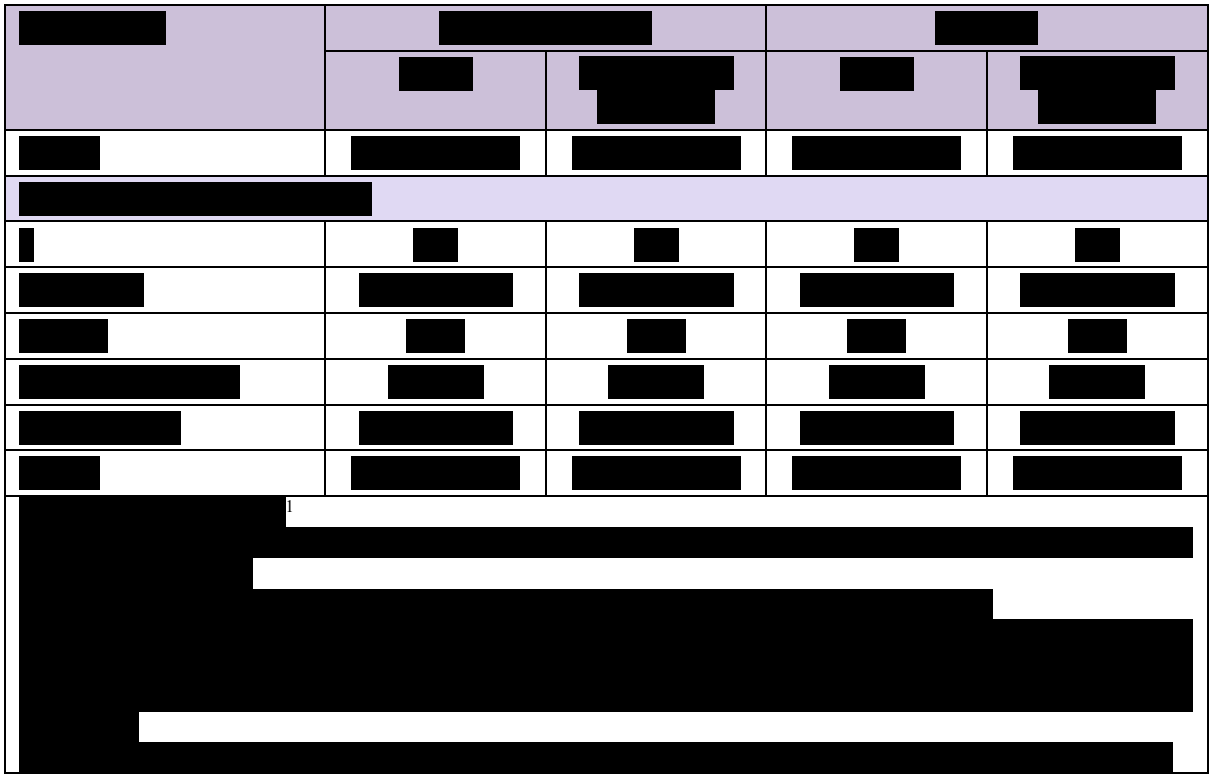
[redacted]

[redacted][1]

[redacted]

Table 3.45: [redacted]

[redacted]	[redacted]		[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]



EAG comment:

- The end of page 66 of Document B discusses discontinuation in relation to the KALM-1 OLE study, stating that it was not due to lack of efficacy, but no positive reasons for discontinuation are listed. The company was asked to comment on the early stopping. The company stated that, "*Reason for discontinuation include adverse events (26 patients), withdrawn consent (17 patients), non-compliance (5 patients) and lost to follow-up (5 patients)"*.[5]
- Similarly, page 96 of Document B[1] mentions early termination of KALM-2 saying that this was due to an administrative decision and not related to efficacy or safety. The company were asked to comment on the early stopping. The company stated that, "*The open-label portion of the KALM-2 study was halted by the sponsor to enable analysis and reporting of the results to be completed for inclusion in the marketing authorisation application to the FDA"*. The EAG is satisfied with this response.
- In Section B.2.6, pooled efficacy results appear to be reported for all four KALM-1 and KALM-2 studies, although these include both double-blinded and open-label extension designs. The company were asked to provide the rationale, methods and analysis of how pooling was executed. The company stated that:[5]
    - "*KALM 1 and 2 were similarly designed studies with similar inclusion and exclusion criteria and similar endpoints. In both studies, significantly greater proportions of participants in the difelikefalin group achieved ≥3- and ≥4-point reductions in weekly means of daily Worst Itching Intensity Numerical Rating Scale (WI-NRS) scores versus the placebo group. Pooled data from the KALM-1 and KALM-2 studies was analysed to obtain a combined estimate of the treatment effects of difelikefalin in HD participants with moderate to severe pruritus, including QoL endpoints. The statistical methods of how pooling was executed are included in the, now fully published, Topf et al. manuscript*[26, 39] *and are as follows.*

- *Efficacy analyses were conducted in the intent-to-treat population from the pooled KALM-1 and KALM-2 studies, which consisted of all randomized participants. Differences between placebo and difelikefalin were analysed using a logistic regression model containing terms for the treatment group, baseline WI-NRS score, use of an anti-itch medication during the week before randomization, presence of specific medical conditions, and geographic region. For the analysis of the proportions of participants who achieved ≥3-point or ≥4-point reductions in the weekly mean WI-NRS scores, missing weekly WI-NRS scores were imputed by multiple imputation under a missing-at-random assumption. Participants who reported <4 daily WI-NRS scores at week 12 or who discontinued treatment early were considered non-responders in the analysis of the complete WI-NRS response. Proportions of participants achieving a ≥5-point improvement in the 5-D Itch total score and a ≥15-point improvement in the Skindex-10 total score were analysed without imputation for missing values. Proportions of participants achieving a ≥5-point improvement in 5-D Itch total score are reported for the pooled population during the placebo-controlled, double-blind period (12 weeks) and the open-label extension period (up to 52 weeks).*
- *Continuous efficacy endpoints were analysed by a mixed model for repeated measures, with terms for treatment, visit, treatment-by-visit interaction, baseline score, use of an anti-itch medication during the week before randomization, the presence of specific medical conditions, and geographic region. An unstructured covariance structure was applied to model the within participant errors. Missing values were not imputed. The mean improvements from baseline in 5-D Itch total score are reported for the pooled population during the placebo-controlled, double-blind period (12 weeks) and the open label extension period (up to 52 weeks).*
- *The subgroup analyses of ≥3-point and ≥4-point reductions from baseline in the weekly mean WI-NRS scores were performed using the same methodology as that employed for the full intent-to-treat population.*"[5]

- The above response appears to imply that pooling was only performed for the RCTs and does not offer a methodology for pooling across the non-RCT data as well. Perusal of Topf 2022[26, 39] suggests that the pooling may only have occurred for the RCTs, but Section B2.6 in the CS[1] is ambiguous about what data were included in the pooling. The OLE studies are mentioned, implying that these were also pooled. Going by the overall available information, the EAG considers that study inclusion in the pooled analysis remains unclear.
- There seem to be some instances of p-values being the only information provided about estimation. For example, at the top of page 63 of Document B in the CS[1] it is stated that: "*the findings for the per protocol population were also in favour of difelikefalin and statistically significant (p<.001)*", with no further information given. Even if the full estimate is in the CSR,[6, 7] full results should be provided in the CS.[1] The company were asked to provide 95% CIs for all between-arm estimates in the CS.[1]. The company responded by fully updating the CS, and this has enabled updating of results in the EAG submission.[5, 29]

3.2.5.3 Sub-groups

All subgroup analyses were pre-planned and were only conducted in the RCTs.

KALM-1

Stratification factors were use of anti-itch medication or not at baseline, and presence or absence of specific medical conditions at baseline. These specific medical conditions included:

- History of fall or fracture (related to fall)
- Confusional state or mental status change or altered mental status or disorientation
- Gait disturbance or movement disorder

Anti-itch medication

The difelikefalin group showed a greater percentage of patients achieving a ≥3-point improvement from baseline in WI-NRS scores at week 12 regardless of use of anti-itch medication, which was considered to be statistically significant in both cases: 48.8% versus 27.2%, p=0.001 for no use of anti-itch medications at baseline; 53.2% versus 29.4%, p=0.005 for use of anti-itch medication (Table 3.46).

Medical conditions

The difelikefalin group showed a greater percentage of patients achieving a ≥3-point improvement from baseline in WI-NRS scores at week 12 regardless of the existence of medical conditions at baseline, which was considered to be statistically significant in both cases: 50.8% versus 31.8%, p=0.001 for no medical conditions at baseline; 63.7% versus 17.6%, p=0.001 for medical conditions at baseline (Table 3.46).

Table 3.46: Summary of week 12 ≥3-point WI-NRS improvement by stratification factors in KALM-1 study

	Placebo	Difelikefalin
Anti-itch medication at baseline = no		
N	111	117
Observed week 12 ≥3-point NRS improvement [1] — (%)		
Yes	31 (31.3%)	53 (53.5%)
No	68 (68.7%)	46 (46.5%)
Missing	12	18
LS means estimate of percent with improvement [2]		
Percent (95% CI)	27.2% (17.8%, 39.2%)	48.8% (36.4%, 61.4%)
Odds ratio (95% CI)		2.55 (1.44, 4.53)
P-value		.001
Anti-itch medication at baseline = yes		
N	78	72
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	20 (30.3%)	29 (50.0%)
No	46 (69.7%)	29 (50.0%)
Missing	12	14
LS means estimate of percent with improvement [2]		
Percent (95% CI)	29.4% (18.8%, 42.8%)	53.2% (39.2%, 66.6%)
Odds ratio (95% CI)		2.73 (1.35, 5.51)
P-value		.005

	Placebo	Difelikefalin
Medical conditions at baseline = no		
N	161	164
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	46 (32.5%)	67 (50.4%)
No	95 (67.4%)	66 (49.6%)
Missing	20	31
LS means estimate of percent with improvement [2]		
Percent (95% CI)	31.8% (24.6%, 40.0%)	50.8% (42.6%, 58.9%)
Odds ratio (95% CI)		2.21 (1.38, 3.55)
P-value		.001
Medical conditions at baseline = yes		
N	28	25
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	5 (20.8%)	15 (62.5%)
No	19 (79.2%)	9 (37.5%)
Missing	4	1
LS means estimate of percent with improvement [2]		
Percent (95% CI)	17.6% (7.4%, 36.3%)	63.7% (43.2%, 80.2%)
Odds ratio (95% CI)		8.20 (2.24, 29.99)
P-value		.001

Based on Table 74, CS[1]
CI = confidence intervals; CS = company submission; LS = least squares; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale

KALM-2

Stratification factors were use or non-use of anti-itch medication at baseline, and presence of absence of specific medical conditions These specific medical conditions included:

- History of fall or fracture (related to fall)
- Confusional state or mental status change or altered mental status or disorientation
- Gait disturbance or movement disorder

Anti-itch medication

The patients using anti-itch medications at baseline had a greater treatment difference (OR = 2.15; 95% CI, 1.09 to 4.25) favouring difelikefalin compared with patients not using anti-itch medications at baseline (OR = 1.36; 95% CI, 0.84 to 2.20) (Table 3.47).

Medical conditions

The patients without medical conditions at baseline had a similar treatment difference (OR = 1.52; 95% CI, 0.99 to 2.32) (trend favouring difelikefalin) compared with patients with medical conditions at baseline (OR = 1.71; 95% CI, 0.67 to 4.36) (trend favouring difelikefalin) (Table 3.47).

Table 3.47: Summary of week 12 ≥3-point WI-NRS improvement by stratification factors in KALM-2 study

	Placebo	Difelikefalin
Anti-itch medication at baseline = no		
n	161	160
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	51 (38.3%)	60 (48.0%)
No	82 (61.7%)	65 (52.0%)
Missing	18	25
LS means estimate of percent with improvement [2]		
Percent (95% CI)	36.0% (23.8%, 50.3%)	43.3% (30.1%, 57.5%)
Odds ratio (95% CI)		1.36 (0.84, 2.20)
P-value		0.213
Anti-itch medication at baseline = yes		
n	85	87
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	26 (35.1%)	35 (53.0%)
No	48 (64.9%)	31 (47.0%)
Missing	11	21
LS means estimate of percent with improvement [2]		
Percent (95% CI)	45.9% (32.4%, 60.1%)	64.6% (49.3%, 77.4%)
Odds ratio (95% CI)		2.15 (1.09, 4.25)
P-value		0.028
Medical conditions at baseline = no		
n	199	195
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	66 (37.7%)	79 (49.4%)
No	109 (62.3%)	81 (50.6%)
Missing	24	35
LS means estimate of percent with improvement [2]		
Percent (95% CI)	41.2% (32.6%, 50.4%)	51.6% (42.2%, 60.8%)
Odds ratio (95% CI)		1.52 (0.99, 2.32)
P-value		0.054
Medical conditions at baseline = yes		
n	37	42
Observed week 12 ≥3-point NRS improvement [1] - n(%)		
Yes	11 (34.4%)	16 (51.6%)
No	21 (65.6%)	15 (48.4%)
Missing	5	11

	Placebo	Difelikefalin
LS means estimate of percent with improvement [2]		
Percent (95% CI)	42.6% (26.6%, 60.3%)	55.9% (37.8%, 72.6%)
Odds ratio (95% CI)		1.71 (0.67, 4.36)
P-value		0.259
Based on Table 75, CS[1] CI = confidence intervals; CS = company submission; LS = least squares; NRS = numerical rating scale; WI-NRS = Worst Itching Intensity Numerical Rating Scale		

For the KALM-2 data only, the proportion of ITT patients achieving a ≥3- and ≥4-point improvement in mean 24-hour WI-NRS from baseline at week 12 by region (United States of America [USA], Asia, Eastern Europe, or Western Europe) was also assessed. The proportion of 3-point and 4-point responders was larger in the difelikefalin group compared to placebo across all regions. Treatment differences between difelikefalin and placebo were similar in the US and Western Europe (ORs of 1.25 and 1.30, respectively, for ≥3-point improvement and ORs of 1.48 and 1.21, respectively, for ≥4-point improvement). The difference between difelikefalin and placebo was generally larger in Eastern Europe than in other regions (OR of 3.06; 95% CI 1.38 to 6.80 for ≥3-point improvement and OR of 2.80; 95% CI 1.21 to 6.46 for ≥4-point improvement).

The number of patients randomised to Asian countries was small (n=20). The point estimates for the OR in Asia varied depending on the endpoint (OR of 1.90; 95% CI 0.21 to 17.07 for ≥3-point improvement and OR of 5.42; 95% CI 0.13 to 226.01 for ≥4-point improvement).

Finally, the proportion of ITT patients achieving a ≥3- and ≥4-point improvement in mean 24-hour WI-NRS from baseline at week 12 was analysed by dialysis type (haemodialysis or haemodiafiltration). Results were numerically similar regardless of dialysis type for ≥3-point improvement (OR of 1.55 and 1.82 for haemodialysis and haemodiafiltration, respectively) and for ≥4-point improvement (PR of 1.70 and 1.72 for haemodialysis and haemodiafiltration, respectively).

EAG comment: Table 75 in Appendix E of Document B[1] reports on the subgroup analysis for KALM-2, as n=246 in the placebo arm and n=247 in the difelikefalin arm which is different from the n=236 and n=237 reported in the results of KALM-2. In addition, the patients not receiving anti-itch medication at baseline do not sum to the total. The company responded by stating that n=246 and n=247 were typos and that this has been updated in the updated version of the CS.[5, 29]

3.2.5.4 General EAG comments on section 3.2.5

- All RCT results given in Section 3.2.5 are ITT analyses where available. Per-protocol analyses were presented in the CS[1] but have not been reproduced in this report because ITT analyses present lower risk of attrition bias.
- The focus of the efficacy section of this report has been the results from the RCTs. Results from the OLE portions of the two KALM studies, and the single arm CLIN3105 study were given less weight as the lack of a control group makes it impossible to differentiate treatment effects from intervening effects (common threats to internal validity such as placebo effects or natural history effects). However, since the economic model uses the 5-D Itch score at baseline (week 12 of double-blind phase) and at week 52, these results were also presented.

3.2.6 Adverse events

3.2.6.1 Adverse events reported in the CS

Although more than 60% of patients experienced an AE with difelikefalin in KALM-1 and KALM-2, the rate and type of AEs observed with difelikefalin treatment were comparable with those observed with placebo.

Table 3.48 summarises the occurrence of treatment-emergent adverse events (TEAEs) and deaths during the double-blind treatment periods of KALM-1 and KALM-2. In the KALM-1 study, the rate of patients experiencing at least one TEAE during the double-blind treatment period was 68.8% in the difelikefalin group versus 62.2% in the placebo group. The findings of the KALM-2 study were consistent with those of KALM-1: 68.1% of patients receiving difelikefalin and 61.4% of those receiving placebo experienced at least one TEAE.

The rate of serious TEAEs was also comparable between difelikefalin and placebo, with 25.9% in the difelikefalin group and 21.8% in the placebo group experiencing at least one serious TEAE during the double-blind treatment period of KALM-1; the equivalent figures for KALM-2 were 24.7% and 21.6%. The number of deaths was very low, and consistent across the treatment arms of both studies. Both deaths in the difelikefalin group of KALM-1 were attributed to sepsis; the two deaths in the placebo group were due to septic shock. In KALM-1, TEAEs led to discontinuation in 4.8% of patients in the placebo group and 7.9% of patients in the difelikefalin group. A similar pattern was observed in KALM-2, suggesting that difelikefalin has a favourable safety profile.

Table 3.48: Summary of TEAEs and deaths during the double-blind treatment periods of KALM-1 and KALM-2 (double-blind safety populations)

	KALM-1		KALM-2	
	Placebo (n=188)	**Difelikefalin (n=189)**	**Placebo (n=236)**	**Difelikefalin (n=235)**
Number of patients with at least one TEAE	117 (62.2%)	130 (68.8%)	145 (61.4%)	160 (68.1%)
Number of patients with at least one serious TEAE	41 (21.8%)	49 (25.9%)	51 (21.6%)	58 (24.7%)
Number of deaths	2 (1.1%)	2 (1.1%)	2 (0.8%)	2 (0.9%)
Number of patients with at least one TEAE resulting in study drug discontinuation	9 (4.8%)	15 (7.9%)	8 (3.4%)	13 (5.5%)
Based on Table 35, CS[1] CS = company submission; TEAE = treatment-emergent adverse event				

Safety data from the OLE of KALM-1 and KALM-2 also show comparable TEAE rates across the treatment groups (Table 3.49).

Table 3.49: Summary of TEAEs during the open-label treatment period of KALM-1 and KALM-2 (open-label safety population)

	KALM-1 OLE		KALM-2 OLE	
	Placebo/ difelikefalin (n=162)	Difelikefalin/ difelikefalin (n=162)	Placebo/ difelikefalin (n=210)	Difelikefalin/ difelikefalin (n=162)
Number of patients with at least one TEAE	132 (81.5%)	125 (82.8%)	117 (61.9%)	256 (64.2%)
Number of patients with at least one serious TEAE	88 (54.3%)	79 (52.3%)	61 (32.3%)	130 (32.6%)
Number of deaths	12 (7.4%)	10 (6.6%)	7 (3.7%)	15 (3.8%)
Number of patients with at least one TEAE resulting in study drug discontinuation	15 (9.3%)	10 (6.6%)	8 (4.2%)	20 (5.0%)

Based on Table 36, CS[1]
CS= company submission; TEAE = treatment-emergent adverse event

Similarly, in the pooled studies, patients reported TEAEs that were mostly mild to moderate in both the placebo-controlled period (difelikefalin: 57.5% [244/424] versus placebo: 52.6% [223/424]) and the OLE period (difelikefalin: 53.6% [427/796]). The incidence rate of common TEAEs and serious TEAEs did not increase with longer-term exposure.

It was demonstrated that 71.2% of patients in the difelikefalin group experienced TEAEs, versus 65.3% in the placebo group (Table 3.50). The rates of TEAEs leading to study drug discontinuation were low, and comparable between the placebo and difelikefalin groups: 4.0% and 6.8%, respectively.[40]. Due to its positive safety profile, difelikefalin is appropriate for the long-term treatment of CKD-aP.

Table 3.50: Summary of TEAEs according to a pooled analysis of the KALM-1 and KALM-2 safety population

	Pooled analysis		
	Placebo-controlled weeks 0-12		Placebo-controlled plus, OLE weeks 0 up to 64
	Placebo (n=424)	Difelikefalin (n=424)	Difelikefalin (n=796*)
Number of patients with any TEAE reported	277 (65.3%)	302 (71.2%)	640 (80.4%)
Number of patients with any non-fatal serious TEAEs reported	96 (22.6%)	107 (25.2%)	354 (44.5%)
Number of patients with any TEAE leading to death	5 (1.2%)	3 (0.7%)	37 (4.6%)
Number of patients with any TEAE leading to study drug discontinuation	17 (4%)	29 (6.8%)	72 (9%)

Based on Table 37, CS[1]
CS = company submission; OLE = open label extension; TEAE = treatment-emergent adverse event
*Number of patients exposed to difelikefalin in either the placebo-controlled period or the OLE n's are based on the safety population, defined during the double-blind period as randomised patients who received at least one

	Pooled analysis		
	Placebo-controlled weeks 0-12		Placebo-controlled plus, OLE weeks 0 up to 64
	Placebo (n=424)	Difelikefalin (n=424)	Difelikefalin (n=796*)
dose of double-blind study drug during the placebo-controlled period, and defined during the OLE period as patients who received at least one dose of study drug during the placebo-controlled or OLE period			

Table 3.51 presents the most commonly reported TEAEs in KALM-1 and KALM-2. In KALM-2, nausea and fall were experienced by ≥5% of patients; however, the rate at which these events occurred were comparable between the difelikefalin and placebo groups. The most commonly reported serious adverse events (SAEs) were hyperkalaemia (2.1% in both groups), pneumonia (1.6% in the difelikefalin group and 2.7% in the placebo group), sepsis (1.6% in the difelikefalin group and 2.1% in the placebo group), hypotension (1.6% in the difelikefalin group and 1.1% in the placebo group), and chronic obstructive pulmonary disease (1.6% in the difelikefalin group and 0.5% in the placebo group).

The TEAEs observed in KALM-1 and KALM-2 are consistent with those observed in CLIN2101 and other studies in the difelikefalin study programme.

Table 3.51: TEAEs ≥5% of any treatment group (double-blind treatment period of KALM-1 and KALM-2)

TEAEs at ≥5% frequency	KALM-1		KALM-2	
	Placebo (n=188)	Difelikefalin (n=189)	Placebo (n=236)	Difelikefalin (n=235)
Diarrhoea	7 (3.7%)	18 (9.5%)	13 (5.5%)	19 (8.1%)
Dizziness	2 (1.1%)	13 (6.9%)	12 (5.1%)	13 (5.5%)
Vomiting	6 (3.2%)	10 (5.3%)	14 (5.9%)	15 (6.4%)
Nasopharyngitis*	10 (5.3%)	6 (3.2%)	N/A	N/A
Fall*	5 (2.7%)	5 (2.6%)	12 (5.1%)	16 (6.8%)
Nausea*	9 (4.8%)	6 (3.2%)	10 (4.2%)	15 (6.4%)
Based on Table 38, CS[1] CS = company submission; N/A = not applicable; TEAE = treatment-emergent adverse event *TEAEs that occurred in ≥5% of patients in only KALM-1 or KALM-2, but results from both studies are reported for consistency				

In the 2-week discontinuation period of KALM-1, patients were evaluated for TEAEs potentially related to opioid withdrawal. The observed TEAE profile showed no suggestion of drug withdrawal following treatment cessation and no evidence of dependence development. Patients were also evaluated for potential signs and symptoms of opioid withdrawal using the Short Opiate Withdrawal Scale (ShOWS) and Objective Opiate Withdrawal Scale (OOWS) during the 2-week discontinuation period. The results from both scales indicated no signs of withdrawal in either treatment group. The proportion of patients experiencing at least one TEAE during the double-blind discontinuation period in the difelikefalin group (19.9%) was comparable with that of the placebo group (24.6%), as presented in Table 3.52. No TEAEs occurred with a frequency of ≥5% during the discontinuation period. Table 3.52 also presents System Organ Classes with a TEAE frequency of ≥5% during the discontinuation period.

Table 3.52: TEAEs by System Organ Class (double-blind discontinuation safety population of KALM-1)

	Placebo (n=179)	Difelikefalin (n=176)
Number of patients with ≥1 TEAE	44 (24.6%)	35 (19.9%)
TEAE by System Organ Class (≥5% frequency in either treatment arm)*		
Gastrointestinal disorders	8 (4.5%)	9 (5.1%)
Infections and infestations	10 (5.6%)	2 (1.1%)

Based on Table 39, CS[1]
CS = company submission; TEAE = treatment-emergent adverse event
*No specific TEAEs occurred at ≥5% frequency in either treatment arm. Data are shown as the overall rates for each system organ class

CLIN3105

Data from CLIN3105 support the safety and tolerability of difelikefalin reported in KALM-1 and KALM-2. Of 222 patients, 143 (64.4%) reported a total of 414 TEAEs over the course of the study. The most common TEAEs reported (≥4% of all patients) were diarrhoea (5.0%), nausea (4.5%), and hyperkalaemia (4.1%); these events were well tolerated. Overall, 6.3% of patients reported TEAEs that resulted in study drug discontinuation. Of the 143 who reported a TEAE, [redacted] 91 serious TEAEs in 45 patients (20.3%) were reported during the study, none of which were considered related to difelikefalin.

EAG comment:

- The CS[1] states that, "*Of the 143 who reported a treatment-emergent adverse event (TEAE), 68 patients (30.6%) had a maximum severity of mild, 56 (25.2%) had a maximum severity of moderate, and 19 (8.6%) had a maximum severity of severe.*" The company was asked to provide the scale used to judge the severity of TEAEs. The company responded as follows: "*The Investigator assessed the severity (i.e., intensity) of each adverse event (serious and non-serious) reported during the study based on his/her clinical judgment. The severity of each adverse event was assigned to one of the following categories:*[5]
  - *Mild: Transient, requires no special treatment, is easily tolerated by the patient, causes minimal discomfort, and does not interfere with the patient's daily activities*
  - *Moderate: Introduces a level of inconvenience or concern to the patient that may interfere with daily activities, but usually is ameliorated by simple therapeutic measures*
  - *Severe: Interrupts a patient's usual daily activity and requires systemic drug therapy or other treatment".*[5] The EAG is satisfied with this response
- The company was also asked to provide a list of TEAEs by severity. In response, the company provided the table below.[5]

Table 3.53: Incidence of TEAEs during the treatment period by MedDRA System Organ Class and Severity Population: Safety

System Organ Class	Severity	CR845 (N=222)
Number of patients with an event (N=143)	Mild	68 (30.6%)

System Organ Class	Severity	CR845 (N=222)
	Moderate	56 (25.2%)
	Severe	19 (8.6%)
Blood and lymphatic system disorders (N=7)	Mild	4 (1.8%)
	Moderate	2 (0.9%)
	Severe	1 (0.5%)
Cardiac disorders (N=19)	Mild	10 (4.5%)
	Moderate	5 (2.3%)
	Severe	4 (1.8%)
Ear and labyrinth disorders (N=3)	Mild	2 (0.9%)
	Moderate	1 (0.5%)
	Severe	0
Endocrine disorders (N=1)	Mild	0
	Moderate	1 (0.5%)
	Severe	0
Eye disorders (N=4)	Mild	4 (1.8%)
	Moderate	0
	Severe	0
Gastrointestinal disorders (N=36)	Mild	20 (9.0%)
	Moderate	14 (6.3%)
	Severe	2 (0.9%)
General disorders and administration site conditions (N=20)	Mild	11 (5.0%)
	Moderate	6 (2.7%)
	Severe	3 (1.4%)
Hepatobiliary disorders (N=1)	Mild	0
	Moderate	1 (0.5%)
	Severe	0
Infections and infestations (N=49)	Mild	27 (12.2%)
	Moderate	16 (7.2%)
	Severe	6 (2.7%)
Injury, poisoning and procedural complications (N=30)	Mild	15 (6.8%)
	Moderate	14 (6.3%)
	Severe	1 (0.5%)
Investigations (N=8)	Mild	5 (2.3%)
	Moderate	2 (0.9%)
	Severe	1 (0.5%)
Metabolism and nutrition disorders (N=20)	Mild	7 (3.2%)
	Moderate	9 (4.1%)
	Severe	4 (1.8%)
Musculoskeletal and connective tissue disorders (N=19)	Mild	9 (4.1%)

System Organ Class	Severity	CR845 (N=222)
	Moderate	9 (4.1%)
	Severe	1 (0.5%)
Nervous system disorders (N=39)	Mild	23 (10.4%)
	Moderate	16 (7.2%)
	Severe	0
Product issues (N=2)	Mild	1 (0.5%)
	Moderate	1 (0.5%)
	Severe	0
Psychiatric disorders (N=5)	Mild	0
	Moderate	5 (2.3%)
	Severe	0
Renal and urinary disorders (N=3)	Mild	2 (0.9%)
	Moderate	1 (0.5%)
	Severe	0
Respiratory, thoracic and mediastinal disorders (N=21)	Mild	12 (5.4%)
	Moderate	7 (3.2%)
	Severe	2 (0.9%)
Skin and subcutaneous tissue disorders (N=9)	Mild	3 (1.4%)
	Moderate	6 (2.7%)
	Severe	0
Surgical and medical procedures (N=1)	Mild	1 (0.5%)
	Moderate	0
	Severe	0
Vascular disorders (N=21)	Mild	10 (4.5%)
	Moderate	8 (3.6%)
	Severe	3 (1.4%)
Based on Table 3 in company response to clarification letter[5] MedDRA = Medical Dictionary for Regulatory Activities; TEAE = treatment-emergent adverse event		

The EAG appreciates the clarity and detail of this response.

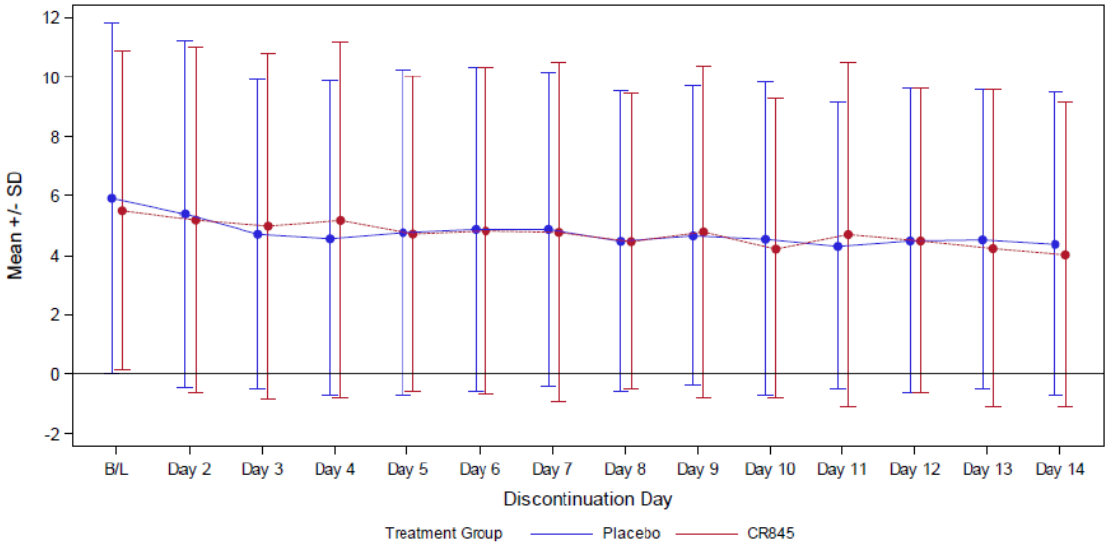
- The company was also asked to discuss AEs that lead to dose reductions, interruptions or discontinuation of difelikefalin treatment. In response, the company stated that, "*A total of 14 subjects (6.3%) experienced at least 1 TEAE that led to study drug discontinuation during the Treatment Period. The most common preferred term of TEAE leading to study drug discontinuation was somnolence (2 subjects [0.9%]). Four subjects (840008011, 840012004, 840018009, 840028005) experienced TEAEs leading to study drug discontinuation that were assessed as related to study drug. These subjects experienced the following study-drug related TEAEs: somnolence (840008011 and 840012004), nausea (840018009), and dizziness (840028005). A supplementary table has been provided detailing adverse events resulting in study drug discontinuation during the study period (please see Appendix U – AEs leading to discontinuation'. Of the 14 subjects who experienced a TEAE that led to study drug discontinuation, 2 subjects (348001001 and 840034003)*

had events with fatal outcomes, and the events for the remaining subjects were reported as recovered/resolved".[5] The supplementary table in Appendix U[24] has not been reproduced here for brevity. However, the two fatal events were enterocolitis-associated clostridium difficile infection and an unknown cause. The EAG accepts that the death due to enterocolitis-associated clostridium difficile infection was unlikely to be related to the study drug but does not understand how a death from unknown causes can be presumed to be not associated with the study drug.

- Table 3.50 presents the pooled adverse reactions results for the two double-blinded and the two open-label studies. The company was asked to specify how pooling was executed. The company stated that, "*The placebo-controlled cohort included participants (848) from the 12-week, pivotal studies (KALM-1 and KALM-2) who received (at least 1 dose of) IV difelikefalin at 0.5 mcg/kg or placebo 3 times per week. The all-difelikefalin-exposure cohort included all participants who received 1 or more doses of IV difelikefalin at 0.5 mcg/kg for up to 64 weeks from the placebo-controlled periods of the pivotal studies (if randomized to difelikefalin) and from the open-label extension periods (up to 52 weeks) of these studies. Safety was evaluated based on adverse events (AEs) and safety assessments (i.e., physical examinations, vital signs, clinical laboratory tests, and electrocardiograms). Safety analyses were summarized descriptively*."[5] The EAG is satisfied with this response.
- Section 4.7 of the Summary of Product Characteristics (SmPC) states that the study drug, "*has minor influence on the ability to drive and use machines.*"[2] As difelikefalin is approved for in-centre use only, the company was asked to provide more details on what measures will be put in place to manage dizziness/somnolence symptoms in patients who drive in for their thrice-weekly haemodialysis appointments, most especially within their first 3 weeks of treatment. The company responded by stating that, "*As per normal prescribing practice, with any medication associated with potential to cause dizziness or somnolence it is expected that clinicians would advise patients in the standard way until the effect of difelikefalin on the patient's ability to drive or operate machinery is known.*"[5] The EAG is satisfied with this response.
- Page 91 of the CS[1] states that potential signs and symptoms of opioid withdrawal were measured with the ShOWS and OOWS. The company was asked to supply the relevant data. The company stated that, "*Data on both ShOWS and OOWS scores for the Double-blind Discontinuation Population of KALM-1 is summarised below:*[5]
    - *ShOWS:*

Figure 6 presents mean total ShOWS scores over time for the Double-blind Discontinuation Population. Both treatment groups showed a slight decrease in mean ShOWS score over time. At baseline, subjects in the difelikefalin and placebo groups reported mean ShOWS scores of 5.5 and 5.9, respectively. On Discontinuation Day 14, difelikefalin and placebo subjects reported mean ShOWS scores of 4.0 and 4.4, respectively, with mean changes from baseline of -1.1 and -1.2, respectively. The largest LS mean treatment group difference in the change in ShOWS score from baseline was 0.9, which was observed at Discontinuation Day 4 and was significant (P = .044). No other treatment group difference in change in ShOWS score from baseline was significant for Discontinuation Days 1 through 6 or for Days >6."[5]

Figure 3.15: Total ShOWS score over time during the double-blind discontinuation



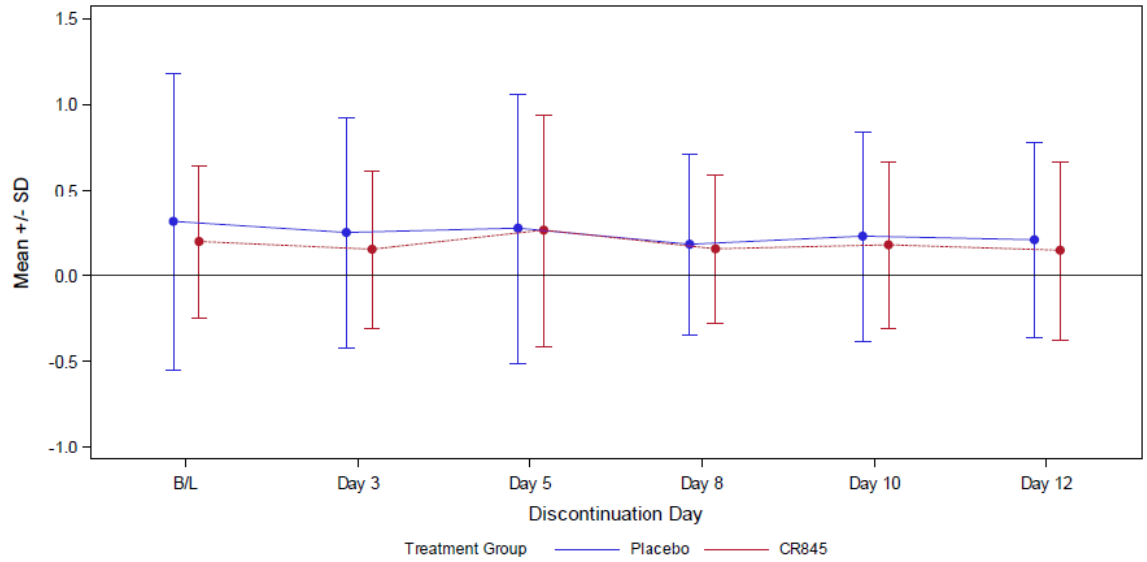
Based on Figure 3 in company response to clarification letter[5]
SD = standard deviation; ShOWS = Short Opiate Withdrawal Scale
Note: The mean and standard deviation for the total observed daily scores are displayed

- *"OOWS:*

Figure 7 presents the total OOWS score over time for the Double-Blind Discontinuation Population. Mean OOWS scores were generally stable over time in both treatment groups. At baseline, subjects in the difelikefalin and placebo groups reported mean OOWS scores of 0.2 and 0.3, respectively. On Discontinuation Day 12 (the last OOWS assessment), difelikefalin and placebo subjects reported mean OOWS scores of 0.1 and 0.2, respectively, with mean changes from baseline of -0.1 in both treatment groups. The largest LS mean treatment group difference in change in OOWS score from baseline was -0.1, which was observed at Discontinuation Day 3 (P = .255). No treatment group difference in change in OOWS score from baseline was significant for Discontinuation Day 3 through Discontinuation Days >6."[5]

- The EAG is satisfied with these responses.

Figure 3.16: Total OOWS score over time during the double-blind discontinuation period (no imputation) – line graph (population: double-blind discontinuation)



Based on Figure 4 in Company response to clarification [5]
OOWS = Objective Opiate Withdrawal Scale; SD = standard deviation
Note: Least squares means, SEs, and 95% CIs come from an ANCOVA model fit at each time point, with treatment group and baseline (Day 85) value as a covariate
Note: The mean and standard deviation for the total observed daily scores are displayed

3.2.6.2 Adverse events reported in Narita 2022

Adverse events were observed in 67% of patients in the placebo group, 72% in the 0.25 mcg/kg group, 77% in the 0.5 mcg/kg group and 85% in the 1.0 mcg/kg group. There appeared to be a dose-response relationship. Adverse events affecting the central nervous system (CNS) were more common at the highest dose. The authors reported that most AEs were mild and occurred early in drug administration. No deaths were reported. Serious AEs occurred in 5% of patients in the 0.25 mcg/kg group, 13% in the 0.5 mcg/kg group and 8% in the 1.0 mcg/kg group. Serious AEs that were thought to be related to study drug were altered state of consciousness in the 0.5 µg/kg of difelikefalin group and obstruction in the small intestine in the 1.0 µg/kg of difelikefalin group. All affected patients recovered after discontinuing or suspending use of the study drug.[13]

Table 3.54: Adverse events in Narita 2022

No. (%)	Placebo	0.25 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
AEs	42 (67)	44 (72)	47 (77)	53 (85)
Adverse drug reactions	7(11)	9 (15)	9 (15)	17(27)
Death	0	0	0	0
Other serious AEs	2(3)	3(5)	8(13)	5(8)
AEs leading to discontinuation	1(2)	0	4(7)	5(8)
AEs leading to interruption	2(3)	0	5(8)	5(8)

Based on Narita 2022[13]

No. (%)	Placebo	0.25 µg/kg of difelikefalin	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin
AEs = adverse events				

3.2.6.3 Adverse events reported in Fishbane 2020

The authors stated that most AEs were mild or moderate. These appeared to be unrelated to dose, but there was a tendency for difelikefalin to lead to higher rates of AEs than placebo (0.5 mcg/kg: 10/44; 1.0 mcg/kg: 6/41; 1.5 mcg/kg 11/44; placebo 4/45). The most prevalent AEs were abdominal pain (6.8%) in the 0.5 mcg/kg group and mental changes (6.8%) in the 1.5 mcg/kg group. Four deaths occurred: one due to respiratory failure (placebo group), two due to cardiac arrest (1.5 mcg/kg group), and one due to septic shock (0.5 mcg/kg group). These were not considered by the authors to be related to the study drug. Seventy eight percent of patients in the combined difelikefalin group and 42% in the placebo group reported TEAEs.[12]

Table 3.55: Adverse events in Fishbane 2020

No. (%)	Placebo	0.5 µg/kg of difelikefalin	1.0 µg/kg of difelikefalin	1.5 µg/kg of difelikefalin
Any TEAE	19 (42.2)	37 (84.1)	29 (70.7)	34 (77.3)
Any serious TEAE	4(8.9)	10 (22.7)	6 (14.6)	11(25)
TEAE leading to discontinuation	1(2.2)	4(9.1)	4(9.8)	7(15.9)
Deaths	1(2.2)	1(2.3)	0	2(4.5)
Based on Fishbane 2020[12] TEAE = treatment-emergent adverse event				

3.2.7 Included studies: supporting evidence

Not applicable.

3.2.8 Ongoing studies

According to the company, there are no ongoing studies that will provide additional evidence in the next 12 months for the indication being appraised.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Not applicable – no ITC was carried out, despite the clear need.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

Not applicable.

3.5 Additional work on clinical effectiveness undertaken by the EAG

Not applicable.

3.6 Conclusions of the clinical effectiveness section

The CS did not report a SLR of clinical effectiveness evidence and instead went straight into the reporting of two RCTs (KALM-1 and KALM-2) which had single-arm extension phases plus an additional single-arm study (CLIN3105).[1] In light of this, the volume of relevant clinical effectiveness

evidence that was omitted from the submission is uncertain as is the impact of this on clinical-and cost effectiveness estimates.

The KALM-1 and KALM-2 RCTs provided the clinical effectiveness data for this submission. The EAG identified two additional RCTs that appeared relevant but were not considered with the submission.[12, 13] The methodological quality of the RCTs was compromised by poor reporting of allocation concealment. The trials demonstrated that difelikefalin 0.5 mcg/kg was more effective than placebo for reducing itching intensity over an 8 or 12-week course of treatment. For QoL, results were more equivocal. Although the KALM-1 study showed a clear benefit over placebo for difelikefalin at 12 weeks, the KALM-2 study showed a more uncertain effect. Meta-analysis of the 8-week 5-D Itch scale score outcome in the Fishbane 2020[12] and Narita 2022[13] RCTs showed a similar level of uncertainty. However, for the 8-week Skindex-10 and Skindex-16 outcomes the Fishbane 2020[12] and Narita 2022[13] trials demonstrated clear benefits for difelikefalin. Longer term benefits on QoL were suggested by the KALM-1 and KALM-2 OLE data, which provided data to 52 weeks and 36 weeks respectively. However, these were one-arm extension studies and so the comparative longer term treatment effects are difficult to estimate. Adverse events were generally non-serious, and the rate of SAEs was similar across study arms.

These results need to be interpreted carefully. One major limitation of the evidence was that the trial data did not compare difelikefalin to ECM (with or without anti-itch medication), as had been requested by the NICE scope. The comparison of difelikefalin plus ECM to placebo plus ECM used in all four RCTs might lead to a spuriously more beneficial effect than the more stringent NICE scope comparison of difelikefalin versus ECM. No indirect treatment comparison was carried out by the company to rectify this issue.

There is a lack of rationale for the methods used in the statistical analysis for the primary outcome in KALM-1 and KALM-2. Decisions were made a priori in the trial protocols with little or no justification, while the validity of assumptions was not effectively tested within modelling.

Another limitation relates to a potential lack of applicability of trial findings to the UK target population. The trial sample had a larger proportion of Black participants than the UK target population. As Black participants were shown on a sub-group analysis to have a better response to difelikefalin than other ethnic groups, this discrepancy may lead to overestimation of the efficacy of difelikefalin in the UK target population. In addition, the company was unable to provide a sub-group analysis of specific anti-itch medications. Therefore, it is unclear if any possible differences between the specific anti-itch medications used in the trial and those used in the UK target population could also reduce the applicability of trial findings.

4. COST EFFECTIVENESS

4.1 EAG comment on company's review of cost effectiveness evidence

This Section pertains mainly to the review of cost effectiveness analysis (CEA) studies. However, the search Section (4.1.1) also contains summaries and critiques of other searches related to CE presented in the CS. Therefore, the following section includes searches for the CEA review, measurement, and evaluation of health effects as well as for cost and healthcare resource identification, measurement, and valuation.

4.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to CE presented in the CS.[1] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[41, 42] The EAG has presented only the major limitations of each search strategy in the report.

The document labelled 'Appendix D, G, H and I - SLR results' and a separate document containing the full search strategies provided details of two sets of searches: one to identify papers regarding treatment pathway and the second set was designed to identify utility evidence, cost and resource use and economic models in CKD-aP.[9, 43] The searches were conducted during April 2022.

A summary of the sources searched is provided in Tables 4.1 and 4.2.

Table 4.1: Data sources for economic evaluations: treatment pathway (as reported in the CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Embase.com	2012-2022/04	18.4.22
Embase	Embase.com	2012-2022/04	18.4.22
MEDLINE In-Process	PubMed	2012-2022/04	18.4.22
Cochrane Clinical Answers, Editorials, and Special Collections	Cochrane (Wiley)	2012-2022/04	18.4.22
CS = company submission			

Table 4.2: Data sources for economic evaluations (as reported in CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Embase.com	2012-2022/04	18.4.22
Embase	Embase.com	2012-2022/04	18.4.22
MEDLINE In-Process	PubMed	2012-2022/04	19.4.22
CENTRAL	Cochrane (Wiley)	2012-2022/04	19.4.22

Resource	Host/Source	Date Ranges	Dates searched
NHS EED	CRD	2012-2022/04*	19.4.22
Conference			
ISPOR		Last 3 years	26.4.22
Other search methods			
Bibliography of relevant reviews			
Citation snowballing			
Hand searching of country-specific websites for relevant objectives			
BAD	bad.org.uk		25.4.22
GIN	https://g-i-n.net/		25.4.22
HTA Organisations & additional economics resources			
NICE	http://www.nice.org.uk/		27.4.22
SMC	https://www.scottishmedicines.org.uk/search/		27.4.22
CADTH	https://www.cadth.ca/		27.4.22
IQWiG	https://www.iqwig.de/		27.4.22
ICER	https://icer-review.org/		27.4.22
PBAC	http://www.pbs.gov.au/pbs/home		27.4.22
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear2n/search/search.aspx		27.4.22
EQ-5D Publications Database	https://euroqol.org/search-for-eq-5d-publications/		27.4.22
ScHARRHUD	https://www.scharrhud.org/index.php?recordsN1&m=search&action=searchRecords		27.4.22

BAD = British Association of Dermatologists; CADTH = Canadian Agency for Drugs and Technologies in Health; CEA Registry = Cost-Effectiveness Analysis Registry; CENTRAL = Cochrane Central Register of Clinical Trials; CRD = Centre for Reviews and Dissemination; CS = company submission; EQ-5D = European Quality of Life-5 Dimensions; GIN = Guidelines International Network; ICER = Institute for Clinical and Economic Review; IQWIG = German Institute for Quality and Efficiency in Health Care; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NHS EED = National Health Service Economic Evaluation Database; NICE = National Institute for Health and Care Excellence; PBAC = Pharmaceutical Benefits Advisory Committee; ScHARRHUD = University of Sheffield Health Utilities Database; SMC = Scottish Medicines Consortium

*This is an archive resource so searches would have been made up to 31 March 2015

EAG comment:

- The CS and response to clarification questions provided sufficient details for the EAG to appraise the two sets of literature searches conducted to identify both treatment pathway and economic and HRQoL data on the treatment of CKD-aP. On the whole, searches were transparent and reproducible and a broad range of databases and grey literature sources were searched.

- Both the treatment pathway search and economic SLR report a single search strategy for both MEDLINE and Embase searches via Embase.com. At clarification the EAG asked the company to confirm whether this was a search of the Embase database conducted on the understanding that it now contains all records from MEDLINE or a multifile search where both resources were searched simultaneously using the same strategy. The company responded stating that "*Because Medline records are also indexed on Embase.com, a single search was performed covering both databases, to avoid duplication of records. It is understood that sometimes the MESH terms differ from EMTREE terms.*"[5] The company pointed out that an additional search of PubMed had been conducted in order to cover MEDLINE In-Process studies and as no limits to in process records appear to have been used this would also have covered any MEDLINE papers missed by the Embase search.
- The EAG queried the inclusion of the Cochrane Database of Systematic Reviews (CDSR) in a list of resources listed in Section 2.2.1 (document Appendix D, G, H, and I - SLR results)[9] as the search strategy document only contained strategies for searches of Cochrane Clinical Answers (clinical pathways search Table 3, document search strategy (SLR)), and a search of Cochrane Central Register of Clinical Trials (CENTRAL) in the economics searches. The EAG asked the company to clarify if this resource was included in error; the company explained that CDSR had been searched in order to identify relevant systematic reviews to validate the searches, rather than being searched as part of review.[5]
- The EAG noted that the Embase.com search strategy reported in Table 5 of the economic SLR appeared to contain several line combination errors.[9] At clarification the company confirmed that this had been caused by a reporting error and provided a correct copy of the strategy used.
- The EAG noted that the search strategy for the National Health Service Economic Evaluation Database (NHS EED) search reported in Section 2.2.1 (document 'Appendix D, G, H, and I - SLR results')[9] and in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart for the treatment pathway review (3.1, Appendix D, G, H and I)[9], was missing from the search strategy document; this was provided at clarification.
- The EAG queried a number of missing search strategies. Section 2.2.3: Other literature sources (document Appendix D, G, H, and I - SLR results) mentioned "*Hand searching of country-specific websites for relevant objectives*".[9] In addition to this, the summary protocol (document Appendix D, G, H, and I - SLR results, Section 10 Appendices) reported additional searches of Health Technology Assessment (HTA) bodies.[9] Details including dates searched and keywords used were provided at clarification. Hits retrieved were reported for the HTA searches but not for the searches of the British Association of Dermatologists (BAD) and Guidelines International Network (GIN) websites.[5]
- The EAG noted a discrepancy regarding the number of records retrieved by both the MEDLINE In-Process (via PubMed) and CENTRAL (via the Cochrane library) searches reported in the PRISMA diagram for the cost resource use and utility review and the corresponding search strategies recorded in the search strategy document (Tables 6 and 7)[43], the company clarified that "*The hits of CENTRAL and MEDLINE retrieved are correct and most recent as reported in PRISMA*".[5]

4.1.2 Inclusion/exclusion criteria

Table 4.3: Study eligibility criteria for the SLRs on treatment options, utilities, costs and resource use and economic modelling

	Description	Justification
Inclusion criteria		
Population	Adult patients with chronic kidney disease-associated pruritus (CKD-aP).	No justification given by the company, but this is the relevant population.
Interventions	Pharmacological interventions only.	No justification given by the company.
Outcomes	**Treatment options**, including recommended guidelines and patterns of care. **Utility outcomes**, including utility values, reported using European Quality of Life-5 Dimensions (EQ-5D), or other methods. **Cost and resource use**, outcomes not listed in the same way as the other systematic literature reviews (SLRs) but appear to be cost and resource data. **Economic outcomes**, including cost and burden studies and economic evaluations.	The company has effectively stratified the SLR into four separate review questions, which cover the outcomes to varying degrees. The first outcome, treatment options, is covered by the following review question: "*What is the standard of care for CKD-aP for treating adults in the UK and Europe?"*[9] This is described by the company as follows: "*A SLR of comparator treatments used in the UK and Europe to determine the appropriate positioning of difelikefalin in the treatment pathway for CKD-aP patients. This information will also support the design of the economic model to be submitted to NICE; highlighting the treatments used in current clinical practice for CKD-aP patients."*[9] The second outcome, utility outcomes, is covered by the following review question: "*What are the utility values for health states experienced by CKD-aP in the UK and Europe on current and emerging treatments?"*[9] This is described by the company as follows: "*A SLR to identify utility values for patients with CKD-aP. This may include outcomes for patients receiving or not receiving treatment for pruritus, and any disutility's associated with disease or treatment related adverse events."*[9] Economic outcomes are covered by two review questions: *1. "What are the direct and indirect costs of CKD-aP patients in the UK and Europe?"*[9] *2. "What are the economic model designs published or submitted for CKD-aP treatment in the UK and Europe?"*[9] These are described by the company, respectively, as follows:

	Description	Justification
		1. "A SLR to highlight the direct and indirect costs of treating CKD-aP patients in Europe."[9] *2. A SLR to highlight economic modelling studies of treating CKD-aP patients in Europe.*[9]
Study design	**Treatment pathway** Published treatment guidelines Relevant reviews and expert summaries **Utility evidence** Cohort studies/longitudinal studies (retrospective/prospective) Case-control studies Clinical trials reporting utility data Economic evaluations reporting utility data **Cost and resource use** Cohort studies/longitudinal studies (retrospective/prospective) Case-control studies Clinical trials reporting cost and resource use data **Economic modelling** Cost effectiveness and cost-utility studies Cost-benefit and cost minimisation analysis Budget impact models	No justification provided by the company.
Language restrictions	Studies published in English language or having English abstract or summary	No justification provided.
Publication timeframe	The last 10 years (2012-2022) to include the latest evidence	No justification provided.
Country	Global, i.e., no restriction with a focus on Europe (specifically the United Kingdom)	No justification provided.

	Description	Justification
Exclusion criteria		
Population	None reported.	Not applicable.
Interventions	None reported.	Not applicable.
Outcomes	None reported.	Not applicable.
Study design	None reported.	Not applicable.
Language restrictions	None reported.	Not applicable.

Based on Table 1, Appendix D, G, H and I -SLR results [9]
CKD-aP = chronic kidney disease-associated pruritus; EQ-5D = European Quality of Life-5 Dimensions; NICE = National Institute for Health and Care Excellence; SLR = systematic literature review

The description of the process used to apply the study selection criteria for the SLRs on treatment pathway/options, utilities, costs and resource use and economic modelling suggested that studies were selected in two stages: title and abstract screening; and full-text screening. Two reviewers screened all references independently during both stages. Disagreements were resolved by consulting a third reviewer.[9]

EAG comment:

- The population described in Table 1 ("*PICOS eligibility criteria*") of 'Appendix D, G, H and I – SLR Results' ("*Adult patients with chronic kidney disease-associated pruritus*")[9] differs to that shown in the NICE final scope ("*Adults with moderate-to-severe pruritus receiving haemodialysis*")[4] and the decision problem "*For the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis, including where established clinical management is insufficient in reducing pruritus*".[1] During the clarification phase the company was asked to explain the differences between these population characteristics. The company responded that, "*Given the limited quantity of published literature in this area the PICOS criteria was kept broad to ensure all relevant papers with useful evidence were captured.*"[5] The EAG is not satisfied with this response. Whilst it makes sense to use a broad range of search terms to maximise sensitivity of the search and ensure all relevant papers are captured, the study selection criteria should be focused to identify studies that address the review question(s). Applying a broader Population, Intervention, Comparison, Outcomes and Study (PICOS) for study selection risks inclusion of studies that are less relevant to the review objectives.
- The comparator in Table 1 ("*PICOS eligibility criteria*") of Appendix D, G, H and I – SLR results is described as "*Pharmacological interventions only*"[9] whereas the NICE final scope and the decision problem both list the comparator as "*Established clinical management without difelikefalin, including gabapentin and pregabalin*".[1, 4] In the clarification letter, the EAG asked the company to explain the differences between these comparator characteristics. The company's response was that "*The SLR description of 'Pharmacological interventions only' was deliberately broad so as to not restrict the results."* The company went on to say that *"The NICE Final Scope and Decision problem definition 'Established clinical management without Difelikefalin, including gabapentin and pregabalin' reflects the comparator definition as per the KALM-1 and KALM-2 trials. Established clinical management could also include 'no treatment'."*[5] The EAG response to this is similar to the above point i.e., that whilst it may be suitable to conduct maximally sensitive searches in order to avoid missing relevant evidence, this should be seen as a distinct process from study selection which should be guided by focused criteria aimed at identifying studies that will answer the review question(s).
- Restrictions according to date (2012 to 2022) and language (had at least an English language abstract available) were applied to the SLRs represented in Table 1 of Appendix D, G, H and I – SLR results.[9] During the clarification phase, the EAG asked the company to explain the impact of these restrictions on the results of all the SLRs. The company stated that, "*It was decided that date restrictions had a low impact on the omission of relevant information, and helped to provide the most recent evidence. Details have been provided of the 741 studies excluded prior to 2012 in an attached document…..It was deemed that language restrictions had a low impact. Details have been provided of the 12 non-English papers excluded*".[5] It was not feasible for the EAG to examine the Excel file listing the 741 pre-2012 studies in detail but the EAG believes that the 2012 date cut-off was arbitrary and may have led to the omission of relevant evidence. Perusal of the Excel file with the 12 papers excluded because of non-English language suggested that none were relevant.

- The company has carried out four different SLRs under one very broad protocol.[9] It is unclear to what extent this protocol was designed *pre-hoc* and whether any *post-hoc* changes were made that could have resulted in study selection bias or other types of bias.
    1. The first SLR, which attempts to determine the positioning of difelikefalin in the patient pathway, did not seek data comparing difelikefalin to comparator treatments but instead used guidelines and systematic reviews to evaluate comparator treatments only.[9] Therefore, this SLR does not directly relate to the comparison described in the NICE final scope: difelikefalin versus ECM.[4]
    2. The question for the second SLR is "*What are the utility values for health states experienced by CKD-aP in the UK and Europe on current and emerging treatments?*"[9] This is the only SLR out of the four described that compares difelikefalin to other treatments in terms of health-related outcomes. However, according to the information given in the review protocol (see Table 4.3 above), this SLR appears to exclude articles that do not report utilities. Therefore, studies containing useful information on clinical effectiveness might have been missed from the SLR.
    3. The third SLR ("*What are the direct and indirect costs of CKD-aP patients in the UK and Europe?*"[9] provides information about the costs of comparator treatments.
    4. The fourth SLR ("*What are the economic model designs published or submitted for CKD-aP treatment in the UK and Europe*")[9] is relevant to this submission. The SLR identified one study that utilised pooled data from the KALM-1 and KALM-2 trials.
- The process undertaken to apply the study selection criteria was satisfactory.

4.1.3 Critique of data extraction

'Appendix D, G, H and I – SLR Results' provided a description of the data extraction process suggesting that two reviewers extracted data independently and that disagreements were resolved by discussion and coming to a consensus.[9] The types of data to be extracted from each study were not described in the methods section although to some extent became apparent on further reading, from details in tables, figures and narrative description.

EAG comment:

- The data extraction process was satisfactory but more detail about the planned data extraction was needed. In light of the lack of information about this in the methods sections of the four SLRs, it is possible that the data extraction was done in an *ad hoc* fashion which may have resulted in bias (i.e., highlighting types of information that were more favourable for the submission rather than adhering to a pre-specified plan for data extraction).

4.1.4 Quality assessment

Quality assessment of the included studies was performed as outlined below:

- Utility review: the risk of bias assessment was performed using the checklists proposed by Papaioannou 2013[44] and NICE Decision Support Unit (DSU) Technical Support Document 9.[45] The description of these checklists included details of the methodology and the results of the scoring. The limitations of all data were the lack of information on loss to follow-up and missing data, but all data were deemed adequate for sample size, selection and recruitment, inclusion and exclusion criteria and the appropriateness of measures. The company considered that the data used in the utility review were of good quality overall.[9]

- Cost and resource use studies: the checklist adapted to cost of illness by Molinier 2008 was used to assess methodological quality.[46] The description of this checklist was accompanied by details of the methodology and results of the scoring. The results suggested that the data used in the utility review were of variable quality. Limitations common to more than half of the papers were insufficient disaggregation of direct and indirect costs, inadequate detail about the sources of cost values, inappropriate valuation of unit costs, and failure to test major assumptions in a sensitivity analysis.[9]
- Economic evaluations: the Drummond checklist was used.[47]

EAG comment:

- The company were asked to provide full details of the critical appraisal process during the clarification phase. The company responded by stating that, '*The critical appraisal conducted during the full-text review phase was assessed by two independent researchers. Three researchers were involved in this process, who each reviewed two-thirds of the articles. Any disagreement between the two researchers were resolved with the third researcher. Disagreements were discussed until a consensus formed over inclusion/quality assessment.*'. The company also stated that, '*The SLR followed a robust methodology that was fully compliant with PRISMA-P1 guidelines and meets the standards described by NICE. The SLR employed a standard two review process and quality control for evidence screening at first (Title/Abstract) and second stage (Full texts). Two investigators, working independently, extracted data for the study population of interest for the final list of selected eligible studies. Any discrepancies observed between the data extracted by the two data extractors were resolved by discussion and coming to a consensus. If a consensus was not formed, the third independent reviewer provided arbitration. Details have been provided in section 2.3.2 of Appendix D, G, H and I..'.* The EAG considers the critical appraisal process to be generally acceptable.
- No quality appraisal appears to have been carried out for the first SLR, which was aimed at determining the appropriate place of difelikefalin in the treatment pathway.[9] The company was asked about this during the clarification phase, and responded by stating that, "*The guideline reviews were not critically appraised as the guidelines are broad. Only a small portion of the guidelines were extracted as not all sections were relevant to CKD-aP, meaning a critical appraisal of the guidelines was not appropriate. Furthermore, the guidelines are mainly issued by national bodies concerned with CKD or pruritis specifically, rather than CKD-aP.*"[5] The EAG does not accept this as a reason to have omitted critical appraisal since the aspects of the guidelines relevant to the SLR could have been subjected to specific evaluation. Appropriate tools would have included the second version of the Appraisal of Guidelines for Research and Evaluation (AGREE II) checklist[48] and/or Risk of Bias in Systematic Reviews (ROBIS).[49]
- The company's overall view that the studies included in the utilities SLR were of good quality may have been overly optimistic. Loss to follow-up and missing data can impact on utility estimates and may have introduced bias. Therefore, the EAG feels that the overall conclusion should have been more cautious.
- The results of the quality appraisal for the economic evaluation SLR (Drummond checklist) are not reported.

4.1.5 Results of the SLRs

The results from each of the four SLRs are summarised below:

SLR 1: What is the standard of care for treating adults with CKD-aP in the UK and Europe?

The SLR included four treatment guidelines from UK, Germany, Canada and Australia and 12 review articles providing recommendations for the treatment and treatment patterns of CKD-aP followed across the real world. All guidelines emphasised adequate dialysis and use of topical anaesthetics and emollients as the first step for the management of pruritus. Gabapentin and pregabalin were the recommended agents by all the guidelines for the management of refractory CKD-aP. Selected antidepressants (e.g., doxepin, amitriptyline), UV therapy, and serotonin antagonists are also recommended for refractory CKD-aP. Similar recommendations were provided in the evidence from review articles.[9]

SLR 2: What are the utility values for health states experienced by adults with CKD-aP in the UK and Europe on current and emerging treatments?

Three studies provided utility values for the UK, US and Pakistan. Using both UK and US tariffs, a significant improvement in utility values was observed with the use of 0.5 mcg/kg difelikefalin at week 8 compared to baseline. A decrease in utility values was observed with increasing disease severity ranging from 0.744 for none to 0.595 for severe CKD-aP using data from the SHAREHD (not explained in full) database, and from 0.625 for none to 0.416 for severe CKD-aP when mapping from 5-D Itch scores collected in the phase III KALM-1 trial.[9]

SLR 3: What are the direct and indirect costs of treating adults with CKD-aP in the UK and Europe?

Seven studies were included detailing cost and resource use in CKD-aP patients. Two studies were conducted in the US, one study each was conducted in Italy, India, Saudi Arabia, Taiwan and globally. Three studies reported both direct cost and resource use, three studies reported only resource use and one study reported only direct cost of patients with CKD-aP. All studies were observational (four retrospective, two prospective and one cross-sectional). The sample size ranged from a minimum of six to 73,124 patients with CKD-aP. Hospital data or databases were the primary source of cost data in the studies. The results suggested that in Italy, the annual direct economic impact on the Italian National Health Service (INHS) due to the healthcare resource consumption was higher for a CKD-aP patient than for a non-CKD-aP patient or control (€37,065 versus €35,988 versus €31,286 respectively). Haemodialysis accounted for >60% of hospitalisation expenses and >77% of outpatient care costs for patients with CKD-aP on haemodialysis. Three studies (Global, Taiwan and US) revealed increased healthcare cost and rate of all-cause hospitalisations for patients who were extremely versus not at all bothered by itchy skin.[9]

SLR 4: What are the economic model designs published or submitted for CKD-aP treatment in the UK and Europe?

One study was included in the economic evaluation SLR discussing the methodological approach to assess the economic value of difelikefalin for the treatment of CKD-aP. Due to lack of economic evaluations evidence in CKD-aP, disease criteria were broadened to include the CKD-aP analogues. The exploratory research included six HTAs and 12 studies assessing economic evaluations in CKD-aP analogues. The estimated cost of dialysis was found to be £30,591 per patient in 2016/17 in a dapagliflozin CKD STA. The annual cost of dialysis was reported to be £32,360.41 whilst the monthly cost reached £2,696.70. A roxadustat CKD HTA reported a cost of £153.52 for haemodialysis. The health state cost of haemodialysis was reported as £461. For intrahepatic cholestasis, the cost per visit to a specialist facility per year on haemodialysis transport was £3,750 for 156 return journeys.[9]

A 1-year decision tree followed by a three state Markov model was used by two of the three atopic - dermatitis HTAs while the third HTA used a cohort Markov state model. A lifetime horizon and discounting of 3.5% was used in all models. A yearly cycle length was used in two models while the third, Baricitinib model, used a 4-week cycle length. Across the 12 modelling analogue studies included in the extrapolatory review, 10 studies used a Markov model, 11 studies assessed the atopic-dermatitis population and the majority of studies assessed tacrolimus (n=4) and dupilumab (n=4).[9]

EAG comment:

- Details of studies excluded at the full text screening stage were not provided as part of the original CS for any of the above reviews.[1, 9] In the clarification letter, the EAG asked the company to provide a list of excluded studies for each of the above SLRs, showing the full bibliographic details for studies excluded at the full text screening stage, together with reasons for exclusion. As part of their response, the company indicated that they had provided a list of excluded studies. On scrutiny, it emerged that details of excluded studies were provided for only two of the four SLRs (treatment pathway and economic evaluations). For the treatment pathway SLR, of 27 references identified, nine were excluded because the outcome was deemed not of interest. For the SLR of economic evaluations, 34 out of 52 identified references were excluded because the outcome was not of interest. A further two references were labelled as "*Review/editorial*" and it was not clear whether these provided any data. Four references were categorised as "*Relevant review*" but were not mentioned in the updated version of 'Appendix D, G, H and I – SLR Results' therefore it is unclear if or how they contributed data. There was no explanation regarding the lack of information on excluded studies for the SLRS on utilities and costs and resource use data.[5, 9]
- The utilities SLR (SLR 2) included a conference abstract[50] and poster,[51] both based on the KALM trials. The former was also included in the economic modelling SLR (SLR 4). However, these SLRs did not include the published report on the KALM-1 study,[16] nor the report on the pooled analysis of data from the KALM-1 and KALM-2 RCTs and OLE phases.[26] Articles reporting on the CLIN3105 study data were also not mentioned. The other two SLRs did not find any papers relevant to the current submission.
- Some details of study inclusion and exclusion were not clear in the above SLRs. Therefore, the risk of missing relevant evidence (and by implication, study selection bias) cannot be discounted.

4.1.6 Conclusions of the cost effectiveness review

The CS and response to clarification provided sufficient details for the EAG to appraise the literature searches conducted to identify economic and HRQoL data on treatment of CKD-aP. As the company did not identify economic evaluation evidence for CKD-aP, search criteria were extended to include CKD and pruritus independently. Searches covered the period from 2012 to April 2022. Searches were transparent and reproducible. MEDLINE, Embase, the Cochrane database and NHS EED were searched. A good range of databases and grey literature resources were further searched to identify further potentially relevant publications.

Overall, 16 studies were identified to provide recommendations for treatment of CKD-aP. Three studies to assess utility values of CKD-aP patients,[50, 52, 53] seven studies to provide cost and resource use data, and one evaluation assessing the economic value of difelikefalin for the treatment of CKD-aP.[50] The expanded economic review identified six NICE HTA assessments, and 12 studies reporting on CEAs in CKD and pruritus. Overall, the EAG has no major concerns about the literature searches conducted.

The PICO used to guide study selection for the four SLRs was broad and appeared to result in retrieval of some material that was not directly relevant to the submission. However, other aspects of study

eligibility risked missing relevant evidence e.g., the date restriction. The information about excluded studies was not completely clear and the risk of study selection bias could not be discounted. No critical appraisal of included studies was presented for the treatment pathway and economic evaluation SLRs. The company's rating of the quality of the studies included in the SLR of utilities may have been overly positive. In light of these methodological shortcomings, the EAG considers that the findings of the four SLRs should be viewed with caution.

4.2 Summary and critique of company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 4.4: NICE reference case checklist

Element of HTA	Reference case	EAG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	As per the reference case
Perspective on costs	NHS and PSS	As per the reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	As per the reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	As per the reference case
Synthesis of evidence on health effects	Based on systematic review	As per the reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	As per the reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	The 5-D Itch score was used to assess itch-related QoL in patients from the KALM-1 and KALM-2 trials. A mapping study was used to map the 5-D Itch to the EQ-5D.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The UK EQ-5D valuation tariff has been used.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	As per the reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	As per the reference case

Element of HTA	Reference case	EAG comment on CS
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	As per the reference case
CS = company submission; EQ-5D = European Quality of Life Five Dimension; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; HTA = Health Technology Assessment; HRQoL = health-related quality of life; PSS = Personal Social Services; QALY = quality-adjusted life year; QoL = quality of life; UK = United Kingdom		

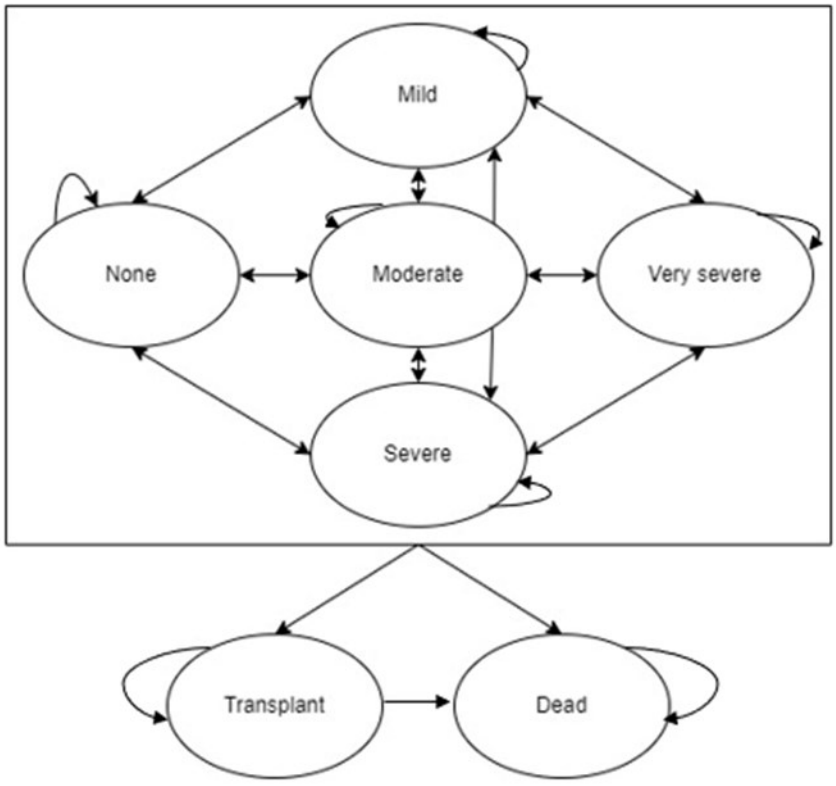
4.2.2 Model structure

4.2.2.1 Health states/events and transitions

The company developed a Markov model in Microsoft Excel® to assess the CE of difelikefalin for the treatment of adult patients with moderate-to-severe CKD-aP who are on haemodialysis.

The model consists of seven mutually exclusive health states as shown in Figure 4.1: 5 core health states defined by the level of itch severity (i.e., none, mild, moderate, severe, and very severe), the health states of renal transplant and death.

Figure 4.1: Model structure



Based on Figure 13 of the CS[1]
CS = company submission

The five core health states of the disease were defined using outcome measures of WI-NRS and 5-D Itch scale scores as these were collected in the KALM-1 and KALM-2 trials. The 5-D Itch scale was the company's preferred measure of treatment efficacy for base case analysis over the WI-NRS questionnaire, as the 5-D Itch scale was used to measure itch severity for up to 64-weeks whereas WI-NRS was only used as a measure of itch severity for the first 12-weeks. Patients in the model can get

renal transplant at any time point and from any of the five core active disease states. Renal transplant is assumed to be the last treatment option for CKD-aP,[54] with patients discontinuing CKD-aP treatment upon receipt of a transplant, whilst death is an absorbing health state.

In the base case analysis, patients with moderate-to-severe CKD-aP who are on haemodialysis enter the model in the respective active disease states. The baseline distribution of patients at model entry is based on the pooled data from the KALM-1 and KALM-2 trials as measured by the total score on 5-D Itch scale: 55.28% in the moderate health state; 34.17% in the severe and 10.55% in the very severe. Following an initial 'run-in' period which reflects short-term treatment decisions and initial response to treatment, patients responding to treatment with difelikefalin remain on treatment for the remainder of their lifetime, while non-responders to difelikefalin treatment continue with ECM only. The 'run-in' period consists of three cycles of 4 weeks length, whereas the cycle length for the long-term lifetime course of treatment for CKD-aP patients is 52 weeks (from cycle 4 onwards).

Costs and utilities are applied to each health state to calculate total costs and QALYs per model cycle. A half-cycle correction is implemented in the model only for the long-term treatment period of all 52-week cycles. The input values of the model and their underlying assumptions are further elaborated in the remaining part of Section 4 of the EAG report.

EAG comment: The EAG considers that the model structure adequately reflects clinical issues related to patients with moderate-to-severe CKD-aP who are on haemodialysis. Therefore, the model structure appears to be appropriate and fit for purpose.

4.2.3 Population

The company's decision problem explored the impact of difelikefalin use '*For the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis, including where established clinical management is insufficient in reducing pruritus*.' This is aligned with the EMA's recommendation to grant marketing authorisation for difelikefalin,[2] but deviates from the population in the NICE scope which was defined as '*Adults with moderate-to-severe pruritus receiving haemodialysis*', without restricting treatment for ICHD patients (see Section 2.1. for further details).

The baseline characteristics applied in the model for the target population, including the average patients' starting age, proportion of male population, weight, and length of time on haemodialysis are based on the KALM-1 and KALM-2 trial populations. Table 4.5 presents these patient population baseline characteristics as estimated by the KALM trials and the UKRR. Patients included in the economic model were assumed to have an average baseline age of 58.3 years, 58.7% to be male, to have a mean weight of 84.4 kg and a mean time on dialysis of 4.78 years, based on the KALM trials population characteristics (using 424 patients in difelikefalin and 424 patients the placebo arms based on B2 clarification question). The UKRR patient population characteristics in Table 4.5 are based on annual UKRR data of approximately 70,000 kidney patients on renal replacement therapy (RRT) in the UK, and were presented in the CS as a secondary option to inform the characteristics of the patient population.[55]

Table 4.5: Baseline characteristics of model population.

Characteristic	Pooled KALM-1 and KALM-2 trials[56]	UKRR[55]
Mean age (SD)	58.3 (12.8)	67.50

Characteristic	Pooled KALM-1 and KALM-2 trials[56]	UKRR[55]
Male (%)	58.7	62.10
Weight (Kg)	84.4 (21.5)	N/A
Time on dialysis (years)	4.78 (4.3)*	3.2[57]

Based on Table 41 of the CS[1]
CS = company submission; Kg = Kilogram; N/A = not applicable; SD = standard deviation; UKRR = UK Renal Registry
*Estimated from pooled KALM trial patient-level data set

EAG comment:

- As detailed in Section 2.1, the population in the company's analysis is narrower the population in the NICE final scope. The company justified the restriction to ICHD use in the decision problem, on the grounds that difelikefalin is restricted for ICHD use only.[2] However, the EAG considers that this deviation is small, because based on the 24th Annual Report of the UKRR, only 5.4% of the patients received home haemodialysis as per 31 December 2020.[58] See the EAG comments in Section 2.1 for a more thorough discussion about the population in the current submission in comparison to the population defined in the NICE final scope.
- The EAG noticed that baseline characteristics based on data from the KALM trials and the UKRR presented some differences, especially in terms of the starting age of the patients and the patient time on dialysis treatment (Table 4.5). Therefore, the EAG questioned if the KALM trial populations are representative of the UK population (Question B2 in the clarification letter).[5] The company responded that the KALM data are considered to be more appropriate for the base case analysis given that pruritus is not regularly registered in current UK clinical practice, deeming the UKRR data more representative of the wider haemodialysis population. Furthermore, in a scenario analysis the company showed that using the UKRR data had only a minor impact on the CE outcomes.[5]

4.2.4 Interventions and comparators

The intervention considered in the model was difelikefalin in combination with ECM. Difelikefalin is administered by IV bolus injection at the end of haemodialysis treatment at 0.5 mcg/kg dry body weight. Therefore, difelikefalin was approved for ICHD use only. Based on the UKRR report of 2019, which estimated that 5.5% of patients were undertaking ICHD less than three times per week, 92.7% exactly three times per week, and 1.8% more than three times per week, the model assumes a weighted frequency of 2.96 dialyses sessions per week (estimated assuming two and four sessions per week for those under and over three ICHD sessions per week).[55]

The modelled comparator is ECM including treatments that focus only on symptom management, given the fact that there are no approved treatments for CKD-aP in the UK. The CS in Section B.3.2.2 stated that the BAD currently recommends use of anti-itch medication such as capsaicin cream, topical calcipotriol, or oral gabapentin and advises against sedative antihistamines and cetirizine.[54] The most common anti-itch medications (in more than >2% of patients at baseline) used in the KALM trials as per company's clarification response (questions A20 and B5) were diphenhydramine (23.5% to 24.4%), hydroxyzine (9.9% to 12.2.%), hydrocortisone (2.6% to 3,8%), cetirizine (1.6% to 2.4%) and clemastine (1.6% to 2.4%).[5]

EAG comment: There seems to be a discrepancy between the drugs recommended by BAD and the ones used in the KALM trials for the ECM drugs. For instance, BAD advises against cetirizine, while as per company's clarification response (questions A20 and B5), cetirizine was used in 1.6% of patients in the DFK arm and 2.4% in the placebo arm. However, it is not clear to what extend UK clinicians adhere to these BAD recommendations. See for a further discussion on the comparability of ECM as used in the UK versus as used in the KALM trials the EAG comments in Section 3.2.1.1.

4.2.5 Perspective, time horizon and discounting

The economic analyses were conducted from the perspective of the NHS and PSS, in line with the NICE reference case.[10] The model has a time horizon of 42 years that is considered appropriate as a lifetime horizon, in line with the NICE reference case, given that the average age of patients at the start of treatment is 58 years. Costs and QALYs were discounted at 3.5% as per the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

4.2.6.1 Run-in phase

The 'run-in' period was used in the model to reflect initial response to treatment with difelikefalin and was implemented using a stopping rule for patients who did not respond at the end of the period. As non-responders were considered those patients experiencing side effects worse than itch and patients for whom treatment with difelikefalin did not improve itch outcomes by week 12 (i.e., the end of the run-in period).

Itch severity during the 'run-in' period was assessed using the WI-NRS scores ranging from 0 (no itching) to 10 (worst imaginable itching) and the 5-D Itch scale from patients in the KALM-1 and KALM-2 trials. The 5-D Itch scale estimated itch severity over the previous 2 weeks, covering five dimensions of itch, including the degree, duration of itch/day, direction (improvement/worsening), disability (impact on activities such as work), and body distribution of itch. The total 5-D Itch scale score ranges from 5 to 25, with higher scores indicating worse levels of itch. A 5-point reduction in the total 5-D Itch score from baseline was assumed to represent a clinically meaningful improvement in patients with CKD-aP undergoing haemodialysis based on the phase 2 CR845-CLIN2101 dataset,[32] which supported that a reduction of ≥3 points on the WI-NRS score represents a clinically meaningful change in pruritus in patients with CKD-aP. In the clarification phase, the company further explained that the phase 2 study demonstrated that the improvement in itch-related QoL measures were highly correlated with a reduction in the WI-NRS score at week 8, with a Pearson coefficient of 0.71 for the 5-D itch total scores (P = <0.0001). A 5-point reduction in 5-D Itch scale was associated with a 4-point reduction in WI-NRS scale, which, as indicated above, was defined as a clinically important reduction in the severity of CKD-aP patients.[32]

Table 4.6 presents the distribution of patients by the end of the 'run-in' period for the difelikefalin treatment arm, estimated based on the number of patients who achieved a clinically meaningful itch score improvement at week 12 divided by the total count of patients in the health state at week 12. A scenario analysis was presented by the company where the stopping rule is instead applied at week 8.

Table 4.6: Distribution of patients after the stopping rule applied, used in model base case.

Characteristic	None	Mild	Moderate	Severe	Very severe	Total
Patients at baseline (N)	0	0	224	129	40	**393**
Patients at week 12 (N)	75	126	161	20	11	**393**

Characteristic	None	Mild	Moderate	Severe	Very severe	Total
Patients at week 12 achieving clinically meaningful threshold (N)	69	88	54	4	0	**215**
Patients remaining on treatment after week 12 (%)	92.00%	69.84%	33.54%	20.00%	0.00%	**54.71%**

Based on Table 44 of the CS[1]
CS = company submission; N = number of patients

EAG comment: The EAG was concerned about the validity of the assumption on the 5-point reduction in the 5-D Itch scale as a clinically meaningful cut-off value and asked the company to provide further support on this cut-off value by using any other relevant appraisals/studies that used the same rating scale of itching. The company had not conducted a literature search in an attempt to identify studies or appraisals that used the 5-D Itch scale and therefore provided no additional evidence (Question B7 in the clarification letter).[5]

4.2.6.2 Transition probabilities and long-term extrapolations

Transition probabilities between CKD-aP severity categories for the difelikefalin and ECM arms were derived from the pooled KALM-1 and KALM-2 trial data using 4-weekly and 52-weekly transition count data.

To estimate transition probabilities between CKD-aP severity levels for each of the treatment arms, first methods of multiple imputation were used to address issues with missing data in the patient-level data set for the total 5-D Itch scale scores. The missing values were estimated based on treatment group, baseline itch score, and patient characteristics including age, sex, diabetes status, ESRD, length of haemodialysis, length of CKD-aP, and use of anti-itch medication at baseline. Table 4.7 shows that in total there were 393 observations at baseline for the difelikefalin arm and 403 for the ECM arm, with the respective numbers dropping to 330 and 359 at the end of the double-blind period.

Table 4.7: Number of data observations included in the analysis at each model cycle from KALM-1 and KALM-2.

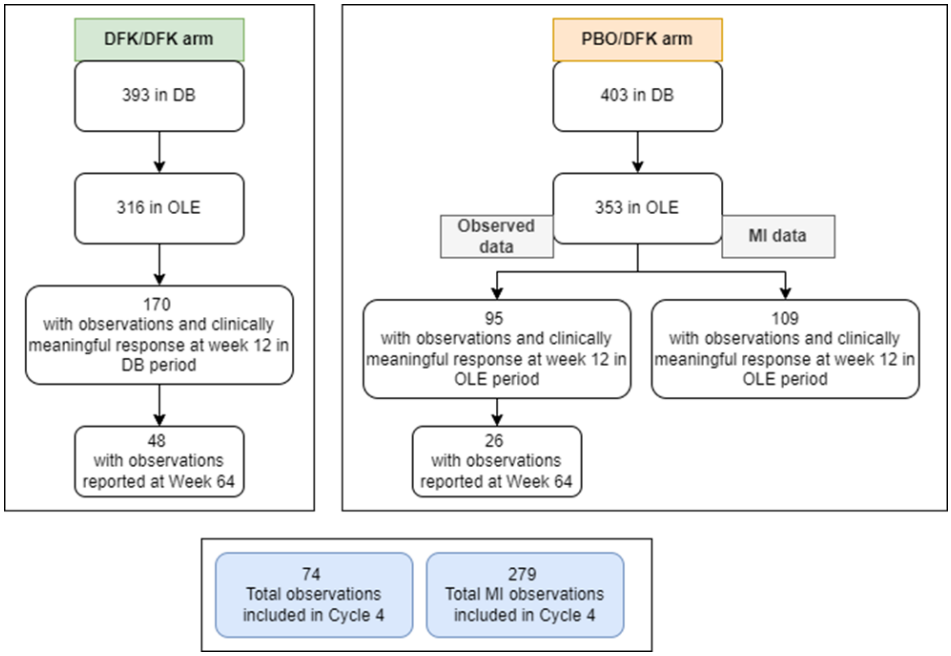
5-D Itch scale total scores	Difelikefalin		Placebo	
	Observed only	Missing data imputation	Observed only	Missing data imputation
Baseline	393	393	403	403
Cycle 1 (Baseline-week 4)	356	393	371	403
Cycle 2 (week 4-week 8)	333	393	357	403
Cycle 3 (week 8-week 12)	330	393	359	403
Cycle 4 (week 12-week 64)	74	279	N/A	N/A

Based on Table 46 of the CS[1]

5-D Itch scale total scores	Difelikefalin		Placebo	
	Observed only	Missing data imputation	Observed only	Missing data imputation
CS = company submission; N/A = not applicable				

In response to clarification response B11, the company further provided Figure 4.2 showing that the number of patients that entered the OLE phase were 316 and 353 in the difelikefalin and placebo arms, respectively.[5] Of the patients in the difelikefalin arm, 170 achieved a clinically meaningful response at week 12 of the double-blind period, whereas of the patients in the placebo arm that were switched to difelikefalin treatment in the OLE phase, 204 achieved a clinically meaningful response as measured in week 12 of the OLE period. The total number of observations at week 64 further dropped to 74 according to Figure 4.2 and Table 4.7, whereas the company indicated that the multiple imputation method was used for 279 observations, including the population that achieved a clinically meaningful treatment response at week 12 in the difelikefalin arm and that entered the OLE period, plus the population from the placebo arm who achieved a clinically meaningful treatment response at week 12 of the OLE period. However, despite the company's response to clarification question B11, the EAG is still unable to reproduce how the company reached to the 279 observations in cycle 4 for the multiple imputation method.

Figure 4.2: Observations included in Cycle 4 for observed and MI data set



Based on Figure 7 of the clarification letter.[5]
DB = double blind; DFK = difelikefalin; MI = multiple imputation; OLE = open label extension; PBO = placebo

Table 4.8 shows transition probability matrices between severity levels of itching that were used in the base case analysis and were estimated from a simulated data set. The simulation method used the mean change from baseline in itch scores by CKD-aP severity levels at baseline: moderate, severe or very severe (Table 47 of the CS). The mean change values were in turn added to the baseline itch scores to estimate the mean itch score for each of the corresponding weeks (week 4, week 8, week 12, and week 64). The simulated itch scores were further grouped into health states and used to estimate the change in health state from the previous cycle, weighted by the distribution of patients across the different severity itching health states at baseline. See Table 4.8 for an illustration of how this approach works for the transitions for week 4 in the difelikefalin group. To find the

probability of remaining in the same health state, the percentages from the column 'score distribution' corresponding to the zeros in column 'change in state' are summed together, so 11.3% + 10.9% + 11.1% = 33.3%. This value can be found in Table 4.8, in cycle 1, for the probability from mild to mild, moderate to moderate, severe to severe and very severe to very severe.

Table 4.8: Illustration of simulation method for transition probabilities

Baseline itch score	Baseline health state	Score distribution week 0	Change in itch score week 4	Week 4 itch score	Week 4 health state	Change in state
12	Moderate	4.5%	-2.9	9.1	Mild	-1
13	Moderate	8.0%	-2.9	10.1	Mild	-1
14	Moderate	9.4%	-2.9	11.1	Mild	-1
15	Moderate	11.3%	-2.9	12.1	Moderate	0
16	Moderate	10.9%	-2.9	13.1	Moderate	0
17	Moderate	11.1%	-2.9	14.1	Moderate	0
18	Severe	10.8%	-5.1	12.9	Moderate	-1
19	Severe	8.5%	-5.1	13.9	Moderate	-1
20	Severe	7.2%	-5.1	14.9	Moderate	-1
21	Severe	7.7%	-5.1	15.9	Moderate	-1
22	Very severe	5.0%	-6.1	15.9	Moderate	-2
23	Very severe	1.9%	-6.1	16.9	Moderate	-2
24	Very severe	2.6%	-6.1	17.9	Severe	-1
25	Very severe	1.0%	-6.1	18.9	Severe	-1

In estimating the transition probabilities, the company assumes that the transition probabilities to a better or worse health state in each cycle independent of the current health state. A result of this simulated approach is that patients will never switch to a worse health state in terms of itching (see Table 4.9). An additional assumption was that patients could switch to a maximum improvement or deterioration of three health states at a time to account for the fact that patients do not experience drastic changes to their itch state and to avoid extreme values.[5] This assumption was validated with KALM trial data in which there were no patients changing four health states in either treatment arm. However, Table 4.9 shows that based on the simulated data, patients always improve up to a maximum of two health states.

Table 4.9: Transition matrices used in the base case analysis based on simulated data

Before	After									
	Difelikefalin arm					ECM arm				
	None	Mild	Moderate	Severe	Very severe	None	Mild	Moderate	Severe	Very severe
Cycle 1 (baseline to week 4)										
None	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Mild	66.71%	33.29%	0.00%	0.00%	0.00%	49.25%	50.75%	0.00%	0.00%	0.00%
Moderate	6.91%	59.80%	33.29%	0.00%	0.00%	5.03%	44.22%	50.75%	0.00%	0.00%
Severe	0.00%	6.91%	59.80%	33.29%	0.00%	0.00%	5.03%	44.22%	50.75%	0.00%
Very severe	0.00%	0.00%	6.91%	59.80%	33.29%	0.00%	0.00%	5.03%	44.22%	50.75%
Cycle 2 (week 4 to week 8)										
None	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Mild	18.47%	81.53%	0.00%	0.00%	0.00%	9.92%	90.08%	0.00%	0.00%	0.00%
Moderate	0.00%	18.47%	81.53%	0.00%	0.00%	0.00%	9.92%	90.08%	0.00%	0.00%
Severe	0.00%	0.00%	18.47%	81.53%	0.00%	0.00%	0.00%	9.92%	90.08%	0.00%
Very severe	0.00%	0.00%	0.00%	18.47%	81.53%	0.00%	0.00%	0.00%	9.92%	90.08%
Cycle 3 (week 8 to week 12)										
None	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Mild	11.81%	88.19%	0.00%	0.00%	0.00%	9.42%	90.58%	0.00%	0.00%	0.00%
Moderate	0.00%	11.81%	88.19%	0.00%	0.00%	0.00%	9.42%	90.58%	0.00%	0.00%
Severe	0.00%	0.00%	11.81%	88.19%	0.00%	0.00%	0.00%	9.42%	90.58%	0.00%
Very severe	0.00%	0.00%	0.00%	11.81%	88.19%	0.00%	0.00%	0.00%	9.42%	90.58%
Cycle 4 (week 12 to week 64)										
None	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Mild	91.83%	8.17%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%

Before	After									
	Difelikefalin arm					ECM arm				
	None	Mild	Moderate	Severe	Very severe	None	Mild	Moderate	Severe	Very severe
Moderate	8.54%	83.29%	8.17%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
Severe	0.00%	8.54%	83.29%	8.17%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
Very severe	0.00%	0.00%	8.54%	83.29%	8.17%	0.00%	0.00%	0.00%	0.00%	100.00%
Cycle 5 (week 64 onwards)										
None	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%	0.00%	0.00%	0.00%
Mild	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%	0.00%	0.00%
Moderate	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%	0.00%
Severe	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%
Very severe	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%

Based on Table 48 and 49 of the CS[1]
CS = company submission; ECM = established clinical management

As a scenario analysis, the company also estimated transition matrices based on the patient-level data, see Table 4.10. In this approach, just as in the simulated approach, it was assumed that the transition probability is only dependent on how many health states the patient moves, and not on the health state the patient departs from. Thus, the probability to go from severe to moderate is the same as from moderate to mild. However, by using the patient-level data, also transitions to worst health states are allow.

Long-term efficacy estimates for the difelikefalin arm, i.e., transition probabilities from week 12 to week 64 in Table 4.9 above, were derived using data from the OLE phase of the KALM-1 and KALM-2 trials, consisting of both patients who had received difelikefalin and patients who had received placebo during the double-blind treatment period. Figure 4.3 presents results for the mean improvement in 5-D Itch scale total score from baseline across the double-blind treatment period and OLE phase, with and without the stopping rule applied.

Figure 4.3

Based on Figure 14 of the CS[1]
CS = company submission; DFK = difelikefalin; PBO = placebo

Treatment efficacy for difelikefalin was assumed to remain constant after week 64 (i.e., patients no longer changed health state) as no data was collected beyond the 52-week OLE phase. This effectively means that no treatment waning is anticipated; however, this was explored by the company in scenario analyses.

Given that all patients in the OLE phase of the KALM-1 and KALM-2 trials received difelikefalin, there were no data available to inform the long-term efficacy for patients in the ECM arm (i.e., transition probabilities from week 12 to week 64). To address this issue, the company considered three possible extrapolation methods for the long-term efficacy of the ECM arm, including the mean difference (MD), the ratio of means (RoM), and no change in efficacy. Figure 4.4 below shows the extrapolated change from baseline for each of these methods. The MD approach estimated the average of the MD between the difelikefalin and the placebo arms at each observation during the double-blind period and used this estimate to predict a mean change in 5-D Itch scale total score from baseline to week 64. This mean change value was then added to the baseline score of patients treated with placebo to simulate a 5-D Itch scale total score in week 64 for the placebo arm. The RoM approach estimated the average of the ratio of means between the placebo arm and the difelikefalin arm at each observation and multiplied this average ratio by the baseline score of patients treated with placebo to simulate a 5-D Itch scale total score in week 64 for the placebo arm. In the base case analysis, it was assumed that efficacy (mean change in itch from baseline) remains unchanged from week 12 onwards. This assumption was mainly

based on expert opinion sought by the company. It was noted that the placebo effect would wane over time in line with the natural progression of the disease, which is likely to get worse over time. The base case analysis further applied a waning effect in the ECM arm equal to a 5% probability for all patients to switch to the next worse health state each year following week 64, which was based on clinical expert opinion.

Figure 4.4: [redacted]

Based on Figure 15 of the CS[1]
CS = company submission; MD = mean difference; PBO = placebo; RoM = ratio of means

EAG comment:

- The company stated that estimating transition probabilities of moving from any one state to each of the other states could lead to unrealistic outcomes due to the potentially small numbers of observation for each probability value.[5] On the other hand, by assuming that the probability of improving or deteriorating CKD-aP in each cycle is equal and independent of health state (i.e. the current approach estimating transition probabilities in the model) implies that the rate of response to treatment is averaged across the population. Therefore, in clarification question B12 the EAG requested the transition probabilities to be estimated directly from the patient-level data allowing for a comparison with those currently used in the model. In the clarification response, the company provided these transition matrices and showed that using transition probabilities of moving from any one state to each of the other states in the model had a relatively small impact on the CE outcomes (question B12 in the clarification letter) compared to the company base case.[5]
- According to the CS, the currently implemented transition matrices estimating the per-cycle probabilities of losing or gaining health from zero to three health states can be estimated either using the simulated data or using observed data. However, transition matrices based on observed data were not presented in the CS, as it was argued that the simulated data set "*offer a better reflection of the underlying trend in the data and a more appropriate quantification of the uncertainty in the mean change in itch scores through probabilistic analysis"*, deeming the simulated data set a better approach for quantifying transition probabilities than the observed dataset. For comparison, the EAG extracted the respective transition probabilities based on observed data from the CE model and presents them in Table 4.10. Compared to the simulated data, the observed data allow patients to switch to both better and worse health states in terms of itching (compared to Table 4.9). Transitions based on observed data also allow patients to switch to a

maximum improvement or deterioration of three health states at time compared to simulated data allowing for a maximum improvement of two health states.

- However, when these matrices are compared to those provided by the company in response to question B12 in the clarification letter, it becomes clear that the idea that the probability of moving one state down does not depend on the current health state is not supported by the data. Hence, the EAG considers it best to not rely on any assumptions and use the transition probabilities as derived from the patient-level data.
- It should be noted that one of the arguments in favour of using the simulated approach, according to the company, was that '*when using the directly observed data, a Dirichlet distribution is required to generate a fully probabilistic transition matrix which may be associated with increased uncertainty'*. The EAG does not agree with this point. The Dirichlet distribution is indeed an appropriate method to propagate uncertainty of the transition matrices in this case, but it is not by itself imposing further uncertainty to model outcomes. Uncertainty in probabilistic estimates may be larger than expected, but that will be owing to the nature of the data.
- It is unclear to the EAG how all transition matrices were derived in light of the multiple imputation applied to account for missing data. For example, when looking at the directly estimated transition probabilities as presented in response to the clarification letter, it is unclear if these probabilities are based on averages over 20 different probabilities, each from a different complete dataset. It is not transparent how analyses per complete dataset were combined in order to find the final estimates, and how uncertainty was estimated (which should be a function of within dataset variation and between dataset variation).
- In the base case, a treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. The company argued that the waning effect in the ECM arm was applied in the model to mitigate the long-term impact of the improved outcomes observed for the placebo arm in the KALM trials, an assumption that lacked face validity according to clinical experts (Question B9 in the clarification letter).[5] That means that itch improvement for the ECM group based on the placebo arms of the KALM trials may overestimate the true impact of ECM arm. However, the company acknowledged that similar trial effects could have also impacted the impact of the difelikefalin arm. In absence of further real-world evidence to support the waning impact of the ECM, the EAG considers this assumption uncertain and removed it from the base case analysis. However, different waning patterns were explored in the EAG's scenario analyses for both the difelikefalin and ECM arms.

Table 4.10: Transition matrices based on patient-level change in state data

Before	After									
	Difelikefalin arm					ECM arm				
	None	Mild	Moderate	Severe	Very severe	None	Mild	Moderate	Severe	Very severe
Cycle 1 (baseline to week 4)										
None	97.96%	2.04%	0.00%	0.00%	0.00%	92.06%	6.45%	1.49%	0.00%	0.00%
Mild	59.80%	38.17%	2.04%	0.00%	0.00%	47.89%	44.17%	6.45%	1.49%	0.00%
Moderate	19.59%	40.20%	38.17%	2.04%	0.00%	14.14%	33.75%	44.17%	6.45%	1.49%
Severe	3.05%	16.54%	40.20%	38.17%	2.04%	0.99%	13.15%	33.75%	44.17%	7.94%
Very severe	0.00%	3.05%	16.54%	40.20%	40.20%	0.00%	0.99%	13.15%	33.75%	52.11%
Cycle 2 (week 4 to week 8)										
None	83.97%	14.50%	1.02%	0.51%	0.00%	86.10%	11.66%	1.99%	0.25%	0.00%
Mild	28.50%	55.47%	14.50%	1.02%	0.51%	30.02%	56.08%	11.66%	1.99%	0.25%
Moderate	3.31%	25.19%	55.47%	14.50%	1.53%	4.22%	25.81%	56.08%	11.66%	2.23%
Severe	0.51%	2.80%	25.19%	55.47%	16.03%	0.50%	3.72%	25.81%	56.08%	13.90%
Very severe	0.00%	0.51%	2.80%	25.19%	71.50%	0.00%	0.50%	3.72%	25.81%	69.98%
Cycle 3 (week 8 to week 12)										
None	84.48%	15.01%	0.51%	0.00%	0.00%	81.89%	17.12%	0.99%	0.00%	0.00%
Mild	27.74%	56.74%	15.01%	0.51%	0.00%	26.05%	55.83%	17.12%	0.99%	0.00%
Moderate	4.83%	22.90%	56.74%	15.01%	0.51%	2.48%	23.57%	55.83%	17.12%	0.99%
Severe	1.02%	3.82%	22.90%	56.74%	15.52%	0.99%	1.49%	23.57%	55.83%	18.11%
Very severe	0.00%	1.02%	3.82%	22.90%	72.26%	0.00%	0.99%	1.49%	23.57%	73.95%
Cycle 4 (week 12 to week 64)*										
None	93.19%	5.02%	1.43%	0.36%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Mild	60.22%	32.97%	5.02%	1.43%	0.36%	0.00%	100.00%	0.00%	0.00%	0.00%

Before	After									
	Difelikefalin arm					ECM arm				
	None	Mild	Moderate	Severe	Very severe	None	Mild	Moderate	Severe	Very severe
Moderate	24.73%	35.48%	32.97%	5.02%	1.79%	0.00%	0.00%	100.00%	0.00%	0.00%
Severe	3.94%	20.79%	35.48%	32.97%	6.81%	0.00%	0.00%	0.00%	100.00%	0.00%
Very severe	0.00%	3.94%	20.79%	35.48%	39.78%	0.00%	0.00%	0.00%	0.00%	100.00%
Cycle 5 (week 64 onwards)*										
None	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%	0.00%	0.00%	0.00%
Mild	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%	0.00%	0.00%
Moderate	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%	0.00%
Severe	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	95.00%	5.00%
Very severe	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%

Based on the CE model submitted by the company[59]
CE = cost effectiveness; ECM = established clinical management
*Transition probabilities for the ECM arm in cycle 4 and for both arms in cycle 5 onwards are based on assumptions

4.2.6.3 Mortality and transplant rates

The model uses annual transplant probability for patients who are alive on haemodialysis and annual mortality probability for patients who are alive on haemodialysis and had not had a transplant. In the base case, both mortality and transplant rates were modelled using the methods adopted in the CEA of haemodiafiltration versus high flux haemodialysis, as presented in NICE guideline for RRT and conservative management (NG107).[60] Annual probabilities of death and transplant were estimated for patients on haemodialysis from 1 to 10 years after initiating dialysis, whilst the annual probability of death was assumed to remain unchanged after year 10 (see Table 50 of the CS). [1] These probabilities in the NG107 CEA were informed by UKRR data from January 2005 to December 2014 and were considered representative given an updated review on mortality rates from 2015 to 2019. The probability of death for patients who received transplant was assumed to be the same as for the general population.

Furthermore, in the base case analysis, an increased mortality risk was applied for patients in the moderate, severe and very severe health states of the CKD-aP population using respective hazard ratios (HRs) of 1.11, 1.02, and 1.24 extracted from the study of Sukul 2021.[25] Sukul 2021 report the adjusted all-cause mortality HR for patients extremely, very much, and moderately bothered by pruritus compared with patients without pruritus, while accounting for potential confounders.

Background mortality was estimated using UK life tables from the Office for National Statistics (ONS).

EAG comment:

- The EAG has concerns around the elevated risk of death assumed in the model for patients in the moderate, severe and very severe health states of the CKD-aP population based on the study by Sukul 2021.[25] Sukul 2021 finds that extreme pruritus is an independent predictor of all-cause and case-specific mortality after controlling for multiple confounders such as patient demographic and clinical characteristics. At the same time, they acknowledge that the possible bidirectionality of the relationship between pruritus and cross-sectional PROs (such as depression, missed dialysis sessions, and poor sleep quality) limits the inferences that can be made and does not allow conclusions about cause-effect relationships. Additionally, the authors state that one cannot make inferences in how changes in pruritus severity may relate to the outcomes in their study.
- In response to the clarification letter, the company explained that increased depression, missed dialysis sessions, poor sleep quality, and skin lesions susceptible to infection are outcomes that could mediate the relationship between extreme pruritis and mortality.[5] The EAG appreciates the company's effort to justify a potential association between pruritus and mortality and has been convinced that a causal relation could exist. However, the evidence currently presented are not substantial enough to establish a causal relationship between a reduction in itching score due to treatment with difelikefalin and a reduction in mortality of these patients. Therefore, the EAG considers that this impact should not be part of the base case computations and instead explored in a scenario analysis.

4.2.6.4 Adverse Events

All commonly reported TEAEs that occurred in ≥2% of participants in the difelikefalin group and with ≥1% higher incidence than the placebo arm were included in the model based on the study by Fishbane 2022.[56] These according to the CS were diarrhoea, dizziness, nausea, gait disturbance including falls, hyperkalaemia, headache, and somnolence. Table 51 of the CS shows the event rates for these AEs, and the derived annual probabilities.[1] From this it can be seen that for all included AEs the probability of occurrence is quite similar in both groups.

Therefore, the treatment-specific cumulative incidences of AEs over the trial duration are used to inform the probability of AEs occurring for each treatment arm, whilst no utility decrement for AEs was considered in the base case analysis (see Section 4.2.7.3).

EAG comment:

- The annual probabilities of AEs included in the model were estimated for the EMC group from the double-blind phase of KALM-1 and KALM-2 whilst for the difelikefalin group data from the double-blind phase was pooled with data from the extension study. The EAG considers this a flawed comparison, as randomisation is broken by this procedure, and potential time effects in the occurrence of AEs can bias the comparison of the annual probabilities. Table 4.11, derived from Fishbane 2022,[56] shows that when only AEs from the double blind phase are compared difelikefalin leads to more AEs.

Table 4.11: Event rates for most common AEs from KALM-1 and KALM-2

	Placebo-controlled cohort				All difelikefalin exposure cohort	
	Placebo n=424 101.1 PY		Difelikefalin n=424 98.0 PY		Difelikefalin N=1,306 811.3 PY	
Commonly reported TEAEs	**n (%)**	**IR/ 1,000 PY**	**n (%)**	**IR/ 1,000 PY**	**n (%)**	**IR/ 1,000 PY**
Diarrhoea	24 (5.7)	267.2	38 (9.0)	469.2	158 (12.1)	266.2
Dizziness	16 (3.8)	188.0	29 (6.8)	316.2	103 (7.9)	151.6
Nausea	19 (4.5)	207.8	28 (6.6)	326.4	147 (11.3)	225.6
Gait disturbances	23 (5.4)	237.5	28 (6.6)	336.6	152 (11.6)	267.5
Hyper-kalaemia	15 (3.5)	158.3	20 (4.7)	234.6	108 (8.3)	157.8
Headache	11 (2.6)	118.7	19 (4.5)	214.2	78 (6.0)	106.0
Somnolence	10 (2.4)	98.9	18 (4.2)	204.0	29 (2.2)	39.4

Based on Fishbane 2022[56]
AEs = adverse events; TEAEs = treatment emergent adverse events

4.2.7 Health-related quality of life

4.2.7.1 Health state utilities

As no generic preference-based measures of health were collected in the KALM-1 or KALM-2 trials, a separate primary data collection study across UK dialysis centres was undertaken to develop a mapping algorithm relating the WI-NRS and 5-D Itch scale to the European Quality of Life-5 Dimensions 3 Levels (EQ-5D-3L) (Appendix J, mapping study).[61]

Full details of the mapping study can be found in Appendix J of the CS. In summary, primary data collection was undertaken between November 2020 and June 2021 across five sites in England on adult patients (18+ years of age) who had been receiving haemodialysis for at least 3 months. The data collected was used to estimate EQ-5D-3L mapping functions from 5-D Itch scale scores, WI-NRS, and 5-D Itch scale scores and WI-NRS combined. All mapping functions included age, sex, diabetes status,

and length of time on dialysis as additional conditioning variables. Despite limitations with missing observations, the 5-D Itch scale score to EQ-5D-3L mapping algorithm was considered the most appropriate option, given the paucity of published data in CKD-aP.

Table 4.12 provides the results of the EQ-5D-3L predictions based on the KALM-1 and KALM-2 data and using the 5-D Itch scale mapping algorithm. In the mapping study, the severe and very severe (unbearable) populations were merged, given the small numbers of observations in each group. In the base case analysis, the utility scores for the severe and very severe populations are set to be equal.

Table 4.12: Health state utility inputs used in the model

	Mean	SE	Source
Health state utility values			
None	[redacted]	[redacted]	Mapping study (Appendix J)[61]
Mild	[redacted]	[redacted]	
Moderate	[redacted]	[redacted]	
Severe	[redacted]	[redacted]	
Very severe	[redacted]	[redacted]	
Kidney transplant	0.71	0.04	TA775 referencing (Lee 2005)[62]
Based on Table 53 of the CS[1] CS = company submission; SE = standard error			

The SLR for utilities did not identify values of HRQoL for transplant. Instead, NICE HTAs identified in the expanded SLR were reviewed. Health state utility values (HSUVs) for kidney transplant were informed by Lee 2005,[62] which was identified in NICE TA775.[63] None of the other HTAs reported kidney transplant utility.

EAG comment:

- According to the NICE methods guide, mapping a disease specific QoL instrument to the EQ-5D is an appropriate approach if no direct EQ-5D data is available. The mapping study that was done to inform the health states utilities is of high quality, based on state-of-the-art methodology.

4.2.7.2 Utilities from the literature

An SLR was undertaken in April-July 2022 to identify HRQoL data for CKD-aP, with a particular focus on the UK and Europe. Full details of the SLR search strategy, study selection process, and results are presented in Appendix G. The review identified three studies providing utility evidence.

The study by Thokala 2021 presents the results of the mapping study discussed above, as well as utility values estimated from the SHAREHD database.[64] Table 4.13 presents these utilities. Note that all these utility are higher than those derived from the mapping study. The impact of using these utility values from the SHAREHD database was explored by the company through a scenario analysis, the results are presented in Section 5.

Table 4.13: Health state utilities as found in SHAREHD

CKD-aP health state	Mean
None	0.744

Mild	0.726
Moderate	0.589
Severe	0.595
Very severe	0.595
Based on Thokala et al. (2021).[64] CKD-aP = chronic kidney disease-associated pruritus; ISPOR = International Society for Pharmacoeconomics and Outcomes Research	

The study by Schaufler 2019 assessed different approaches to mapping individual questions and their rating from the Skindex-10 and the 5-D Itch collected in the Phase 2 difelikefalin trial (CLIN2101) to the five dimensions of the EQ-5D.[51] In the poster that presents these results, no distinction was made according to health state.

The study by Rehman 2021 investigated the effectiveness of zolpidem and acupressure therapy on food acupoints in improving the sleep quality and overall QoL of haemodialysis patients with CKD-aP in Pakistan.[53] Included were adult patients with affected sleep quality, and not on any medication to treat pruritus or sleep. The study observed a numerical improvement in the mean EQ-5D index score in the control group, from 0.49 (±0.30) at baseline to 0.53 (±0.30) at week 8 (p=0.187). In this study, no distinction was made according to health states based on 5-D Itch scores.

EAG comment: Of the discussed studies found in literature, only the study by Thokala 2021 provides alternative utility values for the model health states. These utilities are based on the SHAREHD study, which is a UK randomised study in patients receiving centre-based haemodialysis. However, no further information was provided, such as number of patients available to estimate the utilities and the related standard errors.

4.2.7.3 Adverse event disutilities

The company assumed that health state utility scores reported would include any disutility associated with AEs. Furthermore, they argued, the incremental incidence of AEs reported in Fishbane 2022 for the results of the pooled KALM-1 and KALM-2 trials were small and in general lower in those patients treated with DFK, suggesting that observed AEs are likely to be a feature of underlying disease.[56] As such, any utility decrements associated with their incidence will be implicitly captured in HSUV. To avoid the risk of double counting QALY loss, the company choose to set the disutility to zero for all AEs in the base case analysis.

EAG comment:

- The EAG considers it unlikely that the disutility of AEs are implicitly included in the HSUV, as the latter are derived from the 5-D Itch score. This instrument only contains questions specific to itching and is unlikely to capture AEs such as nausea and dizziness. Also, in the mapping study, where patients filled out both the 5-D Itch score and the EQ-5D, patients were not treated with difelikefalin, so the AEs cannot have been captured with the EQ-5D.
- At clarification, the company provided disutilities for AEs (Table 9, response to clarification letter). Furthermore, the Committee for Medicinal Products for Human Use (CMHP) assessment report (page 68)[65] showed the median duration for some AEs, for somnolence this was 20-30 days, and for the other AEs between 1 and 5 days. Adjusting the disutilities for duration leads to such small loss of QALYs that these have a negligible impact on the ICER.

4.2.8 Resources and costs

The following cost categories were included in the model: drug acquisition costs, drug administration costs for intravenously administered drugs, disease management costs, and costs of treatment-related AEs.

4.2.8.1 Drug acquisition and administration costs

The company estimated the annual cost of the difelikefalin treatment based on the resource use of difelikefalin injection volumes by estimating the number of vials per average patient weight[56] and frequency of ICHD.[55]

The number of difelikefalin vials used were derived from the European Medicines Agency (EMA) recommendation on difelikefalin vials by weight band per the SmPC. The estimation of the dose of difelikefalin was 0.5 mcg/kg dry body weight. The company calculated the total dose volume (mL) required from the vial as such: 0.01 × dry body weight (kg), rounded to the nearest tenth (0.1 mL).[2] To account for wastage, the number of vials per patient were rounded upwards. The average patient weight is given by a base case physical weight (kg) which was based on a pooled average physical weight of the trial populations. The number of vials required per treatment session is derived from the base case mean weight, 84.4 kg (see also Table 4.14).

Table 4.14: Difelikefalin dose, injection volumes and required number of vials

Drug	Weight range (kg) <lower bound>	Injection volume (ml)	Vials required
DFK	40	0.40	1
	45	0.50	1
	55	0.60	1
	65	0.70	1
	75	0.80	1
	85	0.90	1
	95	1.00	1
	105	1.10	2
	115	1.20	2
	125	1.30	2

Based on from CS, Table 54[1]
CS = company submission; DFK = difelikefalin; Kg = kilograms; ml = millilitre
Notes = number of vials required is rounded to account for wastage

The unit cost per vial of difelikefalin presented was the drug price following a Patient Access Scheme (PAS) reduction (or list price). The unit cost was multiplied by the annual number of required vials (derived from the estimated average number of vials required per patient and the weighted frequency of 2.96 ICHD sessions per week), thus deriving an annual cost of the difelikefalin treatment, £5,392.66 at list price, and £[redacted] at PAS price. The company therefore assumes a constant resource use throughout the year.

The company implemented treatment discontinuation as a stopping rule for difelikefalin patients who do not achieve a clinically significant improvement in itch at 12 weeks of treatment. The discontinued patients are assumed to continue with the same progress as patients in the ECM arm for the remainder of the time horizon. The model therefore subjects the discontinued patients to ECM treatment costs.

Treatment with ECM only serves as the comparison to difelikefalin combined with ECM treatment and includes the anti-itch medication consumed by CKD-aP patients. Proportions of patients consuming CKD-aP treatments per health state was based on the KALM trials.[56] The anti-itch medication dose, pack size, and price were based on the British National Formulary (BNF),[66] see Table 4.15. The company does not distinguish between the health states with regards to anti-itch medication doses and assumes the anti-itch medication dose is consumed every day of the week. The company estimated the weekly cost per anti-itch medication by dividing the required dose per week by the pack size of the medication multiplied by the pack unit cost. Upon the EAG request, the company clarified the definitions of the dose input parameters in Table 55 as daily doses and pack unit costs.[1] In this process the company found calculation errors in the ECM drug estimations and updated the model to the accurate calculation. All medications listed were weighted against the frequency at which they are consumed within each health state and the proportion of overall consumption by the specific health state.

The company observed that costs for the moderate health state were lower than the 'mild' and 'none' health states. The company deemed this unlikely, and they suggested that the proportion of moderate CKD-aP patients using anti-depressives in the study is lower than in clinical practice, causing the observed low cost. The company further stated that clinical guidance suggests the anti-itch treatment is additive between health states. For these reasons, the company assumes moderate health state total treatment cost as equal to mild health state total treatment cost by setting the resource use of the moderate health state to equal to the mild health state. Given the low number of observations in the health states for severe and very severe in the KALM trials, these two patient groups were subsequently merged in the mapping study. In the company base case, the company set the ECM medication resource use of the severe and very severe health states as equal (see Table 4.13).

EAG comment: The EAG had concerns regarding the assumption that the moderate CKD-aP patients on average incur the same anti-itch treatment costs as mild CKD-aP patients, but the company showed several scenarios using alternative assumptions (moderate health state set to severe health state and moderate health state set to average between mild and severe health state) that had negligible impact on the ICER. The EAG asked the company to clarify which health state resource use was used as the input for the health state. The company clarified the resource use of very severe patient group was set to equal to the severe group. While the resource use of the most severe patient group might be expected to be higher, the EAG found increasing the resource use had low impact on the ICER.

Table 4.15: Established clinical management resource use and treatment costs

Drug	None	Mild	Moderate	Severe	Very severe	Dose (per day)	Pack size	Pack cost	Weekly cost	Source
Receiving anti-pruritic medication	**40.20%**	**38.50%**	**36.60%**	**55.60%**	**55.60%**					
Topical corticosteroids	2.40%	3.40%	1.60%	7.40%	7.40%	1.50	15	£1.26	£0.88	BNF; hydrocortisone; mild inflammatory skin disorders; 1% ointment AAH Pharmaceuticals
Oral corticosteroids	15.20%	9.40%	8.10%	16.00%	16.00%	1.00	30	£0.86	£0.20	BNF; loratadine; symptomatic relief of allergy such as hay fever, chronic idiopathic urticaria; 10 mg
Antihistamines	4.90%	7.70%	13.80%	24.70%	24.70%	1.50	84	£2.36	£0.30	BNF; hydroxyzine hydrochloride; elderly dose; 10 mg tablets AAH Pharmaceuticals
Gabapentin/ pregabalin	11.00%	7.70%	6.50%	17.30%	17.30%	3.00	100	£2.74	£0.58	BNF; gabapentin; peripheral neuropathic pain; 300 mg capsule; Alliance Healthcare (Distribution) Ltd
Montelukast	1.80%	2.60%	0.80%	1.20%	1.20%	1.00	28	£1.37	£0.34	BNF; montelukast; prophylaxis of asthma; 10 mg tablet; AAH Pharmaceuticals Ltd (NHS indicative price)
Antidepressants	12.50%	17.90%	1.63%	21.00%	21.00%	3.50	30	£13.66	£11.16	BNF; doxepin; pruritus in eczema; xepin 5% cream Cambridge Healthcare Supplies Ltd
Anxiolytic/sedatives	4.30%	4.30%	1.60%	4.90%	4.90%	-	-	-	-	No appropriate cost could be identified

From the company's response to clarification[5] and Table 55 in the post-clarification version of Document B.[29] In turn, the company used the BNF as source material.[66]
BNF: British National Formulary; NHS = National Health Service

Table 4.16: Summary of total weighted treatment costs by health state

CKD-aP severity	ECM arm	DFK arm	
		List price	PAS price
None	£31.98	£5,424.64	[redacted]
Mild	£42.48	£5,435.14	[redacted]
Moderate	£42.48*	£5,435.14	[redacted]
Severe	£75.65	£5,468.31	[redacted]
Very severe	£75.65	£5,468.31	[redacted]

From the company's response to clarification[5] and Table 56 in the post-clarification version of Document B[29].
CKD-aP = chronic kidney disease associate pruritus; DFK = difelikefalin; PAS = Patient Access Scheme; ECM = established clinical management
*ECM costs for moderate CKD-aP were adjusted in model to equal costs for mild CKD-aP; the very severe health state is assumed equivalent to the severe health state

4.2.8.2 Disease management costs

The cost of managing CKD-aP patients in terms of healthcare resources was assumed to include hospitalisation, specialist visits, ICHD treatments, as well as kidney transplant operation and post-transplant treatment. General Practitioner (GP) visits were included as an optional input in the model but was not included in the base case.

Frequencies of healthcare resource use were sourced from clinical opinion and literature, while costs were sourced from the National Cost Collection.[67] The cost and frequencies used to estimate each health state management costs are all listed in Table 4.17 CKD management costs per health state, while description of each cost category and how it is derived is detailed in the subsequent sections. Upon request, the company updated the cost sourced from the National Cost collection from the year 2019/2020 to 2020/21.[67]

Table 4.17: CKD management costs per health state

Cost category	Frequency per health state					Cost used	Source
	None	Mild	Moderate	Severe	Very severe		
GP visit	0.00	0.00	0.00	0.00	0.00	£33.19	GP consultation; CS[1] and PSSRU 2021[68]
Hospitalisation	0.895	0.90	0.95	1.01	1.08	£3,004.43	CS,[1] Sukul 2021[25] and National Cost Collection 2020/2021[67]
Specialist visit (Nephrologist)	4	4	4	4	4	£242.48	Delphi panel,[69] and National Cost Collection 2020/2021[67]
Haemodialysis	154.08	154.08	154.08	154.08	154.08	£169.34	Assumption, National Cost

Cost category	Frequency per health state					Cost used	Source
	None	Mild	Moderate	Severe	Very severe		
							Collection 2020/2021[67]
Transplant						£20,901.72	National Cost Collection 2020/2021[67]
Post-transplant						£5,913.50	NHS Blood and Transplant fact sheet 7 (2009)[70]

Based on updated CS model input.[21, 71]
CKD = chronic kidney disease; CS = company submission; GP = General Practitioner; NHS = National Health Service; PSSRU = Personal Social Services Research Unit

Hospitalisation costs

The company includes in the model hospitalisation of patients for CKD-aP patients compared to CKD patients with no itch issues. The occurrence of all-cause hospitalisation for these patients is captured by hazard rates sourced from Sukul 2021 global study (see Table 4.18: HRs for all-cause hospitalisation).[25]

Table 4.18: Hazard ratios for all-cause hospitalisation

CKD-aP severity	Hazard ratio	95% CI
None	[redacted]	
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]
Severe	[redacted]	[redacted]
Very severe	[redacted]	[redacted]

Based on Table 57 in the CS[1]
CKD-aP = chronic kidney disease-associated pruritus; CI = confidence interval; CS = company submission

The cost of hospitalisation is estimated by company as a weighted average of the following LA08G (Chronic Kidney Disease with Interventions, with CC Score 6+), LA08H (Chronic Kidney Disease with Interventions, with CC Score 3-5), LA08J (Chronic Kidney Disease with Interventions, with CC Score 0-2), LA08K (Chronic Kidney Disease without Interventions, with CC Score 11+), LA08L (Chronic Kidney Disease without Interventions, with CC Score 8-10), LA08M (Chronic Kidney Disease without Interventions, with CC Score 5-7), LA08N (Chronic Kidney Disease without Interventions, with CC Score 3-4), and LA08P (Chronic Kidney Disease without Interventions, with CC Score 0-2) from the National Cost Collection.[67, 69]

Table 4.19: Overview of hospitalisation cost

Cost code and description		Activity	Unit cost	Cost used
LA08G	Chronic Kidney Disease with Interventions, with CC Score 6+	1,363	£8,684.19	
LA08H	Chronic Kidney Disease with Interventions, with CC Score 3-5	743	£6,781.29	
LA08J	Chronic Kidney Disease with Interventions, with CC Score 0-2	612	£5,777.89	
LA08K	Chronic Kidney Disease without Interventions, with CC Score 11+	1,596	£4,417.86	
LA08L	Chronic Kidney Disease without Interventions, with CC Score 8-10	2,087	£3,360.09	
LA08M	Chronic Kidney Disease without Interventions, with CC Score 5-7	3,599	£2,511.25	
LA08N	Chronic Kidney Disease without Interventions, with CC Score 3-4	3,226	£1,900.39	
LA08P	Chronic Kidney Disease without Interventions, with CC Score 0-2	6,350	£1,444.35	
Weighted average hospitalisation cost				£3,004.43

Based on updated CS model input[21, 71] and the National Cost Collection[67]
CS = company submission

Doctor and Specialist visit

Patient review visits to a specialist are assumed based on a modified Delphi panel.[69] In the original model the specialist visits and the associated costs were assumed to occur 17.33 times a year, i.e. once every 3 weeks; after clarification this was corrected to once every 3 months, i.e. 4 times per year. Specialist visits were initially priced using first attendance consultation (nephrologist led, non-admitted face-to-face follow-up (WF01A)) from the National Cost Collection,[67] however, after clarification this was changed to the tariff for follow-up attendance.

The company includes the option of including GP visits in the model. The base case however excludes the visits, and the EAG finds that the impact of including GP visits on the ICER is very small.

Table 4.20: Cost estimation

Cost code and description		Activity	Unit cost	Used cost
Nephrology WF01A	Non-admitted face-to-face attendance, follow-up	391,493	£242.48	
Specialist cost				£242.48

Cost code and description	Activity	Unit cost	Used cost
GP cost			£33.19
Based on updated CS model input,[21, 71] the National Cost Collection[67] and PSSRU[72] CS = company submission; GP = General Practitioner; PSSRU = Personal Social Services Research Unit			

In-centre haemodialysis

Costs of ICHD are estimated based on the cost of LD05A (satellite haemodialysis or filtration, with access via haemodialysis catheter, 19 years and over) and LD06A (satellite haemodialysis or filtration, with access via arteriovenous fistula or graft, 19 years and over) in the National Cost Collection.[67] The frequency of ICHD was based on the same UKRR data as used in the estimation of difelikefalin treatment cost, 2.963.[55]

In the base case however, dialysis costs were not included given the adjustment in risk of mortality for the very severe CKD-aP population and the resulting indirect increase in survival for the difelikefalin treatment arm. This approach was deemed appropriate with reference to NICE guidance (Section 4.4.16) which states that where a technology increases survival in people for whom the NHS is currently providing expensive care, background care costs may be removed.

Table 4.21: Cost estimation

Cost code and description		Number of sessions	National average unit cost	Used cost
LD05A	Satellite haemodialysis or filtration, with access via haemodialysis catheter, 19 years and over	658,152	£162.49	
LD06A	Satellite haemodialysis or filtration, with access via arteriovenous fistula or graft, 19 years and over	1,173,507	£173.19	
Weighted average in-centre haemodialysis cost				£169.34
Based on updated CS model input[21, 71] and the National Cost Collection[67] CS = company submission				

Transplant and post-transplant

Transplant and post-transplant costs are treated as lump sum costs. The frequency of transplantations was gathered from the probability of kidney transplant from NICE guideline NG107.[60] Costs of a transplant for an adult patient was derived as a weighted average of the cost codes of kidney transplant plus the weighted average of codes of pre-transplant work-up and examination post-transplant from the National Cost Collection for adult patients.[67] Post-transplant management was informed by a one-off cost estimate from the NHS Blood and Transplant fact sheet 7 (2009).[70] All cost codes and description of transplant are listed in Table 4.22, as well as the post-transplant management estimate.

Table 4.22: Cost estimation transplantation and post-transplantation

Cost code and description		Activity	Unit cost	Used cost
LA01A	Kidney transplant, 19 years and over, from cadaver non-heart-beating donor	457	£21,181.92	£20,106.37
LA02A	Kidney transplant, 19 years and over, from cadaver heart-beating donor	899	£19,249.13	

Cost code and description		Activity	Unit cost	Used cost
LA03A	Kidney transplant, 19 years and over, from live donor	266	£21,155.76	
LA11Z	Kidney pre-transplantation workup of live donor	1,821	£660.77	£482.50
LA12A	Kidney pre-transplantation workup of recipient, 19 years and over	7,616	£438.63	
LA13A	Examination for post-transplantation of kidney of recipient, 19 years and over	66,181	£312.23	£314.85
LA14Z	Examination for post-transplantation of kidney of live donor	2,693	£353.64	
Weighted average Transplant management cost				£20,901.72
Post-transplant management cost				£5,913.50
Based on updated CS model input[21, 71] and the National Cost Collection[67] CS = company submission				

EAG comment: As indicated earlier, the company did not include dialysis costs in the base case given the (indirect) increase in survival for the difelikefalin treatment arm. This was justified by the company by referring to NICE guidance (Section 4.4.16)[10] which states that where a technology increases survival in people for whom the NHS is currently providing expensive care, background care costs may be removed. However, the NICE guidance only states that *"In cases where a technology increases survival in people for whom the NHS is currently providing care that is expensive or would not be considered cost effective at NICE's normal levels, the committee may consider alongside the reference-case analysis a non-reference-case analysis with the background care costs removed. The committee will consider in its decision making both the reference-case and non-reference-case analyses, taking into account the nature of the specific circumstances of the evaluation including the population, care pathway and technology, as well as: the extent to which the cost effectiveness of the technology is driven by factors outside its direct costs and benefits if the NHS is already providing care that would not be considered cost effective at NICE's normal levels if the high-cost care is separate from direct, intrinsic consequences of the technology (such as a side effect or administration cost) the extent to which commercial solutions would address the issue."*

Based on the above, it is clear that removal of the ICHD costs should be done only for a non-reference case analysis, alongside an analysis that includes these costs. It should be noted that in an earlier comment the EAG already argued that there was insufficient evidence that treatment with difelikefalin would improve survival, rendering the above discussion moot (see Section 4.2.6.3).

4.2.8.3 Adverse event costs

In Section 4.2.6 the incidence rates for several AEs were presented, i.e. diarrhoea, dizziness, nausea, gait disturbance (falls), hyperkalaemia, headache, and somnolence, based on the KALM-1 and KALM-2 pooled data.[56] The company assigned the listed AEs each the cost of a single GP appointment, at £33.19, from which annual AE costs for DFK (£36.38) and ECM (£38.32) were calculated. The

company justifies using GP appointment as the associated cost of AE as there were no other relevant or appropriate costs for AEs identified in either the SLR or the adapted SLR.

EAG comment:

- The company used the AE incidence rate for difelikefalin based on the all-difelikefalin-exposure cohort from Fishbane 2022[56] while the IR for placebo was based on placebo-controlled cohort. Fishbane 2022 also report the IR for AE while on difelikefalin in the placebo-controlled cohort, which in general are rather higher than the used input. For example, IR for diarrhoea while exposed to difelikefalin is 469.2 in the placebo-controlled cohort, and 266.2 in the all-difelikefalin-exposure cohort. In the EAG comments of Section 4.2.6 it was already set out that the approach used by the company is flawed. However, when implementing the difelikefalin AE incidence rates of the placebo-controlled cohort, the impact on the ICER is negligible.

4.2.9 Disease severity

The new NICE process and methods manual describes disease severity as a decision modifier, i.e. depending on the severity of the disease, a higher threshold ICER may be used to consider if a new technology offers value for money. [10] According to the manual:

The committee will consider the severity of the condition, defined as the future health lost by people living with the condition with standard care in the NHS (including use of other available treatments, diagnostics, or best supportive care). The extent of unmet health need is reflected within the severity definition.

When assessing the severity of the condition in technology appraisals, the committee will consider the associated absolute and proportional QALY shortfall.

The QALY weightings for severity are applied based on absolute and proportional shortfall, whichever implies the greater severity level. If either the proportional or absolute QALY shortfall calculated falls on the cut-off between severity levels, the higher severity level will apply.

The cut-off point between severity levels are shown in Table 4.23.

Table 4.23: QALY weightings for severity

QALY weight	Proportional QALY shortfall	Absolute QALY shortfall
1	Less than 0.85	Less than 12
x 1.2	0.85 to 0.95	12 to 18
x 1.7	At least 0.95	At least 18
Based on NICE manual 2022[10] NICE = National Institute for Health and Care Excellence; QALY = quality-adjusted life year		

In their submission, the company stated that the technology is not expected to meet the criteria for a severity weight.

EAG comment: The EAG has used the tool developed by the institute for Medical Technology Assessment (iMTA Disease Burden Calculator - iDBC), which does not only provide the proportional and absolute QALY shortfall, but also the probability of the QALY shortfall falling in one of the three bands (see Table 4.23) using the PSA QALY results for the comparator group.[73]

In line with the tool developed by the university of York and Sheffield, "QALY Shortfall Calculator", the iDBC requires the average age of the population of interest and the male – female distribution. Furthermore, 1,000 PSA values for the QALYs in the comparator group are needed. This is in contrast with the "QALY Shortfall Calculator" which only requires the deterministic estimate for the QALYs in the comparator group.[74]

Since the QALYs in the comparator group can differ between the company base case and the EAG base case, the iDBC was run for each set of PSA QALYs separately. The results are shown in Table 4.24; the results of the QALY Shortfall Calculator are also shown for validation purposes. Comparing these results to the cut-off point in Table 4.23, the conclusion is that the QALY weight should remain 1. And this conclusion remains when taking the uncertainty into account.

Table 4.24: Estimated absolute and proportional QALY shortfall

	iDBC		QALY Shortfall Calculator	
	Company	**EAG**	**Company**	**EAG**
Remaining QALYs with standard treatment	2.74	2.86	2.74	2.86
QALYs without disease†	12.98	12.98	13.04	13.04
Absolute QALY loss	10.24	10.12	10.30	10.18
Proportional shortfall	0.79	0.78	0.79	0.78
Probabilistic results				
Weight = 1	99.8%	100%	Not available	
Weight = 1.2	0.2%	0%		
Weight = 1.7	0%	0%		

iDBC = iMTA Disease Burden Calculator; EAG = External Assessment Group; QALY = quality-adjusted life year
†Age was set at 58 years; percentage of male patients was set to 59%

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

5.1.1 Main results original company submission

Table 5.1 shows the company's deterministic base case results from the original submission. All results are discounted and include a single PAS discount value of [redacted] for difelikefalin. Total costs associated with difelikefalin treatment combined with established clinical management (DFK plus ECM) were estimated at [redacted] and total costs associated with ECM only were estimated at £30,442, indicating that addition of DFK to the ECM treatment increases total costs by [redacted]. Total QALYs associated with DFK plus ECM were estimated at [redacted] and total QALYs associated with ECM were estimated at 2.75, indicating an incremental number of [redacted] QALYs gained for patients treated with DFK plus ECM. These give an ICER for DFK plus ECM versus ECM only of £24,293 per QALY gained. The disaggregated results are shown in Table 5.2

Table 5.1: Company base case deterministic CE results, original submission (PAS price for difelikefalin)

Technologies	Total costs	Total LYG	Total QALYs	Inc. costs	Inc. LYG	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	4.65	[redacted]				
ECM	[redacted]	4.59	[redacted]	[redacted]	0.06	[redacted]	£24,293

Based on: Table 61 in CS[1]
CE = cost effectiveness; CS = company submission; DFK = difelikefalin; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

Table 5.2: Disaggregated base case results

Outcomes	DFK plus ECM	ECM	Incremental
Costs			
Treatment (list price)	£8,832	£476	£8,356
Treatment (PAS price)	[redacted]	£476	[redacted]
Adverse events	£101	£101	-£1
Management (total)	£30,102	£29,864	£238
None	£6,420	£1,173	£5,247
Mild	£3,882	£4,584	-£703
Moderate	£4,029	£6,269	-£2,239
Severe	£2,286	£3,868	-£1,582
Very severe	£719	£1,250	-£531
Transplant	£12,766	£12,721	£46
Health outcomes			
LYs (total)	**4.65**	**4.59**	**0.06**
None	1.02	0.19	0.83
Mild	0.61	0.72	-0.11

Outcomes	DFK plus ECM	ECM	Incremental
Moderate	0.62	0.97	-0.35
Severe	0.35	0.58	-0.24
Very severe	0.11	0.18	-0.08
Transplant	1.95	1.94	0.01
QALYs (total)	[redacted]	**2.75**	[redacted]
None	[redacted]	0.11	[redacted]
Mild	[redacted]	0.42	[redacted]
Moderate	[redacted]	0.50	[redacted]
Severe	[redacted]	0.25	[redacted]
Very severe	[redacted]	0.08	[redacted]
Transplant	[redacted]	1.38	[redacted]
Adverse events	[redacted]	0.00	[redacted]

Based on Table 62 and Table 63 in CS[1]
CS = company submission; DFK = difelikefalin; ECM = established clinical management; LYs = life years; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

5.1.2 Main results based on model after the request for clarification

During the clarification phase, the EAG asked the company about some inconsistencies and potential errors on the calculation of costs and resource use as implemented in the model. Furthermore, the company discovered calculation errors in the weekly dose estimations of ECM treatment and corrected these. Following the EAG's suggestion, the company updated management costs inputs from the National Cost Collection from 2019/2020 to 2020/2021.[67] In addition, the company adjusted the estimated average kidney transplant cost to only include costs for adult patients in the updated model. The EAG also noted discrepancies in the disease management cost category for specialist visits to a nephrologist between the CS report and the CE model, as the report mentioned 3-monthly visits, and the model estimated 3-weekly visits. The EAG also questioned the patient review visits cost being based on first attendance consultation (WF01B). The company noted in the response that the visit discrepancy was an error and updated the model for 3-monthly visits. The company also changed the visit cost to follow-up consultation (WF01A).

With these changes, the revised company base case following the clarification phase are presented in Table 5.3.

Table 5.3: Company base case deterministic cost effectiveness results, after clarification (discounted, with PAS price)

Technologies	Total costs	Total LYG	Total QALYs	Inc. costs	Inc. LYG	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	4.65	[redacted]				
ECM	£23,644	4.59	2.75	[redacted]	0.06	[redacted]	£23,277

Based on: Table 61 in CS clarification updates[29]
CS = company submission; DFK = difelikefalin; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; LYG = life years gained; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

As a consequence of the revision, the ICER has slightly decreased from £24,293 per QALY gained to £23,277 per QALY gained. Table 5.4 presents the disaggregated of the revised base case.

Table 5.4: Disaggregated base case results, after clarification

Outcomes	DFK plus ECM	ECM	Incremental
Costs			
Treatment (list price)	£8,545	£136	£8,408
Treatment (PAS price)	[redacted]	£136	[redacted]
Adverse events	£17	£37	-£20
Management (total)			
None	£3,717	£679	£3,038
Mild	£2,256	£2,664	-£408
Moderate	£2,384	£3,710	-£1,325
Severe	£1,385	£2,344	-£959
Very severe	£447	£777	-£330
Transplant	£13,345	£13,297	£48
Health outcomes			
LYs (total)	**4.65**	**4.59**	**0.06**
None	1.02	0.19	0.83
Mild	0.61	0.72	-0.11
Moderate	0.62	0.97	-0.35
Severe	0.35	0.58	-0.24
Very severe	0.11	0.18	-0.08
Transplant	1.95	1.94	0.01
QALYs (total)	[redacted]	**2.75**	[redacted]
None	[redacted]	0.11	[redacted]
Mild	[redacted]	0.42	[redacted]
Moderate	[redacted]	0.50	[redacted]
Severe	[redacted]	0.25	[redacted]
Very severe	[redacted]	0.08	[redacted]
Transplant	[redacted]	1.38	[redacted]
Adverse events	[redacted]	0.00	[redacted]

Based on Table 62 and Table 63 in CS[29]
CS = company submission; DFK = difelikefalin; ECM = established clinical management; LYs = life years; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

5.2 Company's sensitivity and scenario analyses

In this Section, only the results of the revised company analyses will be presented.

5.2.1 Probabilistic sensitivity analysis

After the corrections from the clarification phase, the PSA results in Table 5.5 below show that probabilistic results are still well aligned with the deterministic base case. The CE plane in Figure 5.1 shows that most of the simulations fell in the north-east quadrant. Based on the cost effectiveness acceptability curve (CEAC) in Figure 5.2, the probability that difelikefalin combined with ECM is CE at thresholds of £20,000 and £30,000 per QALY gained is 24% and 84%, respectively. Those percentages are slightly higher than the previous results before clarification.

Table 5.5: Company base case probabilistic CE results, after clarification (PAS price for difelikefalin)

Technologies	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	[redacted]			
ECM	[redacted]	[redacted]	[redacted]	[redacted]	£23,253

Based on Table 64 in CS after clarification[29]
CE = cost effectiveness; CS = company submission; DFK = difelikefalin; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

The EAG noted a difference between the totals QALYs for EMC in Table 64 of the CS (2.77 QALYs) and the results in the model (2.74 QALYs). The EAG assumed that the reported total QALYs are correct and report that value in Table 6.3.

Figure 5.1: * [redacted]

[redacted]

Based on Figure 18 of the updated CS after clarification[29]
CE = cost effectiveness; CS = company submission; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 5.2: [redacted]

Based on Figure 19 of the updated CS after clarification[29]
BSC = best supportive care (which is establish clinical management in most of the report); CS = company submission; CEAC = cost effectiveness acceptability curve; DFK = difelikefalin; PSA = probabilistic sensitivity analysis

5.2.2 Deterministic sensitivity analysis

The results of the company's deterministic sensitivity analysis (DSA) are displayed in Figure 5.3. Parameters relating to health state utilities had the largest impact on the ICER. The ICER exceed a willingness-to-pay (WTP) threshold of £30,000/QALY only when changing the utility score for the 'none' CKD-aP severity health state to its lowest value.

Figure 5.3: [redacted]

Based on Figure 20 of the CS after clarification[29]
BSC = best supportive care; CS = company submission; CI = confidence interval; DFK = difelikefalin; ICER = incremental cost-effectiveness ratio

5.2.3 Scenario analyses

Company scenario analysis results are presented Table 5.6. The rationale for each scenario is outlined in Table 65 of the updated CS after clarification.[29] The scenario with the largest impact on the results at PAS price is still using the observed data directly rather than the estimated transition probabilities from a simulated data set. This scenario increased the ICER to £37,913 per QALY gained. In the scenario where the MD was used in long-term extrapolation method for the ECM arm the ICER increased to £30,054. In all other scenarios the ICER remained below £30,000.

Table 5.6: Scenario analysis results at PAS price after clarification

Scenario	Assumption	Incr. costs	Incr. QALYs	ICER per QALY
Base case		[redacted]	[redacted]	£23,277
ECM extrapolation	Mean difference long-term extrapolation for ECM arm	[redacted]	[redacted]	£30,054
	Ratio of means long-term extrapolation for ECM arm	[redacted]	[redacted]	£25,409
Stopping rule	Stopping rule applied in week 8	[redacted]	[redacted]	£23,077
KALM-1 and KALM-2 data separately	KALM-1 trial data only	[redacted]	[redacted]	£25,817
	KALM-2 data only	[redacted]	[redacted]	£19,805
Observed data	Using the observed data directly (instate of a simulated data set)	[redacted]	[redacted]	£37,913
Efficacy plateau	Plateau after Year 2	[redacted]	[redacted]	£21,475
Treatment waning	In cycle 5, assume a probability of 5% for patients to gain a health state (deteriorate) each cycle	[redacted]	[redacted]	£25,915
	In cycle 5, assume a probability of 10% for patient to gain a health state (deteriorate) each cycle	[redacted]	[redacted]	£26,016
SHAREHD utility scored	SHAREHD utility scores	[redacted]	[redacted]	£21,584

Based on Table 65 of the updated CS after clarification[29] and the updated CE model[21, 71]
CE = cost effectiveness; CS = company submission; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; Incr. = incremental; PAS = Patient Access Scheme; QALY = quality-adjusted life year

5.2.4 Subgroup analysis

The following subgroups were included in the model:

- Patients only receiving anti-itch medication at baseline

- Patients not receiving anti-itch medication at baseline
- Patients only with severe or very severe itch at baseline

Subgroup analyses were performed to explore the effects of variation in the target population. Subgroup analysis results are displayed in Table 5.7. In all subgroup analyses the ICER remained below £30,000 per QALY gained.

Table 5.7: Results of the cost effectiveness analysis within the subgroups

Subgroup	Incr. costs	Incr. QALYs	ICER per QALY
Only receiving anti-itch medication at baseline	[redacted]	[redacted]	£23,993
Not receiving anti-itch medication at baseline	[redacted]	[redacted]	£25,922
Only with severe or very severe itch at baseline	[redacted]	[redacted]	£18,642

Based on Table 66-68 of the updated CS after clarification[29]
CS = company submission; Incr. = incremental; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life years

5.3 Model validation and face validity check

Validation efforts conducted on the economic model were shortly discussed in the validation section of the CS (Section B.3.13).[1] The company indicated in Section B.3.13. of the CS that most of the validation efforts focused on internal quality assurance measures undertaken throughout the model development phase. The model was validated by using extreme values and formula auditing to ensure the consistency of model estimates.

The company further solicited expert opinion to validate key model inputs and assumptions from a clinical perspective. For instance, clinical opinion was requested to validate the natural progression of CKD-aP and the potential trend in the mean change in itch score that could be expected in the extrapolation period for patients receiving placebo in the KALM trials. The model structure and inputs were also critiqued and validated by an external health economics consultant. The CS states that overall, the validation process did not identify issues with the structural or computational accuracy of the model.

EAG comment: In clarification Question B.34 the company was asked to provide an internal validation test showing the extent to which the model results match the observed data for the first 64 weeks of the KALM trials. In response to this request the company presented the distribution of patients in the DFK arm at week 12 and week 64 when using the 'change in state' transition estimates (base case) and the directly observed transitions as estimated from the pooled patient level data for the KALM trials. The results were quite similar and the EAG had no major concerns about this evidence.

6 IMPACT ON THE COST-CONSEQUENCE ANALYSIS OF ADDITIONAL EXPLORATORY CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE EAG

6.1 Exploratory and sensitivity analyses undertaken by the EAG

6.1.1 Explanation of the company adjustments after the request for clarification

During the clarification phase, the EAG requested the company to make the following amendments to the model:

- Update management costs inputs from National Cost Collection from 2019/2020 to 2020/2021.
- The company identified an error in the calculation of weekly dose of ECM costs in the model. The calculation error was corrected for the updated model. In addition, the cost of hydrocortisone 1% cream was updated to £1.26.
- Correct discrepancies in the disease management cost category for specialist visits to a nephrologist between the CS report and the CE model, as the report mentioned 3-monthly visits, and the model estimated 3-weekly visits. The company responded that the visit discrepancy was an error and updated the model for 3-monthly visits.
- The EAG questioned the patient review visits cost being based on first attendance consultation (WF01B). The company changed the visit cost to follow-up consultation (WF01A).
- The percentage of male patients was the correct input from Fishbane 2022.[56]

6.1.2 Explanation of the EAG adjustments

The changes that the EAG can make (to the model received with the response to the clarification letter) can be subdivided into the following three categories (according to Kaltenthaler 2016[75]).

- Fixing errors (FE) (correcting the model where the company's electronic model is unequivocally wrong).
- Fixing violations (FV) (correcting the model where the EAG considers that the NICE reference case, scope, or best practice has not been adhered to).
- Matters of judgement (MJ) (amending the model where the EAG considers that reasonable alternative assumptions are preferred).

In the current assessment, there were no errors identified in the model following the clarification phase, and five MJ played a role. After the proposed changes were implemented in the company's model, additional scenario analyses were explored by the EAG in order to assess the impact of alternative assumptions on the CE results.

6.1.2.1 Fixing errors

There were no errors identified in the model following the clarification phase.

6.1.2.2 Fixing violations

No violations were applicable to this appraisal.

6.1.2.3 Matters of judgement

The EAG's preferences regarding alternative assumptions led to the following changes to the company base case analysis:

- The EAG prefers to use the observed data to directly estimate the transition matrices from all to all health states (see Section 4.2.6.2).

- The company assumed a treatment waning for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In absence of real-world evidence to support the waning impact of the ECM and the continuing impact of difelikefalin, the EAG considers this assumption uncertain and removed it from the base case analysis (see Section 4.2.6.2).
- The EAG does not agree with the company's approach to use an increased risk of death assumed for patients in the moderate, severe and very severe health states of the CKD-aP population based on the study by Sukul 2021[25] as the EAG considers the evidence presented not substantial to establish a causal relationship between pruritus and mortality of these patients. Therefore, the EAG removed this elevated risk of death for these patients from the model (see Section 4.2.6.3).
- The cost of haemodialysis treatment are included in the model for completeness. However, this only impacts the costs per treatment group, and not the incremental costs, unless a scenario with survival benefit is explored.

The overview of the changes and the bookmarks for the justification of the EAG changes are presented in Table 6.1.

Table 6.1: Company and EAG base case preferred assumptions

Base case preferred assumptions	Company	EAG	Justification for change
Transition probabilities	Estimated based on simulated data	Estimated based on observed data	Section 4.2.6.2
Waning effect for the ECM arm	5% probability of deteriorating per year	No waning	Section 4.2.6.2
Elevated risk of death for patients in moderate, severe and very severe health states	Increased risk of death assumed based on the study of Sukul 2021	No increased risk of death	Section 4.2.6.3
Cost of haemodialysis	Costs of haemodialysis excluded	Costs of haemodialysis included	Section 4.2.8.2
EAG = Evidence Assessment Group; ECM = established clinical management			

6.1.3 Additional scenarios conducted by the EAG

The EAG conducted a series of scenario analyses to explore the impact of key assumptions and uncertainties within the CE analyses.

6.1.3.1 Scenario 1: Transition probabilities

The EAG explored the impact on the ICER of using the transition probabilities as estimated based on the simulated data and the change of state observed data.

6.1.3.2 Scenario 2: Treatment waning

The EAG explored the impact of treatment waning for both the ECM and the DFK arms equal to a 5% and 10% probability of deteriorating per year. Additionally, in line with the company base case, we also explored the impact of 5% and 10% waning for just the EMC patients.

6.1.3.3 Scenario 3: Increased risk of mortality

The impact of an increased risk of death assumed for patients in the moderate, severe and very severe health states of the CKD-aP population based on the study by Sukul 2021[25] was investigated in scenario analysis.

6.1.3.4 Scenario 4: Health state utilities

Utility values estimated from the SHAREHD database were used to inform health state utilities for each of the five health states of itching as shown in Table 4.13.

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the EAG

6.2.1 Results of the EAG preferred base case scenario

The EAG preferred base case incremental CE results, provided in Table 6.2, indicate that the ICER, compared to the company base case, has substantially increased. The company base case ICER after clarification amounted to £23,277 per QALY gained, whereas the ICER for the EAG preferred base case is £35,048 per QALY gained.

Table 6.2: EAG base case deterministic CE results (discounted)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	4.78	[redacted]	[redacted]	0.00	[redacted]	£35,048
ECM	£97,611	4.78	2.88				
Based on the EAG preferred version of the electronic model (version in response to question B12)[21] CE = cost effectiveness; EAG = Evidence Assessment Group; Incr. = incremental; DFK = difelikefalin; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years							

Results of the PSA in Table 6.3 below show that probabilistic results are higher than the EAG deterministic base case. This is due to the skewness in the distribution around the transition probabilities whenever these are very close to zero, i.e., a 0% transition in the deterministic analysis will become a small but non-0% transition in the PSA. The CE plane in Figure 6.1 shows that all the simulations fell in the north-east quadrant. Based on the CEAC in Figure 6.2, the probability that difelikefalin combined with ECM is cost effective at thresholds of £20,000 and £30,000 per QALY gained are 0% and 13%, respectively, using the EAG base case assumptions.

Table 6.3: EAG base case probabilistic CE results (discounted)

Technologies	Total costs (£)	Total QALYs	Incr. costs (£)	Incr. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	[redacted]	[redacted]	[redacted]	£41,157
ECM	£97,695	2.86			
Based on the EAG preferred version of the electronic model[21] CE = cost effectiveness; DFK = difelikefalin; EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; Incr. = incremental; LYG = life years gained; QALYs = quality-adjusted life years					

Figure 6.1: PSA CE plane – EAG base case

CE = cost effectiveness; EAG = Evidence Assessment Group; PSA = probabilistic sensitivity analysis; QALYs = quality-adjusted life years; WTP = willingness-to-pay

Figure 6.2: PSA CEAC EAG base case

BSC = best supportive care; CEAC = cost effectiveness acceptability curve; DFK = difelikefalin; EAG = Evidence Assessment Group; PSA = probabilistic sensitivity analysis

6.2.2 Results of the EAG additional exploratory scenario analyses.

The results of the scenario analyses are provided in Table 6.4.

Table 6.4: Results of exploratory scenario analyses by the EAG

Scenario	DFK plus ECM		ECM		Incr. costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
EAG base case	[redacted]	[redacted]	£97,611	2.88	[redacted]	[redacted]	£35,048
Scenario 1a: observed change of state	[redacted]	[redacted]	£97,697	2.77	[redacted]	[redacted]	£51,521
Scenario 1b: simulated change of state	[redacted]	[redacted]	£97,650	2.76	[redacted]	[redacted]	£29,879
Scenario 2a: 5% waning for the ECM arm	[redacted]	[redacted]	£97,692	2.85	[redacted]	[redacted]	£30,092
Scenario 2b: 10% waning for the ECM arm	[redacted]	[redacted]	£97,769	2.83	[redacted]	[redacted]	£26,646
Scenario 2c: 5% waning for both the ECM and the DFK arms	[redacted]	[redacted]	£97,692	2.85	[redacted]	[redacted]	£34,855
Scenario 2d: 10% waning for both the ECM and the DFK arms	[redacted]	[redacted]	£97,769	2.83	[redacted]	[redacted]	£35,437
Scenario 3: increased risk of mortality with more severe itching	[redacted]	[redacted]	£92,832	2.78	[redacted]	[redacted]	£38,283
Scenario 4: health state utilities from SHAREHD	[redacted]	[redacted]	£97,611	3.21	[redacted]	[redacted]	£32,892

Based on the EAG preferred version of the electronic model (using 'scenario B12')[21]
DFK = difelikefalin; EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; Incr. = incremental; QALYs = quality-adjusted life years

6.2.3 Subgroup analysis

The EAG did not conduct a subgroup analysis based on anti-itch treatment at baseline, since the relevant transition matrices were not available for the direct measured transitions.

The subgroup analysis in which only patients with severe and very severe pruritus start treatment with difelikefalin showed an ICER of £30,274 per QALY gained, which is clearly lower than the ICER for the population that also includes patients with moderate pruritus (see table 6.5).

Table 6.5: Results of subgroup analysis for patients with severe and very severe pruritis only

Preferred assumption	DFK plus ECM		ECM		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
EAG's preferred base case	[redacted]	[redacted]	£97,611	2.88	[redacted]	[redacted]	£35,048
EAG's model, severe and very severe only	[redacted]	[redacted]	£97,742	2.83	[redacted]	[redacted]	£30,274

Analysis by company, using the EAG preferred base case provided as part of the company's response to the FAC.[76]
DFK = difelikefalin; EAG = Evidence Assessment Group; ECM = established clinical management; FAC = factual accuracy check; ICER = incremental cost-effectiveness ratio; Incr. = incremental; QALYs = quality-adjusted life years

6.3 EAG preferred assumptions

Table 6.6 shows the step-by-step changes made by the EAG to the company base case. The change with by far the largest impact on the results was changing the source for the transition probabilities using the observed instead of the simulated data. This change leads to increase of the ICER to over £30,000. Also important is the removal of the elevated risk of death for patients in moderate, severe and very severe health states which is also related to change on the costs of haemodialysis costs. The impact of disregarding the treatment waning impact for the ECM arm and incorporation of the AEs disutilities is relatively small, with the first change leading to an increase in the number of QALYs gained and the second to a decrease in the number of QALYs gained.

Table 6.6: Individual impact of EAG preferred assumptions

Preferred assumption	DFK plus ECM		ECM		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
Company base case (original)	[redacted]	[redacted]	£30,442	2.75	[redacted]	[redacted]	£24,293
Company base case (after clarification)	[redacted]	[redacted]	£23,644	2.75	[redacted]	[redacted]	£23,277
EAG change on transition probabilities	[redacted]	[redacted]	£23,590	2.76	[redacted]	[redacted]	£25,792
EAG change on waning effect for the ECM arm	[redacted]	[redacted]	£23,626	2.78	[redacted]	[redacted]	£26,320
EAG change on elevated risk of death for patients in moderate, severe and very severe health states	[redacted]	[redacted]	£24,476	2.84	[redacted]	[redacted]	£27,566

Preferred assumption	DFK plus ECM		ECM		Incr. Costs (£)	Incr. QALYs	ICER (£/QALY)
	Costs (£)	QALYs	Costs (£)	QALYs			
EAG change on cost of haemodialysis	[redacted]	[redacted]	£92,732	2.75	[redacted]	[redacted]	£33,723
Based on the EAG preferred version of the electronic model DFK = difelikefalin; EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; Incr. = incremental; QALYs = quality-adjusted life years							

6.4 Conclusions of the cost effectiveness section

The company developed a Markov model in Microsoft Excel® to assess the CE of difelikefalin for the treatment of adult patients with moderate-to-severe CKD-aP who are on haemodialysis. The model consists of seven mutually exclusive health states, five core health states defined by the level of itch severity (i.e. none, mild, moderate, severe, and very severe), and the health states of renal transplant and death.

The EAG considers that the model structure adequately reflects clinical issues related to patients with moderate-to-severe CKD-aP who are on haemodialysis. Therefore, the model structure appears to be appropriate and fit for purpose. The CE analysis was performed in line with the NICE Reference Case in terms of perspective, time horizon and discounting

The population in the company's analysis are adult patients receiving in-centre haemodialysis for CKD who suffer from moderate-to-severe pruritus, where ECM is insufficient in reducing pruritus. The treatments compared were difelikefalin with ECM versus ECM alone.

The model consists of a short period of three 4-week cycles, after which an assessment is made if difelikefalin has led to enough improvement to warrant continuation of the treatment. Transition probabilities for this period of 12 weeks were derived from the double-blind phase of the KALM-1 and KALM-2 studies. The 5-D Itch score was used to establish which health state a patient would occupy. After this period, the model works with cycles of 1 year. Transition probabilities for the first full year (week 12 to 64) of difelikefalin were derived from the OLE study, whereas for the ECM group is was assumed that patient would not change their health state. For all years thereafter it was assumed that patients receiving difelikefalin would remain in the health state they were in at 64 weeks, whereas ECM patients would slowly move to worse health states (5% each year would move to one state worse).

To estimate transition probabilities between the severity health states for each of the treatment arms, first methods of multiple imputation were used to address issues with missing data in the patient-level data set for the total 5-D Itch scale scores. The missing values were estimated based on treatment group, baseline itch score, and patient characteristics including age, sex, diabetes status, ESRD, length of haemodialysis, length of CKD-aP, and use of anti-itch medication at baseline.

The transition probabilities were estimated from a simulated data set. The simulation method used the mean change from baseline in itch scores by CKD-aP severity levels at baseline: moderate, severe or very severe. The mean change values were in turn added to the baseline itch scores to estimate the mean itch score for each of the corresponding weeks (week 4, week 8, week 12, and week 64). The simulated itch scores were further grouped into health states and used to estimate the change in health state from the previous cycle, weighted by the distribution of patients across the different severity itching health states at baseline. With this approach it was assumed that the probability to move to one, two or three states better or worse would not depend on the current health state.

During clarification, the company also provided the direct estimates of all transition probabilities. When the matrices of the simulated method were compared to these direct estimates, the EAG concluded that the idea that the probability of moving one state down does not depend on the current health state is not supported by the data. Furthermore, the simulated method only works with average changes from baseline in the 5- D Itch score and thus reduces the variation observed in the data. Hence, the EAG considers it best to not rely on any assumptions and use the transition probabilities as directly derived from the patient-level data.

It is unclear to the EAG how all transition matrices were derived by the company in light of the multiple imputation applied to account for missing data. For example, when looking at the directly estimated transition probabilities as presented in response to the clarification letter, it is unclear if these probabilities are based on averages over 20 different probabilities, each from a different complete dataset. It is not transparent how analyses per complete dataset were combined in order to find the final estimates, and how uncertainty was estimated (which should be a function of within dataset variation and between dataset variation).

The model also includes an annual transplant probability and annual mortality probabilities, the later split into mortality for patients on haemodialysis and mortality for patients after a transplant. Annual probabilities of death and transplant were estimated for patients on haemodialysis from 1 to 10 years after initiating dialysis; after year 10 the annual probability of death was assumed to remain unchanged whilst the probability of transplant was assumed to be 0. The probability of death for patients who received transplant was assumed to be the same as for the general population.

Furthermore, in the base case analysis, an increased mortality risk was applied for patients in the moderate, severe and very severe health states of the CKD-aP population using respective HRs of 1.11, 1.02, and 1.24 extracted from the study of Sukul 2021.[25]

The EAG has concerns around the elevated risk of death assumed in the model for patients in the moderate, severe and very severe health states of the CKD-aP population. Sukul 2021[25] find that extreme pruritus is an independent predictor of all-cause and case-specific mortality after controlling for multiple confounders such as patient demographic and clinical characteristics. At the same time, they acknowledge that the possible bidirectionality of the relationship between pruritus and cross-sectional patient-reported outcomes (such as depression, missed dialysis sessions, and poor sleep quality) limits the inferences that can be made and does not allow conclusions about cause-effect relationships. Additionally, the authors state that one cannot make inferences in how changes in pruritus severity may relate to the outcomes in their study.

The company explained that increased depression, missed dialysis sessions, poor sleep quality, and skin lesions susceptible to infection are outcomes that could mediate the relationship between extreme pruritis and mortality and that thus a causal relationship could exist. However, the evidence currently presented are not considered substantial enough to establish a causal relationship between a reduction in itching score due to treatment with difelikefalin and a reduction in mortality of these patients. Therefore, the EAG considers that this impact should not be part of the base case computations and instead explored in a scenario analysis.

As no generic preference-based measures of health were collected in the KALM-1 or KALM-2 trials, a separate primary data collection study across UK dialysis centres was undertaken to develop a mapping algorithm relating the 5-D Itch scale to the EQ-5D-3L.

This primary data collection was undertaken between November 2020 and June 2021 across five sites in England on adult patients who had been receiving haemodialysis for at least 3 months. The data collected was used to estimate EQ-5D-3L mapping functions from 5-D Itch scale scores. All mapping functions included age, sex, diabetes status, and length of time on dialysis as additional conditioning variables. Despite limitations with missing observations, the 5-D Itch scale score to EQ-5D-3L mapping algorithm was considered the most appropriate option, given the paucity of published data in CKD-aP.

The EAG considers the mapping study that was done to inform the health states utilities of high quality, based on state-of-the-art methodology. According to the NICE methods guide, mapping a disease specific QoL instrument to the EQ-5D is an appropriate approach if no direct EQ-5D data is available.

Another potential source of EQ-5D utilities was the study by Thokala 2021[64]. They present the results of the mapping study discussed above, as well as utility values estimated from the SHAREHD database. All these utilities are higher than those derived from the mapping study. The impact of using these utility values was explored through a scenario analysis.

The following cost categories were included in the model: drug acquisition costs, drug administration costs for intravenously administered drugs, disease management costs, and costs of treatment-related AEs.

The annual cost of the difelikefalin treatment are based on the used injection volumes by estimating the number of vials per average patient weight of 84.4 kg[56] and frequency of ICHD (2.96 ICHD sessions per week).[55] To account for wastage, the number of vials per patient were rounded upwards. An annual cost of the difelikefalin treatment of £5,392.66 at list price was derived, and [redacted] at PAS price.

Treatment with ECM includes only the anti-itch medication consumed by CKD-aP patients. Proportions of patients consuming CKD-aP treatments per health state was based on the KALM trials.[56] The anti-itch medication dose, pack size, and price were based on BNF.[66] All medications listed were weighted against the frequency at which they are consumed within each health state and the proportion of overall consumption by the specific health state.

The cost of managing CKD-aP patients in terms of healthcare resources was assumed to include hospitalisation, specialist visits, ICHD treatments, as well as kidney transplant operation and post-transplant treatment. The occurrence of all-cause hospitalisation varied based by itch severity.

The company opted not to include the costs of dialysis in their base case. This was justified by the company by referring to NICE guidance (Section 4.4.16)[10] which states that where a technology increases survival in people for whom the NHS is currently providing expensive care, background care costs may be removed. However, the NICE guidance actually states that removal of such costs should be done only for a non-reference case analysis, alongside an analysis that includes these costs. It should be noted that in an earlier comment the EAG already argued that there was insufficient evidence that treatment with difelikefalin would improve survival, rendering the above discussion moot. For completeness, the EAG opted to include the ICHD costs in their base case.

The company's deterministic base case analysis showed that the total costs associated with difelikefalin treatment combined with ECM were estimated at [redacted] and total costs associated with ECM only were estimated at £23,644, indicating that addition of difelikefalin to the ECM treatment increases total costs by [redacted]. Total QALYs associated with DFK plus ECM were estimated at [redacted] and total QALYs associated with ECM were estimated at 2.75, indicating an incremental number of [redacted] QALYs gained for patients treated with DFK plus ECM. These give an ICER for DFK plus ECM versus ECM only of

£23,277 per QALY gained. All results are discounted and include a single PAS discount value of [redacted] for difelikefalin.

The PSA the company did showed that the probability that difelikefalin combined with ECM is cost effective at thresholds of £20,000 and £30,000 per QALY gained is 24% and 84%, respectively.

The company performed various scenario analyses to assess the impact of alternative assumptions on the ICER. For most scenarios the ICER was close to the base case ICER. There were two scenarios with a slightly larger impact. The first concerns the assumption about the transition probabilities between 12 and 64 weeks for the ECM group. In the base case patients were assumed to stay in the same health state during that period. In the scenario, these transitions were derived using the MD between difelikefalin and ECM in the first 12 weeks and applying this to the difelikefalin transitions to estimate the ECM transitions. With this approach, the ICER was £30,054. In the second scenario, transitions were no longer based on the simulated data but the observed data about change of state. This yielded an ICER of £37,913.

The EAG's preferences regarding alternative assumptions for the model led to a number of changes to the company base case analysis. Most importantly, the EAG prefers to use the observed data to directly estimate the transition matrices from all to all health states. In addition, the EAG considers that there is a clear lack of any evidence for the company assumption of a treatment waning for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. Thus, in absence of real-world evidence to support the waning impact of the ECM and the continuing impact of difelikefalin, the EAG considers this assumption uncertain and removed it from the base case analysis. Furthermore, the EAG does not agree with the company's approach to assume an increased risk of death for patients in the moderate, severe and very severe health states of the CKD-aP population based on the study by Sukul 2021[25] as the EAG considers the evidence presented not substantial enough to establish a causal relationship between pruritus and mortality of these patients. Therefore, the EAG removed this elevated risk of death for these patients from the model. Finally, the cost of haemodialysis treatment is included in the model for completeness. However, this only impacts the costs per treatment group, and not the incremental costs, unless a scenario with the above-mentioned survival benefit is explored.

These changes in the model lead to the following EAG preferred base case incremental cost effectiveness results. The total costs for difelikefalin amount to [redacted], versus £97,611 for ECM. At the same time [redacted] and 2.88 QALYs are accumulated, for difelikefalin and EMC, respectively. This leads to an ICER of £35,048, which is higher than the company ICER of £23,277 per QALY gained.

The probabilistic ICER, £41,157 per QALY gained, is higher than the EAG deterministic base case. This is due to the skewness in the distribution around the transition probabilities whenever these are very close to zero, i.e., a 0% transition in the deterministic analysis will become a small but non-0% transition in the PSA. The PSA shows that the probability that difelikefalin combined with ECM is cost effective at thresholds of £20,000 and £30,000 per QALY gained are 0% and 13%, respectively, using the EAG base case assumptions.

The EAG assessed disease severity as a potential decision modifier. It was found that the probability of a QALY weight of 1 being applicable was 100%.

Several scenarios were explored, and most of these led to only small changes in the ICER. The most substantial change occurred when transition probabilities were derived using the observed data to

estimate to probability of a change of state, independent on the current state. This scenario yielded an ICER of £51,521 per QALY gained.

8. REFERENCES

[1] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Document B. Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 11.8.22]

[2] European Medicines Agency. *Kapruvia 50 micrograms/mL solution for injection. EPAR - Product Information. Annex I. Summary of product characteristics [Internet]*. Amsterdam, The Netherlands: European Medicines Agency, 2022 [accessed 11.8.22] Available from: https://www.ema.europa.eu/en/documents/product-information/kapruvia-epar-product-information_en.pdf

[3] Kidney Research UK. *Kidney health inequalities in the UK: an agenda for change [Internet]*. Peterborough Kidney Research UK, 2018 [accesed 11.8.22] Available from: https://www.kidneyresearchuk.org/wp-content/uploads/2019/09/Health_Inequalities_lay_report_FINAL_WEB_20190311.pdf

[4] National Institute for Health and Care Excellence. *Single Technology Appraisal: Difelikefalin for treating pruritus in people having haemodialysis: final scope [Internet]*. London: NICE, 2022 [accessed 20.6.22] Available from: https://www.nice.org.uk/guidance/gid-ta10793/documents/final-scope

[5] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Response to request for clarification from the EAG*, 2022 [accessed 20.9.22]. 82p.

[6] Cara Therapeutics Inc. *CR845-CLIN3102 Clinical study report – double-blind phase (AIC): a multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of intravenous CR845 in hemodialysis patients with moderate-to-severe pruritus, with a 52-week open-label extension [PDF provided by company]*, 2020 [accessed 11.8.22]

[7] Cara Therapeutics Inc. *CR845-CLIN3103 Clinical study report – double-blind phase (AIC): multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of intravenous CR845 in hemodialysis patients with moderate-to-severe pruritus, with a 52-week open-label extension [PDF provided by company]*, 2020 [accessed 11.8.22]

[8] Wheeler K. FDA approves treatment for uremic pruritis Korsuva is approved for treating moderate to severe pruritis associated with chronic kidney disease. *Drug Topics* 2022;166(4):35.

[9] Vifor Pharma UK, Initiate Consultancy Ltd. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Appendix D, G, H and I: Systematic Literature Review of Chronic kidney disease-associated pruritus (CKD-aP)*, 2022 [accessed 11.8.22]

[10] National Institute for Health and Care Excellence. *NICE health technology evaluations: the manual. Process and methods [PMG36] [Internet]*. London: National Institute for Health and Care Excellence, 2022 [accessed 14.10.22]. 181p. Available from: https://www.nice.org.uk/process/pmg36

[11] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[12] Fishbane S, Mathur V, Germain MJ, Shirazian S, Bhaduri S, Munera C, et al. Randomized controlled trial of difelikefalin for chronic pruritus in hemodialysis patients. *Kidney Int Rep* 2020;5(5):600-10.

[13] Narita I, Tsubakihara Y, Uchiyama T, Okamura S, Oya N, Takahashi N, et al. Efficacy and safety of difelikefalin in Japanese patients with moderate to severe pruritus receiving hemodialysis: a randomized clinical trial. *JAMA Network Open* 2022;5(5):e2210339.

[14] Feng WW, Yuan B, Shen FY, Fan WY, Mei DS, Bao BY, et al. Efficacy of uremic pruritus treatment in patients undergoing hemodialysis, a network meta-analysis for randomized clinical trials. *Nephrol Ther* 2021;17(1):30-4.

[15] Wala K, Szepietowski JC. Difelikefalin in the treatment of chronic kidney disease-associated pruritus: a systematic review. *Pharmaceuticals (Basel)* 2022;15(8):934.

[16] Fishbane S, Jamal A, Munera C, Wen W, Menzaghi F. A phase 3 trial of difelikefalin in hemodialysis patients with pruritus. *N Engl J Med* 2020;382(3):222-32.

[17] National Institute for Health and Care Excellence. *Single technology appraisal and highly specialised technologies evaluation: user guide for company evidence submission template [Internet]*. London: NICE, 2015. Last updated: 10 February 2022 [accessed 14.10.22]. 28p. Available from: https://www.nice.org.uk/process/pmg24/

[18] Cara Therapeutics Inc. *CR845-CLIN3102 Clinical study report – open-label extension phase (AIC): multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of intravenous CR845 in hemodialysis patients with moderate-to-severe pruritus, with a 52-week open-label extension [PDF provided by company]*, 2020 [accessed 11.8.22]

[19] Cara Therapeutics Inc. *CR845-CLIN3103 Clinical study report – open-label extension phase (AIC): a multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of intravenous CR845 in hemodialysis patients with moderate-to-severe pruritus, with a 52-week open-label extension [PDF provided by company]*, 2020 [accessed 11.8.22]

[20] Cara Therapeutics Inc. *CR845-CLIN3105 Clinical study report (AIC): an open-label, multicenter, study to evaluate the safety and effectiveness of intravenous CR845 in hemodialysis patients with moderate-to-severe pruritus [PDF provided by company]*, 2020 [accessed 11.8.22]

[21] Vifor Pharma UK, Initiate Consultancy Limited. *Cost-effectiveness analysis of difelikefalin (DFK) for the treatment of CKD-aP (ID3890 CE Model PAS price__clarificationquestions_direct transitions) [Excel file updated following clarification questions]*, 2022 [accessed 20.9.22]

[22] Weisshaar E, Szepietowski JC, Dalgard FJ, Garcovich S, Gieler U, Giménez-Arnau AM, et al. European S2k guideline on chronic pruritus. *Acta Derm Venereol* 2019;99(5):469-506.

[23] Hercz D, Jiang SH, Webster AC. Interventions for itch in people with advanced chronic kidney disease. *Cochrane Database Syst Rev* 2020:Issue 12. Art. No.: CD011393. DOI: 10.1002/14651858.CD011393.pub2.

[24] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Response to request for clarification from the EAG: Additional appendices (P-V) [folder provided by the company]*, 2022 [accessed 20.9.22]

[25] Sukul N, Karaboyas A, Csomor PA, Schaufler T, Wen W, Menzaghi F, et al. Self-reported pruritus and clinical, dialysis-related, and patient-reported outcomes in hemodialysis patients. *Kidney Med* 2021;3(1):42-53.e1.

[26] Topf J, Wooldridge T, McCafferty K, Schömig M, Csiky B, Zwiech R, et al. Efficacy of difelikefalin for the treatment of moderate-to-severe pruritus in hemodialysis patients: pooled analysis of KALM-1 and KALM-2 phase 3 studies [PDF provided by company]. *Kidney Med* 2022;PrePrint 2022 June 12.

[27] Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. *Biometrics* 1999;55(3):853-7.

[28] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). MedDRA (Medical Dictionary for Regulatory Activities). Japanese Edition, September 2018 - Version 21.1 [Internet]. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, 2018 [accessed 14.10.22]. Available from: https://www.meddra.org/how-to-use/support-documentation/japanese/welcome

[29] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Document B [revised after clarification]. Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 20.9.22]

[30] Altman DG. Rapid response to: exercise training meta-analysis of trials in patients with chronic heart failure (ExTraMATCH). *BMJ* 2004;328:189

[31] Vernon MK, Swett LL, Speck RM, Munera C, Spencer RH, Wen W, et al. Psychometric validation and meaningful change thresholds of the Worst Itching Intensity Numerical Rating Scale for assessing itch in patients with chronic kidney disease-associated pruritus. *J Patient Rep Outcomes* 2021;5(1):134.

[32] Vernon M, Ständer S, Munera C, Spencer RH, Menzaghi F. Clinically meaningful change in itch intensity scores: an evaluation in patients with chronic kidney disease-associated pruritus. *J Am Acad Dermatol* 2021;84(4):1132-4.

[33] Mathur VS, Lindberg J, Germain M, Block G, Tumlin J, Smith M, et al. A longitudinal study of uremic pruritus in hemodialysis patients. *Clin J Am Soc Nephrol* 2010;5(8).

[34] Storck M, Sandmann S, Bruland P, Pereira MP, Steinke S, Riepe C, et al. Pruritus Intensity Scales across Europe: a prospective validation study. *J Eur Acad Dermatol Venereol* 2021;35(5):1176-85.

[35] Reich A, Heisig M, Phan NQ, Taneda K, Takamori K, Takeuchi S, et al. Visual analogue scale: evaluation of the instrument for the assessment of pruritus. *Acta Derm Venereol* 2012;92(5):497-501.

[36] Cara Therapeutics Inc. *CR845-CLIN3102 Statistical Analysis Plan - Double-Blind Phase (AIC) [PDF provided by the company]*: Cara Therapeutics Inc, 2019 [accessed 11.8.22]

[37] Dworkin RH, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;113(1-2).

[38] Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. *Br J Dermatol* 2010;162(3):587-93.

[39] Topf J, Wooldridge T, McCafferty K, Schomig M, Csiky B, Zwiech R, et al. Efficacy of difelikefalin for the treatment of moderate to severe pruritus in hemodialysis patients: pooled analysis of KALM-1 and KALM-2 phase 3 studies. *Kidney Med* 2022;4(8):100512.

[40] Fishbane S, Wen W, Munera C, Lin R, Bagal S, McCafferty K, et al. Safety and tolerability of difelikefalin for the treatment of moderate-to-severe pruritus in hemodialysis patients: pooled analysis from the phase 3 clinical trial program [Pre-proof provided by the company]. *Kidney Med* 2022 [accessed 11.8.22].

[41] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. *J Clin Epidemiol* 2016;75:40-6.

[42] Canadian Agency for Drugs and Technologies in Health. *PRESS - Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E) [Internet]*. Ottawa: CADTH, 2016 [accessed 16.11.20] Available from: https://www.cadth.ca/resources/finding-evidence/press

[43] Vifor Pharma UK, Initiate Consultancy Ltd. *Systematic Literature Review of chronic kidney disease-associated pruritus (CKD-ap): search strategy document [Word doc provided by the company]*, 2022 [accessed 11.8.22]

[44] Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. *Value Health* 2013;16(4):686-95.

[45] Papaioannou D, Brazier JE, Paisley S. *NICE DSU Technical Support Document 9: The identification, review and synthesis of health state utility values from the literature*. Sheffield: Decision Support Unit, ScHARR, 2011. 64p. Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD9-HSUV-values_FINAL.pdf

[46] Molinier L, Bauvin E, Combescure C, Castelli C, Rebillard X, Soulie M, et al. Methodological considerations in cost of prostate cancer studies: a systematic review. *Value Health* 2008;11(5):878-85.

[47] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313(7052):275-83.

[48] Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol* 2010;63(12):1308-11.

[49] Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34.

[50] Soro M, Thokala P, Fotheringham J. POSC312. A methodological approach to assess the economic value of difelikefalin to treat chronic kidney disease associated pruritus (CKD-AP). *Value Health* 2022;25(1 Suppl):S209.

[51] Schaufler T, Munera C, Menzaghi F. PUK28. Emulating EQ-5D data from available pros to optimize clinical trial planning in chronic kidney disease-associated pruritus. *Value Health* 2019;22(Suppl 3):S918.

[52] Schaufler T, Munera C, Menzaghi F. PUK31. Estimating the impact of treating chronic kidney disease-associated pruritus on quality of life in the absence of direct utility measures. *Value Health* 2019;22(Suppl 3):S919.

[53] Rehman IU, Ahmed R, Rahman AU, Wu DBC, Munib S, Shah Y, et al. Effectiveness and safety profiling of zolpidem and acupressure in CKD associated pruritus: an interventional study. *Medicine* 2021;100(21):e25995.

[54] Millington GWM, Mustapa MF, Levell NJ, Collins A, Lovell CR, Leslie TA, et al. British Association of Dermatologists' guidelines for the investigation and management of generalized pruritus in adults without an underlying dermatosis, 2018. *Br J Dermatol* 2018;178(1):34-60.

[55] UK Renal Registry. *UK Renal Registry 23rd Annual Report - data to 31/12/2019 [Internet]*. Bristol: The Renal Association, 2021 [accessed 11.8.22] Available from: https://ukkidney.org/audit-research/annual-report/23rd-annual-report-data-31122019

[56] Fishbane S, Wen W, Munera C, Lin R, Bagal S, McCafferty K, et al. Safety and tolerability of difelikefalin for the treatment of moderate to severe pruritus in hemodialysis patients: pooled analysis from the phase 3 clinical trial program. *Kidney Med* 2022;4(8):100513.

[57] The Renal Association. *Clinical practice guideline - haemodialysis [Internet]*, 2019 [accessed 11.8.22] Available from: https://ukkidney.org/sites/renal.org/files/FINAL-HD-Guideline.pdf

[58] UK Renal Registry. *UK Renal Registry 24th Annual Report - data to 31/12/2020 [Internet]*. Bristol: The Renal Association, 2022 [accessed 10.10.22] Available from: https://ukkidney.org/audit-research/annual-report

[59] Vifor Pharma UK, Initiate Consultancy Limited. *Cost-effectiveness analysis of difelikefalin (DFK) for the treatment of CKD-aP. Version 1. Submission model. 09.08.2022. Version 1.1. Submision model - external link and DSA amendment. 05.09.2022 [Excel]*: Initiate, 2022 [accessed 20.9.22]

[60] National Institute for Health and Care Excellence. *Renal replacement therapy and conservative management. NICE guideline 107 [Internet]*. London: National Institute for Health and Care Excellence 2018 [accessed 11.8.22] Available from: https://www.nice.org.uk/guidance/ng107

[61] Vifor Pharma UK, School of Health and Related Research (ScHARR). *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Appendix J: Developing EQ-5D-3L mapping models in patients on haemodialysis: Comparisons across different pruritus patient reported outcome measures*. Sheffield: ScHARR, University of Sheffield, 2021 [accessed 11.8.22]

[62] Lee AJ, Morgan CL, Conway P, Currie CJ. Characterisation and comparison of health-related quality of life for patients with renal failure. *Curr Med Res Opin* 2005;21(11):1777-83.

[63] National Institute for Health and Care Excellence. *Dapagliflozin for treating chronic kidney disease. NICE technology appraisal guidance 775 [Internet]*. London: National Institute for Health and Care Excellence 2022 [accessed 10.10.22] Available from: https://www.nice.org.uk/guidance/ta775

[64] Thokala P, Fotheringham J, Soro M. POSC312. A methodological approach to assess the economic value of difelikefalin to treat chronic kidney disease associated pruritus (CKD aP). Poster presented at ISPOR Europe 2021; 30 Nov-3 Dec 2021; Virtual [PDF provided by company]. 2021.

[65] Committee for Medicinal Products for Human Use (CHMP). *Kapruvia - Assessment Report (EMA/152430/2022)*. Amsterdam: European Medicines Agency, 2022 [accessed 11.8.22] Available from: https://www.ema.europa.eu/en/documents/assessment-report/kapruvia-epar-public-assessment-report_en.pdf

[66] Joint Formulary Committee. British National Formulary (BNF) [Internet] London: BNF Publications, 2022 [accessed 14.10.22]. Available from: https://bnf.nice.org.uk/

[67] NHS England. National Cost Collection for the NHS: 2020/21[Internet]. [accessed 12.10.22]. Available from: https://www.england.nhs.uk/national-cost-collection/

[68] Jones K, Burns A. *Unit costs of health and social care 2021 [Internet]*. Canterbury: Personal Social Services Research Unit, University of Kent, 2021 [accessed 11.8.22]. 237p. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-of-health-and-social-care-2021/

[69] Initiate Consultancy Limited, Vifor Pharma UK. *Chronic kidney disease-associated pruritus (CKD-AP) modified delphi*: Initiate Consultancy Limited, 2022 [accessed 11.8.22]

[70] NHS Blood and Transplant. NHS Blood and Transplant fact sheet 7 [As referenced in the CS]. NHS Blood and Transplant, 2009.

[71] Vifor Pharma UK, Initiate Consultancy Limited. *Cost-effectiveness analysis of difelikefalin (DFK) for the treatment of CKD-aP (ID3890 CE Model PAS price__clarificationquestions) [Excel file updated following clarification questions]*, 2022 [accessed 20.9.22]

[72] Jones K, Burns A. *Unit Costs of Health and Social Care 2021*. Personal Social Services Research Unit Kent, UK, 2021. 185p.

[73] Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-adjusted probability of being cost effective. *Pharmacoeconomics* 2019;37(9):1155-63.

[74] McNamara S, Schneider PP, Love-Koh J, Doran T, Gutacker N. Quality-adjusted life expectancy norms for the English population. *Value Health* 2022;Online ahead of print.

[75] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

[76] Vifor Pharma UK. Difelikefalin for treating pruritus in people having haemodialysis [ID3890]. EAG report – factual accuracy check and confidential information check. 2022 [accessed 3.11.22].

Appendix 1: EAG search strategies

Embase (Ovid): 1974-2022/09/26
Searched 28.9.22

Search A) Difelikefalin

1 Difelikefalin/ or (ckd 943or ckd943 or cr 845 or cr845 or Difelikefalin or fe 202845 or fe202845 or korsuva or mr 13a9 or mr13a9).af. (149)

Search B) Pruritus + Hemodialysis + RCTs/Obs

1 exp Pruritus/ or exp antipruritic agent/ (170412)
2 (Itch$ or pruritis or pruritus or prurigo).ti,ab,ot. (58882)
3 (antiprurigo or antipruritic$ or antipruritus or "counterirritant agent").ti,ab,ot. (1136)
4 Gabapentin/ or (dineurin or dm 1796 or dm 5689 or dm1796 or dm5689 or gabalept or gabaliquid or geriasan or gabapen or gabatin or gantin or go 3450 or go3450 or goe 3450 or goe3450 or gralise or kaptin or keneil or neurontin or neurotonin or nupentin or sefelsa or serada).ti,ab,ot. (35663)
5 ("μ-receptor antagonists" or "κ agonists").ti,ab,ot. (191119)
6 or/1-5 (408858)
7 Hemodialysis/ or hemodialysis patient/ or (Hemodialysis or haemodialysis or hemodialyse or hemorenodialysis or hemotrialysate).ti,ab,ot. (173557)
8 ((blood or center or centre or department or unit$ or extracorporeal or patient$ or renal) adj3 dialysi?).ti,ab,ot. (65888)
9 or/7-8 (201301)
10 6 and 9 (3023)
11 Clinical study/ (160452)
12 Case control study/ (193188)
13 Family study/ (25692)
14 Longitudinal study/ (178735)
15 Retrospective study/ (1312962)
16 Prospective study/ (797464)
17 "randomized controlled trial (topic)"/ (235240)
18 16 not 17 (788060)
19 Cohort analysis/ (899840)
20 (Cohort adj (study or studies)).mp. (422254)
21 (Case control adj (study or studies)).tw. (158015)
22 (follow up adj (study or studies)).tw. (70422)
23 (observational adj (study or studies)).tw. (227065)
24 (epidemiologic$ adj (study or studies)).tw. (117548)
25 (cross sectional adj (study or studies)).tw. (302748)
26 or/11-15,18-25 (3565865)
27 Randomized controlled trial/ (729886)
28 Controlled clinical study/ (467140)
29 random$.ti,ab. (1838689)
30 randomization/ (95175)
31 intermethod comparison/ (287723)
32 placebo.ti,ab. (347218)
33 (compare or compared or comparison).ti. (575316)

34 ((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab. (2574168)
35 (open adj label).ti,ab. (100738)
36 ((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab. (261075)
37 double blind procedure/ (199152)
38 parallel group$1.ti,ab. (30124)
39 (crossover or cross over).ti,ab. (118312)
40 ((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant$1)).ti,ab. (389121)
41 (assigned or allocated).ti,ab. (458814)
42 (controlled adj7 (study or design or trial)).ti,ab. (419313)
43 (volunteer or volunteers).ti,ab. (271661)
44 human experiment/ (594923)
45 trial.ti. (370766)
46 or/27-45 (5917856)
47 (random$ adj sampl$ adj7 ("cross section$" or questionnaire$1 or survey$ or database$1)).ti,ab. not (comparative study/ or controlled study/ or randomi?ed controlled.ti,ab. or randomly assigned.ti,ab.) (9127)
48 Cross-sectional study/ not (randomized controlled trial/ or controlled clinical study/ or controlled study/ or randomi?ed controlled.ti,ab. or control group$1.ti,ab.) (322012)
49 (((case adj control$) and random$) not randomi?ed controlled).ti,ab. (20281)
50 (Systematic review not (trial or study)).ti. (223256)
51 (nonrandom$ not random$).ti,ab. (18121)
52 "Random field$".ti,ab. (2784)
53 (random cluster adj3 sampl$).ti,ab. (1467)
54 (review.ab. and review.pt.) not trial.ti. (1027708)
55 "we searched".ab. and (review.ti. or review.pt.) (43848)
56 "update review".ab. (124)
57 (databases adj4 searched).ab. (53836)
58 (rat or rats or mouse or mice or swine or porcine or murine or sheep or lambs or pigs or piglets or rabbit or rabbits or cat or cats or dog or dogs or cattle or bovine or monkey or monkeys or trout or marmoset$1).ti. and animal experiment/ (1167135)
59 Animal experiment/ not (human experiment/ or human/) (2449901)
60 or/47-59 (4056397)
61 46 not 60 (5236440)
62 26 and 10 (540)
63 61 and 10 (757)
64 62 or 63 (1134)

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by the end of **2 November 2022** using the below comments table.

All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and separately highlight information that is submitted as [redacted]' in turquoise, all information submitted as [redacted]' in yellow, and all information submitted as '[redacted]' in pink.

Issue 1 Inappropriate reporting of Key Issue 3.

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Key Issue 3 Inappropriate reporting of issue. Page 16; Table 1.4 The EAG report that a SLR was not carried out for clinical effectiveness, and note in the expected effect on the cost-effectiveness estimates that: "*The cost effectiveness (CE) might have been spuriously increased".*	This should instead be reported as: *"The direction of the impact (if any) on the cost effectiveness estimates is unknown".*	The EAG have implied that by not conducting a clinical SLR, the company may have knowingly reported a more favourable ICER. This is incorrect. The impact of not conducting a clinical SLR can only be described as unknown.	Whilst the EAG does not agree that this is a factual inaccuracy, the text has been amended to clarify that an increase or decrease in cost-effectiveness may be possible because of the absence of an SLR.

Issue 2 Incorrect interpretation and reporting of Key Issue 5

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Key Issue 5 Page 17; Table 1.16 The EAG report in the expected effect on the cost-effectiveness estimates that: "*The cost effectiveness (CE) might have been spuriously increased"* . Page 50 and 53 The EAG state that: *"the company mentioned age, sex and race as potential effect modifiers in their original CS (see page 22 in Document B, CS"* Page 67 In reference to statistical analysis done on the trials, the EAG state that:	The company do not report age, sex and race to be effect modifiers. The reference paragraph is: *"People in lower socio-economic groups are more likely to develop chronic kidney disease, progress towards kidney failure, and die earlier with CKD. People from black, Asian and minority ethnic populations are more likely to progress to kidney failure faster and less likely to receive a transplant."* This states that race is a prognostic factor in the progression of CKD, rather than CKD-aP, which is irrelevant to the decision problem. Furthermore, the EAG do not correctly report the subgroup analysis, as there was no statistically significant difference	As noted by EAG, uncertainty in subgroup estimates is to be expected as clinical trials are not designed or powered to robustly measure between subgroup effects. As such, nominal differences in treatment effect are expected between subgroups. There is no significant difference in treatment effect (>= 3 point improvement in WI-NRS at week 12) when stratifying KALM-1 & 2 data by patient race and as such no convincing evidence for an impact on treatment effect if generalised to a population with fewer black people. Furthermore, that the point estimates happen to be higher in Black patients in both trials in contrast with	In Key issue 5 (Table 1.6) regarding the expected effect on the cost effectiveness estimates, the EAG statement is as follows: "*The cost effectiveness (CE) is likely to have been spuriously increased.*" The EAG notes that the company's entire statement about equality considerations in the CS (pages 21-22) is as follows (emphasis added here by the EAG): "*People in lower socio-economic groups are more likely to develop chronic kidney disease, progress towards kidney failure, and die*

"The EAG is surprised that the prognostic factor of race was not considered, given that it was suggested, pre-hoc, as a potential effect modifier by the company , and was later shown by a sub-group analysis to have an effect on outcome. Page 73 *"Based on the sub-group analysis results, being Black or African American improves outcome with difelikefalin relative to placebo (see Section 3.2.1.1)."* Page 126 *"As Black participants were shown on a sub-group analysis to have a better response to difelikefalin than other ethnic groups, this discrepancy may lead to overestimation of the efficacy of difelikefalin in the UK target population."*	in treatment effect associated with race in any subgroup analyses. All noted references to previous identification of race as a potential effect modifier should be removed, as well as all references that describe race as a confirmed effect modifier within the KALM-1 & 2 trials.	age and sex is not convincing of significant treatment effect modification associated with race.	*earlier with CKD. People from black, Asian and minority ethnic populations are more likely to progress to kidney failure faster and less likely to receive a transplant. Women are more likely to be diagnosed with CKD, but less likely to start dialysis. Older people with CKD are less likely to receive a kidney transplant than their younger counterparts.* ***These populations are at greater risk of developing CKD-aP and experiencing symptoms for longer while on dialysis.*** *Therefore, guidance on the use of difelikefalin could have a different impact on people with protected characteristics compared to the wider*

			population (1). Difelikefalin is also restricted for in-centre haemodialysis use only, which may be considered to represent a barrier to some patients for whom in-centre haemodialysis is less accessible." The CS clearly acknowledges these populations to be at greater risk of developing CKD-aP. Therefore, the EAG's statement is not a factual inaccuracy. In addition, as stated in the EAG report "*Effect modification may be inferred from sub-group analysis.*" (page 52). The subgroup analysis presented in Topf et al. (2022) substantiates this statement.

			The EAG did not state that there was a statistically significant difference in treatment effect associated with race in a subgroup analysis. This is not a factual inaccuracy. The company's overall concerns around uncertainty and generalizability connected to the results of the subgroup analysis are covered and noted within the EAG's report. For example, the EAG comments in section 3.2.1 state that, "*Whilst the EAG accept the uncertainty in the sub-group estimates in Table 3.9 and realise that it cannot be definitively concluded that race is an effect modifier, the EAG believe there is enough*

			evidence to suggest consideration of this point by the Committee." (p. 53). Related to the above, the EAG suggests that a lack of power in the sub-group analyses means that type II errors are more likely; that is, that estimates of between-group differences that suggest no statistically significant difference do not necessarily indicate no difference. The EAG therefore maintain that any apparent differences in point estimates observed in sub-group testing should be viewed as potential true differences, even if this cannot be definitively confirmed on statistical testing. In summary, the lack of power

			increases the need for vigilance about attention to possible sub-group differences, rather than diminishing it.

Issue 3 Reference to ICHD restriction

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Decision problem Page 23; Table 2.1 EAG state: *"The company's decision to narrow the population to those having ICHD use appears to be sensible, if difelikefalin is 'restricted to in-centre haemodialysis use only'. However, there are no references in the company submission (CS) to back up the statement that difelikefalin is restricted for ICHD use only. "*	This is incorrect as several references are made to the SmPC which states that DFK be restricted to ICHD only.	As per company submission. The EAG also recognise that this is stated in the SmPC later in the document.	Amended.

Issue 4 Incorrect statements

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 126 In section 3.3 in response to the critique of trials, the EAG states: *No ITC was carried out, despite the clear need.*	The statement should be changed to: "No ITC was conducted by the company".	Inappropriate conclusion. The company argue that no clear need has been established and this sentence should be altered to only reflect that no ITC was conducted. Furthermore, cost-effectiveness analysis is based on clinical trial data providing the best possible direct evidence of relative outcomes for patients treated with DFK in addition to ECM in comparison with ECM alone, aligned to treatment positioning in the CS.	Not a factual inaccuracy. The EAG has demonstrated the feasibility and necessity of an ITC analysis throughout the report (e.g., see Table 1.3 and Section 3.2 of the EAG report).
Page 126 In section 3.6, in reference to the KALM trials, the EAG state:	This statement should be removed and further clarification on the trials allocation concealment methodology can be requested.	Inappropriate conclusion. Overall methodological quality of the trials is not impacted by lack of understanding of treatment allocation processes. Additional detail	Not a factual inaccuracy. The EAG assessment of methodological quality was based on information in the CS.

"The methodological quality of the RCTs was compromised by poor reporting of allocation concealment." Further information on the allocation concealment within the KALM trials may be able to be provided. Suggested lack of detail in reporting does not compromise the methodological quality of the trials.		describing allocation concealment can be requested.	

Issue 5 Critique of SLRs informing model

Description of problem	Description of proposed amendment	Justification for amendment	EAG comments
Page 138, In their critique of the companies SLRs, the EAG state that the utilities SLR (SLR 2) or the economic modelling SLR (SLR 4) "*did not include the published report on the KALM-1 study,15 nor the report on the pooled analysis of data from the*	Text to be removed.	The company do not agree that this is an appropriate conclusion as the KALM studies did not report generic HRQoL measures or resource use/cost estimates. Therefore, it is not expected that an SLR focusing on utilities or economic models	Last bullet point on page 138 of EAG report was amended.

KALM-1 and KALM-2 RCTs and OLE phases." The EAG then follow this point and state that: *"As the SLRs did not identify all the papers relevant to the submission, it is possible that the review methods were suboptimal, and the possibility of study selection bias cannot be discounted."*		would identify these papers for inclusion.	

Issue 6 EAG Subgroup Analysis

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>EAG comment</th></tr>
<tr>
<td>Page 179,
The EAG report the results of the subgroup analysis based on only patients with severe and very severe at the start of treatment.

This was not possible to be conducted by the EAG as the direct transition matrices for</td>
<td>The company have updated the EAG model and re-run the scenario using the appropriate transition matrices.
The results are as follows:
<table>
<tr><th rowspan="2">Preferred assumption</th><th colspan="2">DFK plus ECM</th><th colspan="2">ECM</th><th rowspan="2">Incr. Costs (£)</th><th rowspan="2">Incr. QALYs</th><th rowspan="2">ICER (£/QALY)</th></tr>
<tr><th>Costs (£)</th><th>QALYs</th><th>Costs (£)</th><th>QALYs</th></tr>
<tr><td>EAG's preferred base case</td><td>[redacted]</td><td>[redacted]</td><td>£97,611</td><td>2.88</td><td>[redacted]</td><td>[redacted]</td><td>£35,048</td></tr>
<tr><td>EAG's model, severe and very severe only</td><td>[redacted]</td><td>[redacted]</td><td>£97,742</td><td>2.83</td><td>[redacted]</td><td>[redacted]</td><td>£30,274</td></tr>
</table>
</td>
<td>Required correction.</td>
<td>We thank the company for providing this subgroup analysis, with all transition matrices based on the specified subgroup.
This table has been added to the EAG report as Table 6.5.</td>
</tr>
</table>

this population were not provided to the EAG.			

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG comment
Give full details of inaccurate marking - document title and page number	Give details of incorrect confidential marking	Please copy the impacted section here, with your amended marking.	
Page 12 and 149	Title for Fig 4.3 to be marked up CIC	[redacted]	Amended
Page 12, 171 and 172	Titles for Figs 5.1 – 5.3 to be marked up AIC	[redacted]	Amended

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is the end of on **16 December 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	Garth Baxter
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	CSL Vifor
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	Nothing to disclose

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
1. The population in the decision problem and the included trials appears narrower than that in the National Institute for Health and Care Excellence (NICE) final scope. The decision problem and trial populations preclude first line treatment and are restricted to people receiving in-centre haemodialysis (ICHD). The NICE scope makes	No	For clarity, the positioning of difelikefalin can be split into two parts: 1) The licensed indication which is for adults with moderate-to-severe CKD-aP restricted to in-centre haemodialysis only. 2) The alignment with current UK clinical practice which is where established clinical management is insufficient in reducing pruritus Regarding the alignment with current UK clinical practice, as stated in the company submission, there are currently no approved treatments for CKD-aP. When discussed with clinicals in a modified Delphi, [redacted] Clinical advisors highlighted that treatment is additive and as such, difelikefalin would be used as an adjunct to first line treatments in patients whose itch persists despite the use

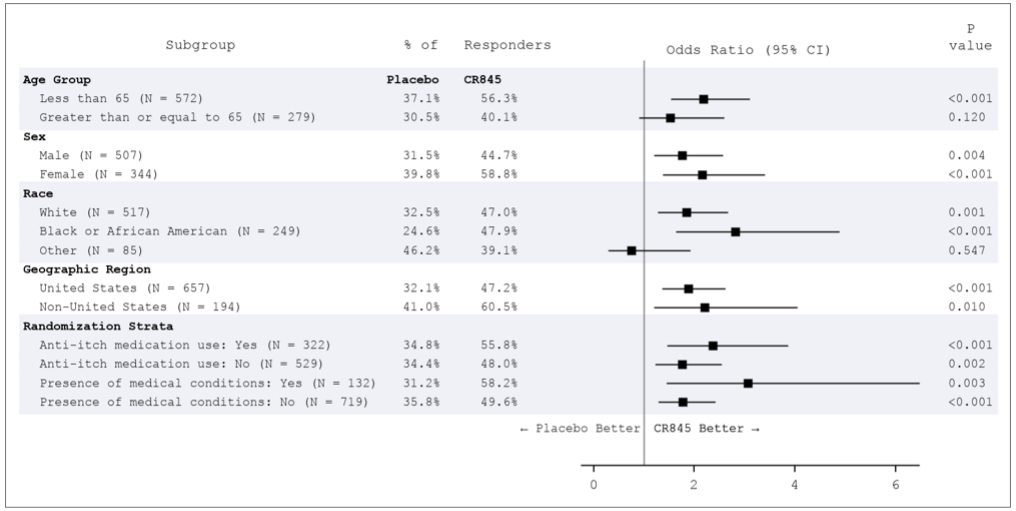
no restrictions in terms of ICHD and treatment line.		of current treatments. The company are not seeking a recommendation as a first line treatment and have used the term established clinical management to reflect that first line treatments may vary across the UK. The company agree with the EAGs statement that the defined population in the decision problem is narrower than the NICE scope but that it is: 1) consistent with the license for difelikefalin, and 2) is reflective of current UK clinical practice.
2. The comparison in the included trials is difelikefalin plus established clinical management (ECM) versus placebo plus ECM, whereas the comparison in the NICE final scope and the decision problem is difelikefalin versus ECM. The nature of the treatment comparison in the trials may lead to a more optimistic impression of the study drug's benefits compared with the NICE final scope/decision problem.	Yes *Updated subgroup analysis*	The company accept the comparison presented in the company submission is different to the NICE scope. The company highlight that this is for two reasons: 1) The key clinical data evidencing the effect of difelikefalin is versus placebo in patients with and without anti-itch medication, and 2) As stated in the answer for key issue 1, the anticipated place in therapy is difelikefalin with established clinical management. In the company submission, subgroup analyses were presented for patients with and without anti-itch medication at baseline in the KALM trials. These subgroup analyses have been incorporated into the company's updated economic model and the EAG preferred basecase and presented below. **The cost-effectiveness results are consistent when considering only patients with and without anti-itch medication at baseline.** This is also consistent with the subgroup analysis for the odds of achieving a ≥3-point reducing in WI-NRS at week 12 as presented in Topf et al. 2022 and Figure 12 below.
3. A systematic literature review (SLR) of clinical effectiveness evidence	Yes	The company have conducted a clinical effectiveness SLR and attached the results as a separate report. In summary, the SLR resulted in the inclusion of 30 publications that link to seven clinical trials (four phase III trials and three phase II trials). All the relevant

was not carried out. This made it difficult to determine whether all relevant studies were included in the clinical effectiveness part of the submission. The Evidence Assessment Group (EAG) identified two potentially relevant randomised controlled trials (RCTs) that had not been considered within the company submission (CS). Data from these RCTs have been added to the report by the EAG.		clinical studies included for full review were conducted by the originator company Cara Therapeutics or it's Japanese licensee Kissei Pharmaceutical and were identified in the company's responses to the clarification questions. The company would like to thank the EAG for including details of two of the Phase II trials in its report and agrees with the EAG's conclusion that the 8-week data from these studies should not been pooled with the 12-week data from the KALM RCTs because of the relatively large difference in follow-up time that might have an impact on outcome. **The results of the clinical SLR have no impact on the cost-effectiveness estimates presented by the company or EAG.**
4. Differences between ECM in the included trials and the United Kingdom (UK) target population may limit the generalisability of clinical effectiveness evidence from the trials. The company did not provide the results of sub-group analyses in relation to specific anti-itch	No	Please see response for Key issues 1 and 2. **The company would like to highlight that there is no current standard of care for CKD-aP. This was highlighted in the SLR for clinical guidelines and confirmed by clinical opinion in the modified Delphi.** A subgroup analysis of anti-itch medications has been provided by the company in Appendix S (subgroups included those receiving general anti itch medication, steroids, gabapentin/pregabalin, opioids or antihistamines versus patients not receiving those medications). The company agrees with the EAG conclusion that no significant difference was found between groups receiving anti-itch medication versus those not. The results of the pooled analysis of KALM-1 and KALM-2 reported by Topf et al., 2022 also show that there is no significant difference in the mean effect of difelikefalin when stratified by anti-itch medication use at baseline.

medications other than difelikefalin. This hindered the evaluation of the impact of differences in the use of non-difelikefalin anti-itch medications between the included trials and the UK target population.		The requested sub-group analysis by specific anti-itch medication at a class level is unlikely to aid the further evaluation of the generalisability as there is no clinical basis to suggest that 1) use of anti-itch medication is consistent across UK clinical practice, and 2) one medication out of a sub-group of treatments (e.g. antihistamines) is preferred over another. Moreover, given the large variation of medication used in the KALM trials, any such analysis would be informed by very small observation numbers and subject to high levels of uncertainty. **The company believe that it is therefore inappropriate to conduct this sub-group analysis. Results from the Delphi panel further demonstrate the generalisability of the KALM trials to UK clinical practice.**
5. The included trials recruited a larger proportion of Black participants relative to those seen in the UK target population. Results from sub-group analyses suggested that Black participants tend to have better difelikefalin outcomes than other ethnic groups. This may further affect the generalisability of the overall trial results.	No	There is no evidence in our studies or in the literature of race being a prognostic factor for itch in patients with ESRD. However, black patients are at increased risk of CKD progression, and as such are over-represented in the ESRD population in comparison with the general population. Similarly, as CKD-aP is a common complication of ESRD, it should also be expected that Black patients are more likely to experience CKD-aP as a consequence of increased risk of ESRD and CKD progression more generally. However, this does not equate to either an expectation that Black patients in receipt of HD are more likely to experience CKD-aP than other ethnicities also in receipt of HD, or that Black participants of the KALM trials would be expected to experience a different relative treatment effect than participants of other ethnicities. This position is supported by the patient organisation submission from Kidney Research UK who state “kidney disease disproportionally affects people from deprived communities and ethnic minority groups and people in these cohorts progress faster to end stage renal failure. Evidence shows that fewer kidney patients from deprived communities are treated with peritoneal dialysis, with more treated with haemodialysis. There are therefore likely to be proportionally more people from these cohorts on haemodialysis, experiencing pruritus and likely to benefit from this treatment”. The sample size for the pooled data was determined to be more appropriate to ensure the validity of these analyses. For the pooled analysis, models were adjusted for study/region combined variable. In the forest plot included in the Integrated Summary of Efficacy

		(Figure 12 and Figure 13) and reproduced below, the difelikefalin drug effect is evident in both White and Black or African American patients; the confidence intervals overlap suggesting that effects are not different between the 2 groups. The drug effect in the "Other" race category is limited as the number of "other race" subjects was small (85 subjects) compared to the White (517 subjects) and Black or African American subgroups (249 subjects) and is composed of different race groups (Asians and Native Americans, for example). The difference in the LS means estimate of the proportion of black/African American and white patients responding to difelikefalin in the pooled analysis was 47.9% vs 47.0% respectively (Topf et al., 2022). The nonsignificant difference on odds ratios was due to a lower placebo response in black participants. **The results of the KALM clinical trials did not display a statistically significant difference in the treatment effect for participants of different ethnicities.** **This issue is suggesting discriminative effects which aren't present, therefore, the comparison of ethnicities is not appropriate.**

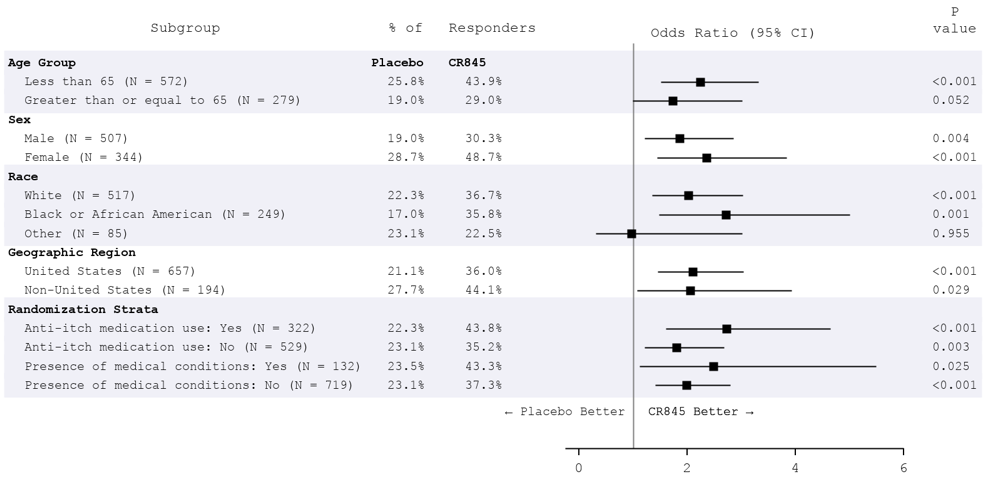
Figure 12: Subjects with ≥3-point Improvement from Baseline to Week 12 with Respect to the WI-NRS Score by Subgroup (Population: ITT- Pooled Dataset)



MAR = missing at random; MI = multiple imputation

Note: Estimated percentages, odds ratio, and confidence intervals were estimated based on a logistic regression with terms for treatment group, baseline WI-NRS score, region/study, use of anti-itch medication during the week prior to randomization, and the presence of specific medical conditions. For the analysis by geographic region and randomization strata, the corresponding subgroup variables were removed from the model, and Study ID was included in the model of United States. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.

Figure 13 Subjects with ≥4-point Improvement from Baseline to Week 12 with Respect to the WI-NRS Score by Subgroup (Population: ITT- Pooled Dataset)



MAR = missing at random; MI = multiple imputation

Note: Estimated percentages, odds ratio, and confidence intervals were estimated based on a logistic regression with terms for treatment group, baseline WI-NRS score, region/study, use of anti-itch medication during the week prior to randomization, and the presence of specific medical conditions. For the analysis by geographic region and randomization strata, the corresponding subgroup variables were removed from the model, and Study ID was included in the model of United States. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.

6. The rationale for the statistical analysis in the included trials (specifically, multiple imputation (MI) and	No	A multiple imputation (MI) procedure was chosen for the treatment of missing data as single imputation methods like Last Observation Carried Forward (LOCF) and Baseline Observation Carried Forward (BOCF) would not be adequate in this setting and would underestimate the variability of the results due to missing data. We observed in our Phase 2 studies (CLIN2005 and CLIN2101) that patients randomized to placebo do improve, on average, compared to baseline and therefore a BOCF imputation that assumes a patient's itch returns to baseline level after discontinuation would underestimate a

<table>
<tr><td>logistic regression) is not transparent. Lack of information about model inputs and outputs in both instances has hindered the EAG's assessment of the quality of the statistical analyses.</td><td></td><td>possible treatment effect. An LOCF imputation on the other hand may overestimate a treatment effect. The use of MI is also consistent with recommendations of the Panel on Handling Missing Data in Clinical Trials convened by the National Research Council of the National Academies. Data from the completed Phase 2 study CLIN2101 did not suggest any obvious missing data patterns. Therefore, the primary analysis of the pivotal Phase 3 studies KALM-1 and KALM-2 imputed missing weekly WI-NRS scores using an MI procedure under a Missing At Random (MAR) assumption. Under this assumption, patients who discontinued the double-blind treatment early would have similar weekly WI-NRS scores as other patients in their respective treatment group that have complete data. Since assumptions about missing data are difficult to assess, sensitivity analyses of the primary endpoint were added to evaluate the robustness of study results under different assumptions for missingness.

As presented in the table following Table 2 Key Issues, in general, results were consistent and support the efficacy of difelikefalin compared to placebo in the treatment of moderate to severe CKD-aP when counts are based on non-missing data.</td></tr>
<tr><td>7. Clinical effectiveness data from the KALM-1 and KALM-2 trials were pooled without adjusting for differences between the trials. This may have resulted in biased estimates of treatment effectiveness.</td><td>Yes
Updated scenario analysis</td><td>The company would like to apologise for a typo in clarification response A33. The pooled analysis was adjusted for a region/study combined covariate. Therefore, model was adjusted for differences between the trials.

The table below summarises the covariates adjustment for primary efficacy analysis:
<table>
<tr><th>Covariate</th><th>Multiple imputation KALM1</th><th>Logistic regression KALM1</th><th>Multiple imputation KALM2</th><th>Logistic regression KALM2</th><th>Logistic regression pooled data KALM1/ KALM2</th></tr>
<tr><td>Non-missing WI-NRS scores for each week</td><td>x</td><td></td><td>x</td><td></td><td></td></tr>
<tr><td>Treatment group</td><td></td><td>x</td><td></td><td>x</td><td>x</td></tr>
</table></td></tr>
</table>

Baseline WI-NRS score	x	x	x	x	x
Use of anti-itch medication during the week prior to randomization	x	x	x	x	x
Presence of specific medical conditions	x	x	x	x	x
Region			x	x	
Region/study combined*					**x**

* As all sites in CLIN3102 were in the United States, a single model parameter will be used to account for the effects of region and study that will have 3 levels: CLIN3102 United States, CLIN3103 United States, and CLIN3103 ROW (Western Europe/ Western European Origin [Canada, United Kingdom, Germany, Australia, New Zealand], Eastern Europe [Poland, Hungary, Romania, Czech Republic], Asia [Taiwan, South Korea]).

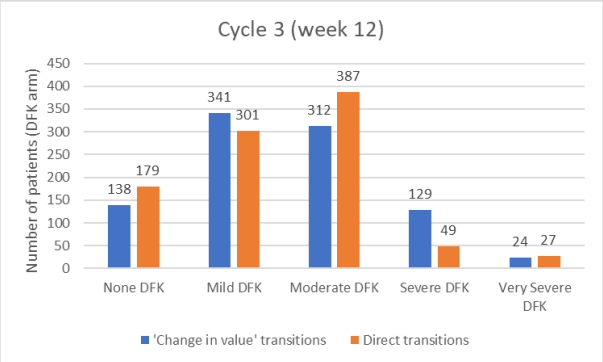
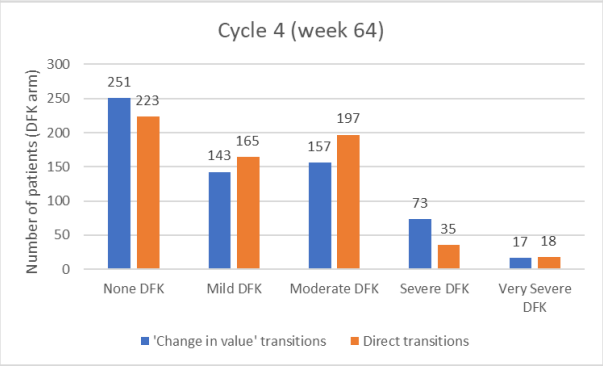
Please find below frequency table of the region/study combined variable included in the pooled analysis:

Region/study combined variables (pooled analysis)	Frequency	Percent
United States/KALM1	378	44,42%
Rest of World/KALM2	194	22,80%
United States/KALM2	279	32,78%

For the model, merged patient level data and mean count estimates were used. Clinical data pooling was done by Cara therapeutics but not relied upon for modelling – this was conducted separately. The pooled data would therefore only be relevant if it used point estimates, and this is not included in modelling estimates.

For further reference, in their report (page 89-90), the EAG conclude that a more valid pooled result for the odds of achieving a ≥3-point reducing in WI-NRS would be an OR (95% CI) of 2.07 (1.24 to 3.45) as compared with an OR (95% CI) of 1.93 (1.44 to 2.57)

		as reported by Topf et al., 2022. This suggests that any potential bias in estimated treatment effect may underestimate the benefit of difelikefalin. In the company submission, scenario analyses were presented using patient-level data taken from each trial independently. **The cost-effectiveness results are consistent when trial data are used individually and pooled.**
8. The company assumed in the base case and an alternative scenario that transitions can be modelled by only looking at the probability of shifting between one, two or three health states up or down, regardless of the current health state. This assumption did not seem to be supported by the directly estimated transition probabilities.	No	The company understand the EAGs preference for using the directly estimated transition probabilities, however, still note that there is uncertainty associated with both methods. The company would like to draw attention to the probabilistic results of the EAG preferred base case which were materially higher than the EAG deterministic base case results (£41,157 versus £35,048). The EAG report that *this is due to the skewness in the distribution around the transition probabilities whenever these are very close to zero, i.e., a 0% transition in the deterministic analysis will become a small but non-0% transition in the PSA.* This concern was raised by the company at submission wherein creating a matrix which calculates the probability of moving from any one state to each of the other states can result in small observation numbers estimating a single probability value, which may lead to unrealistic outcomes. Further, the company disagree with the EAG statement that in the company's base case, the assumption that transitions can be modelled by only looking at the probability of shifting between one, two or three health states up or down, regardless of health state did not seem to be supported by the directly estimated transition probabilities. At clarification the company presented the distribution of patients in the DFK arm at week 12 and week 64 when using the 'change in state' transition estimates (base case) and the directly observed transitions as estimated from the pooled patient level data for the KALM trials (figures presented below for reference). **The EAG agreed that results were quite similar and that they had no major concerns about this evidence.** The company also note that although it is implied that the rate of response to treatment is averaged across the population, by estimating treatment response by CKD-aP severity at baseline, the average treatment response is weighted by the distribution of patients at baseline (i.e. the number of patients with moderate, severe, or very severe CKD-aP). The limitation of this is that as the numerical benefit of treatment with difelikefalin was larger in

		more severe patients, the rate of transitions from more severe states to less severe states may be underestimated whilst the rate of transitions from less severe states to more severe states may be overestimated. **However, as seen in the validation results presented at clarification** (and below for reference) **this limitation has little impact on the overall movement of patients across health states.** **The company believe that the 'change in state' transition estimates (base case) are still the most appropriate for decision making.**  

<table>
<tr><td></td><td></td><td></td></tr>
<tr><td>9. In the estimation of the transition probabilities, use has been made of data that was imputed multiple times to account for missing data. It is unclear to the EAG how all transition matrices were derived in light of the multiple imputation.</td><td>Yes
Additional scenario analysis</td><td>Multiple imputation was used to estimate missing weekly values in the 5-D Itch total score for patients included in the KALM-1 and KALM-2 trials. Missing values were estimated based on patient demographics including age, sex, presence of type 2 diabetes mellitus, duration of ESRD, and duration of HD and treatments received including concomitant anti-pruritic medications and study treatment arm, as well as non-missing 5-D Itch total scores. Missing values were generated through predictive mean matching, where mean predictions from regression models based on chained equations are matched to the closest observation from the data, with that matched value being imputed. This approach ensures imputed values are consistent with the observed data, and that no observations are imputed outside possible ranges for 5-D Itch total score.
Additional scenario analyses have been conducted without MI to account for missing data and presented below in the company’s updated basecase and EAG preferred basecase.
For reference, the number of data observations included in the analysis at each model cycle from the pooled KALM-1 and KALM-2 data are presented here (Table 46 of the company submission).

<table>
<tr><th>5-D Itch Scale total scores</th><th colspan="2">Difelikefalin</th><th colspan="2">Placebo</th></tr>
<tr><th>Cycle</th><td>Observed only</td><td>Missing data imputation</td><td>Observed only</td><td>Missing data imputation</td></tr>
<tr><td>Baseline count</td><td>393</td><td>393</td><td>403</td><td>403</td></tr>
<tr><td>Cycle 1 (baseline to Week 4)</td><td>356</td><td>393</td><td>371</td><td>403</td></tr>
<tr><td>Cycle 2 (Week 4 to Week 8)</td><td>333</td><td>393</td><td>357</td><td>403</td></tr>
<tr><td>Cycle 3 (Week 8 to Week 12)</td><td>330</td><td>393</td><td>359</td><td>403</td></tr>
</table>
</td></tr>
</table>

<table>
<tr>
<td></td>
<td></td>
<td>
<table>
<tr><td>Cycle 4 (Week 12 to Week 64)</td><td>74</td><td>279</td><td>N/A</td><td>N/A</td></tr>
</table>

In the company’s updated base case, the cost-effectiveness results are consistent. In the EAG base case the cost-effectiveness estimates have a more material increase (approx. £2,800/QALY). The company believe this is due to the method of transition matrix estimation as the reliance on a small number of observations informing the Cycle 4 transition is more heavily weighted, increasing the uncertainty of the outcomes.
</td>
</tr>
<tr>
<td>10. In the base case, treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In the absence of further real-world evidence to support the waning impact of ECM and/or the lack of waning over time with difelikefalin, the EAG considers this assumption uncertain.</td>
<td>Yes</td>
<td>
The company thank the EAG for considered the waning affect in their scenario analysis and accept their comment that real world data could provide information about the long-term disease development in patients with pruritus who receive established clinical management.
The company do not have any further real-world evidence to support the long-term progression of CKD-aP, however sought to review NICE TAs on atopic dermatitis that were identified in the companies expanded economic literature review.
For reference, the expanded SLR identified the following:
<table>
<tr><th>HTA ID</th><th>Title</th><th>Indication</th></tr>
<tr><td>HST17</td><td>Odevixibat for treating progressive familial intrahepatic cholestasis [ID1570]</td><td>Liver transplant</td></tr>
<tr><td>TA775</td><td>Dapagliflozin for treating chronic kidney disease [ID3866]</td><td>CKD</td></tr>
<tr><td>TA807</td><td>Roxadustat for treating anaemia in people with chronic kidney disease [ID1483]</td><td>CKD</td></tr>
<tr><td>TA534</td><td>Dupilumab for treating moderate to severe atopic dermatitis after topical treatments [ID1048]</td><td>Atopic dermatitis</td></tr>
<tr><td>TA681</td><td>Baricitinib for treating moderate to severe atopic dermatitis [ID1622]</td><td>Atopic dermatitis</td></tr>
<tr><td>TA814</td><td>Abrocitinib, tralokinumab or upadacitinib for treating moderate to severe atopic dermatitis [ID3960]</td><td>Atopic dermatitis</td></tr>
</table>
</td>
</tr>
</table>

The aim of the initial review of HTAs was to provide insight and guidance on the model structure and identify any relevant costs and utilities.

The company have conducted a second review of the 3 atopic dermatitis HTAs, in which itching (pruritus) is the most disruptive symptom affecting sleep and causing anxiety or depression (TA534 ACM1 slides). This was done to identify any relevant data regarding the long-term loss of clinical benefit on the relevant established clinical management (standard of care/best supportive care) treatment arm, following discontinuation from the clinical trial.

Following the review, it was identified that all 3 HTAs had used the same scenario analysis to reflect a reduction in clinical benefit of best supportive care that had originated in TA534. For clarity, the company assumed that in the best supportive care state, 25% of the benefit would be lost in year 2, 50% in year 3, 75% in year 4, and 100% in year 5 and beyond. These estimates were applied to adjust down the utility value for the best supportive care state which was used as a driver of benefit in the atopic dermatitis models. Therefore, where the health state utility was 0.80 during the trials and 0.66 at baseline, by the end of year 5, everyone in the best supportive care arm returned to the baseline utility (0.66) for the remainder of their time in the model. The table below provides a summary of this benefit loss and sensitivity analysis values that were also included:

	Proportion of patients losing benefit in BSC arm			
	Year 2	**Year 3**	**Year 4**	**Year 5+**
Base case	25%	50%	75%	100%
SA 1	82%	90%	94%	96%
SA 2	57%	82%	92%	97%

In TA534, the committee concluded that values for sensitivity analyses 1 and 2 were the most plausible analyses for decision-making.

		In TA681, the company presented results using the 2 committee-preferred sensitivity analysis from TA534, and the ERG presented no benefit waning on best supportive care. The committee concluded that the loss of clinical benefit for best supportive care over time is likely to be between the company and ERG base case. The committee concluded that the uncertainty associated with including a waning effect was acceptable. In TA814, after the first committee consultation, in which the committee asked for more exploration of waning, the EAG presented the scenario analysis accepted in TA534 and subsequently TA681. The scenario assumed that by year 5, 97% of people had returned to baseline utility (SA2). The committee concluded that there is uncertainty when modelling BSC waning without evidence of natural history of the disease or use of further treatments. It was difficult to identify a clear conclusion on whether the waning was accepted, however given that these treatments were given a positive approval at ACM2 following a negative approval at ACM1, it is likely that the waning factor was accepted. **Following a review of the 3 atopic dermatitis TAs, the company have conducted additional analysis on the long-term benefit of established clinical management to reflect the scenarios that were accepted in TA534, TA681 and TA814.** Further details are presented below. **Given the results of this scenario analysis, the company have subsequently updated the base case model to include a waning of 10% in the established clinical management arm.** The company believe that this is a highly conservative estimate of waning given the evidence from TA534, TA681 and TA814, and the additional scenarios presented.
11. The company applies an increased risk of death for patients in the moderate, severe and very severe health states of the	No	The company thanks the EAG for acknowledging the relationship between mortality and extreme pruritus (Page 154 of the EAG report) and understands the decision that there is uncertainty in linking a reduction in itching score to a reduction in mortality. **The company have updated its base case analysis to remove the mortality adjustment for more severe CKD-aP states.** For completeness, the company have also

chronic kidney disease-associated pruritis (CKD-aP) population, based on an observational study. The EAG considers the evidence presented not substantial enough to establish a causal relationship between pruritus and mortality of these patients.		included the costs of dialysis (as per the EAG model) but note that this has no impact on the incremental results.

Key issue 6 additional table of results

Primary efficacy analysis and associated sensitivity analyses: Subjects with ≥3-Point Improvement from Baseline at Week 12 With Respect to the Worst Itching Intensity Numerical Rating Scale Score (Populations: ITT and Per Protocol)

	CLIN3102		CLIN3103	
	Placebo (N = 189)	**CR845 (N = 189)**	**Placebo (N = 236)**	**CR845 (N = 237)**
Primary analysis (combined estimates) [1]				
Week 12				
Observed ≥3-point NRS improvement [2] - n (%)				
Yes	51 (30.9%)	82 (52.2%)	77 (37.2%)	95 (49.7%)
No	114 (69.1%)	75 (47.8%)	130 (62.8%)	96 (50.3%)
Missing	24	32	29	46
LS means estimate of percent with improvement [3]				
Percent (95% CI)	27.6% (20.2%, 36.6%)	51.0% (42.9%, 58.9%)	42.2% (32.5%, 52.5%)	54.0% (43.9%, 63.9%)
LH odds ratio (95% CI)		2.72 (1.72, 4.30)		1.61 (1.08, 2.41)
CHW *P* value		<.001		0.020
Sensitivity Analyses				
(1) Subjects who discontinued early as nonresponders [4]				

	CLIN3102		CLIN3103	
	Placebo (N = 189)	CR845 (N = 189)	Placebo (N = 236)	CR845 (N = 237)
N	189	189	236	237
LS mean percent with improvement (95% CI)	26.0% (19.0%, 34.5%)	44.6% (35.4%, 54.2%)	37.2% (27.8%, 47.6%)	43.7% (33.4%, 54.7%)
LH odds ratio (95% CI)		2.29 (1.46, 3.60)		1.31 (0.89, 1.94)
CHW *P* value		<.001		0.168
(2) Multiple imputation with MNAR assumption [4]				
N	189	189	236	237
LS mean percent with improvement (95% CI)	27.6% (20.2%, 36.4%)	47.0% (37.1%, 57.3%)	39.9% (30.6%, 50.1%)	50.7% (41.2%, 60.1%)
LH odds ratio (95% CI)		2.33 (1.47, 3.71)		1.55 (1.05, 2.28)
CHW *P* value		<.001		0.029
(3) Tipping point [4]				
N	189	189	236	237
Highest shift parameter without tipping	6.50	6.50	0.75	0.75
Percent with improvement (95% CI)	29.1% (21.5%, 38.1%)	42.8% (33.7%, 52.4%)	41.9% (32.0%, 52.4%)	52.1% (42.5%, 61.5%)
LH odds ratio (95% CI)		1.82 (1.16, 2.86)		1.51 (1.01, 2.25)
CHW *P* value		.009		0.044
Additional Analyses				
Per Protocol Population [4]				
N	169	163	213	205
LS mean percent with improvement (95% CI)	27.0% (19.1%, 36.6%)	50.4% (47.1%, 53.6%)	39.7% (29.7%, 50.7%)	52.0% (43.8%, 60.2%)
LH odds ratio (95% CI)		2.74 (1.71, 4.41)		1.65 (1.08, 2.51)
CHW *P* value		<.001		0.019
No CHW adjustment for interim analysis [4]				
N	189	189	236	237
LS mean percent with improvement (95% CI)	28.3% (21.0%, 37.1%)	50.9% (41.6%, 60.2%)	42.6% (33.4%, 52.3%)	53.4% (43.7%, 62.8%)
LH odds ratio (95% CI)		2.62 (1.68, 4.09)		1.54 (1.05, 2.27)
P value		<.001		0.027

CHW = Cui, Hung, Wang; CI = confidence interval; ITT = Intent-to-treat; LH = Lawrence, Hung; LS = least squares; NRS = Numerical Rating Scale; MNAR = missing not at random

[1] Primary analysis was performed on ITT population and use Multiple Imputations With Missing at Random Assumption

[2] Counts and percentages were based on non-missing data.

[3] Estimated percentage, odds ratio and P value used a logistic regression model with terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomization, and the presence of specific medical conditions. Missing values were imputed using MI under MAR missing data assumption for interim subjects and post-interim subjects separately.

[4] Analysis based on interim and post-interim subjects combined.

Note: Combined analysis used the separate interim and post-interim results to generate an adjusted overall estimate and P value using the LH/CHW methodology. See the Statistical Analysis Plan for full details.

Reference: CLIN3102, CLIN3103 CSRs

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Response
Additional question from EAG report (page 67): *The covariates used in the logistic regression analysis were: trial group, baseline WI-NRS score, baseline use of antipruritic medication, and history of prespecified medical conditions in KALM-1 plus region in KALM-2. The company was asked to provide the rationale and validity of using these variables and asked to discuss if other baseline characteristics were considered for use in the logistic regression models, such as gender, race and age. In response the company reiterated the response to the same question on the MI analysis in the previous comment. In a further inquiry on how the variables were selected in the model the company responded that "The variables to be included in the logistic model were specified a priori in the study protocol and statistical analysis plan. There was no additional selection or de-selection of variables in the model.". Similar to the MI analysis there is a noteworthy lack of rationale and justification on the conceptualization of the model. The EAG has highlighted this as a key issue (Aligned with key issue 6).*	During the design of the studies, variables that were thought to potentially affect treatment response were included as stratification factors (use of antipruritic medication, specific medical conditions) or covariates (baseline WI-NRS score). The integrated analysis in the ISE/SCE took the study/geographical region into account. While the demographic subgroups defined by age, sex and race were not identified pre-hoc as potential effect modifiers, the robustness of the treatment effects across subgroups were assessed in the ISE/SCE. As stated a priori in the ISE/SCE statistical analysis plan, these subgroup analyses were contemplated only in the integrated dataset and repeated the model used for the primary analyses in the individual studies. While enrolled subjects were required to have moderate to severe itch (WI-NRS > 4), the following table indicates that baseline itch levels were homogeneous for all subgroups.

Mean and 95%CI of Baseline WI-NRS by subgroup (Population: ITT)

Baseline WI-NRS	Pooled (N=851)
Overall	
Mean (95% CI)	7.17 (7.08, 7.27)
Median (Range)	7.13 (4.1, 10.0)
By Age in years	
<65 (n)	572
Mean (95% CI)	7.18 (7.06, 7.30)
Median (Range)	7.13 (4.1, 10.0)
≥ 65 (n)	279
Mean (95% CI)	7.13 (6.98, 7.32)
Median (Range)	7.13 (4.1, 10.0)
By sex	
Male (n)	507
Mean (95% CI)	7.09 (6.97, 7.22)
Median (Range)	7.00 (4.1, 10.0)
Female (n)	344
Mean (95% CI)	7.29 (7.14, 7.44)
Median (Range)	7.25 (4.1, 10.0)
By race	
White (n)	517
Mean (95% CI)	7.15 (7.03, 7.28)

Median (Range)	7.00 (4.1, 10.0)
Black or African American (n)	249
Mean (95% CI)	7.23 (7.06, 7.41)
Median (Range)	7.25 (4.3, 10.0)
Other (n)	85
Mean (95% CI)	7.11 (6.81, 7.42)
Median (Range)	7.00 (4.4, 10.0)

WI-NRS = Worst Itching Intensity Numerical Rating Scale.
CI = confidence interval; Range corresponds to minimum and maximum

Reference: CLS Vifor ad hoc table 29.1.1

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Company's base case ICER			**£23,277**
Key issue 10	The company applied a **5%** treatment waning effect for the ECM arm only.	The company applied a **10%** treatment waning effect for the ECM arm only.	£21,072
Key issue 11	The company included a mortality adjustment for the moderate, severe, and very severe CKD-aP health state, and excluded dialysis costs.	The company has removed this mortality adjustment and included dialysis costs.	£31,865
Company's updated base case	**Incremental QALYs:** [redacted]	**Incremental costs:** [redacted]	**£24,552**

Sensitivity analyses around revised base case

Sensitivity analyses are presented for the following key issues:

Key issue 2: Results presented separately for patients only **with** and **without** anti-itch medication at baseline in the KALM trials (update of subgroup analysis A and B in the company submission).

Model	Incremental costs	Incremental QALYs	ICER
Only patients receiving anti-itch medication at baseline Efficacy is updated to reflect the subgroup and established clinical management treatment costs are updated to reflect 100% use of treatment costs			
Company updated model	[redacted]	[redacted]	£25,339
EAG model	[redacted]	[redacted]	£33,991
No patients receiving anti-itch medication at baseline Efficacy is updated to reflect the subgroup and established clinical management treatment costs are updated to reflect 0% use of treatment costs			
Company updated model	[redacted]	[redacted]	£26,956
EAG model	[redacted]	[redacted]	£36,934

Key issue 7: Results presented separately for patients in KALM-1 only and KALM-2 only (update of scenario analysis 3.a and 3.b in the company submission)

Model	Incremental costs	Incremental QALYs	ICER
KALM-1 only			
Company updated model	[redacted]	[redacted]	£30,389
EAG model	[redacted]	[redacted]	£34,562
KALM-2 only			
Company updated model	[redacted]	[redacted]	£23,115
EAG model	[redacted]	[redacted]	£35,355

Key issue 9: Results presented for data without MI to account for missing data (new scenario)

Model	Incremental costs	Incremental QALYs	ICER
Without MI to account for missing data			
Company updated model	[redacted]	[redacted]	£24,516
EAG model	[redacted]	[redacted]	£37,830

Key issue 10: Results presented for an additional waning effect (new scenario)

Following a review of the 3 atopic dermatitis TAs, the company have conducted additional analysis on the long-term benefit of established clinical management to reflect the scenarios that were accepted in TA534, TA681 and TA814.

In TA534, the company presented analysis in which 100%, 96% and 97% of patients had returned to baseline utility by year 5. Within the company model for TA534, treatment related utilities were modelled to reflect the clinical benefit associated with being treated with the intervention and best supportive care.

To reflect a similar level of treatment waning in this analysis, the company have presented results whereby a waning effect is applied separately to the none, mild, moderate, and severe CKD-aP health states so that the proportion of people in each health state at year 5 (sensitivity analysis 1) and at year 10 (sensitivity analysis 2) equals the proportion of people in each health state at baseline. This approach assumes that the rate of waning may vary depending on CKD-aP severity.

Method

- To ensure that the proportions match, mortality effects must first be excluded then waning effect adjusted before re-adjusting the mortality effects to understand the impact on the ICER.
- The waning effect proportions were then adjusted starting with the none and mild state to match the baseline proportions. For simplicity this was rounded to 2 decimal places. All relevant analyses can be seen from cell V64 and below in the ECMEngine sheet.

- Goal seek was then used to adjust the waning effect for the moderate and severe proportions of patients.
- This was repeated for the 5- and 10-year estimates in both the EAG and company updated base case model.

The values of for the waning effect and the relative patient numbers excluding mortality are presented in the tables below. For reference, the company have also provided a table showing the proportion of patients in state at years 5 and 10 in the company's updated model with the 10% waning applied across all 4 states.

Year	None	Mild	Moderate	Severe	Very severe
EAG model					
Basecase proportion in state	0.00%	0.00%	55.28%	34.17%	10.55%
Year 5 - waning					
Annual waning applied to Cycle 5	*92.00%*	*100.00%*	*10.89%*	*8.27%*	*n/a*
Proportion of patients in state	0.00%	0.00%	55.35%	34.18%	10.47%
Year 10 - waning					
Annual waning applied to Cycle 5	*60.00%*	*91.00%*	*4.62%*	*3.37%*	*n/a*
Proportion of patients in state	0.00%	0.00%	55.27%	34.24%	10.49%

Year	None	Mild	Moderate	Severe	Very severe
Company updated model					
Basecase proportion in state	0.00%	0.00%	55.28%	34.17%	10.55%
Year 5 - waning					
Annual waning applied to Cycle 5	*95.00%*	*95.00%*	*9.30%*	*6.35%*	*n/a*
Proportion of patients in state	*0.00%*	*0.00%*	*55.32%*	*34.17%*	*10.50%*

Year 10 - waning					
Annual waning applied to Cycle 5	*60.00%*	*85.50%*	*3.92%*	*2.72%*	*n/a*
Proportion of patients in state	*0.00%*	*0.00%*	*55.20%*	*34.11%*	*10.68%*

Year	None	Mild	Moderate	Severe	Very severe
Company updated model					
Basecase proportion in state	0.00%	0.00%	55.28%	34.17%	10.55%
Annual waning applied to Cycle 5	*10.00%*	*10.00%*	*10.00%*	*10.00%*	*n/a*
Proportion of patients in state year 5	*5.25%*	*21.48%*	*34.85%*	*25.38%*	*13.04%*
Proportion of patients in state year 10	*3.10%*	*14.41%*	*28.01%*	*28.02%*	*26.46%*

There are a range of values that could be used to match the proportion of patients by the chosen year in the none and mild cohort as these had 0 patients at baseline. Thus, the lowest estimate of waning was selected to delay the transitions over the years.

The company recognise that this is a pragmatic and simplified approach, but nonetheless, can be used to inform decision making as to the direction of the ICER if a waning effect similar to those proposed in TA534, TA681, and TA814 is considered.

The impact on the ICER results is presented below. The company note that as the waning effect is continuous (given the model structure), the 10-year results are likely more informative as at this point only 9% of the population are alive, therefore any sustained impact of waning beyond these years is likely to be minimal.

Model	Incremental costs	Incremental QALYs	ICER
Waning effect applied to match baseline at year 5 A waning effect is applied to match the proportion of patients in the ECM arm at year 5 to the proportion of patients at baseline.			
Company updated model	[redacted]	[redacted]	£18,613
EAG model	[redacted]	[redacted]	£19,248
Waning effect applied to match baseline at year 10 A waning effect is applied to match the proportion of patients in the ECM arm at year 10 to the proportion of patients at baseline.			
Company updated model	[redacted]	[redacted]	£20,668
EAG model	[redacted]	[redacted]	£21,625

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis ID3890

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with chronic kidney disease-associated pruritis or caring for a patient with chronic kidney disease-associated pruritis. The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary (section 1.3) at the beginning of the EAR.

A patient perspective could help either:

- resolve any uncertainty that has been identified OR
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

Patient expert statement

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

Patient expert statement

The deadline for your response is **5pm** on **<insert deadline>.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with chronic kidney disease-associated pruritis

Table 1 About you, chronic kidney disease-associated pruritis, current treatments and equality

1. Your name	Faizan Awan
2. Are you (please tick all that apply)	☒ A patient with chronic kidney disease-associated pruritis? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with chronic kidney disease-associated pruritis? ☐ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	Kidney Research UK
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☒ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☐ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement ☐ I agree with it and **will be** completing
5. How did you gather the information included in your statement? (please tick all that apply)	☒ I am drawing from personal experience ☐ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: ☐ I have completed part 2 of the statement **after attending** the expert

Patient expert statement

	engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☒ I have not completed part 2 of the statement
6. What is your experience of living with chronic kidney disease-associated pruritis? **If you are a carer (for someone with chronic kidney disease-associated pruritis please share your experience of caring for them**	For me itch was a real pain, it felt as if my body was on fire and endless nights and some days of rigorous itching, using anything I could find or think of to relieve the itch, cold showers, using the carpet to get between my toes, moisturisers and even got to the point I tried utensils from the kitchen. My sleep was disturbed, my family members would wake during the night as I was doing all I could to get rid of the sensation of itch, sleep was inconsistent and when I did finally get to sleep, I would miss the days as I had not slept during the night. Blood stains would require family members and myself to constantly wash bedding and my clothes.
7a. What do you think of the current treatments and care available for chronic kidney disease-associated pruritis on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	**For me they didn't work, so were inadequate I was given moisturisers mainly. And extra phosphate binders as was told it was due to chemical imbalance in my bloods.**
8. If there are disadvantages for patients of current NHS treatments for chronic kidney disease-associated pruritis (for example, how they are given or taken, side effects of treatment, and any others) please describe these	The main disadvantage is that it did not relieve the burden of itch. I wasn't sleeping well days drifted away and my general quality of life suffered as I couldn't do what I wanted to in the day as I had not slept in the night. My concentration in days I could go and volunteer was not the best, the effect on family members who woke in the night too as I was having cold showers or using the carpet affect their lives also.
9a. If there are advantages of chronic kidney disease-associated pruritis over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others?	

9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why? **9c. Does difelikefalin help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	
10. If there are disadvantages of difelikefalin over current treatments on the NHS please describe these. For example, are there any risks with difelikefalin If you are concerned about any potential side effects you have heard about, please describe them and explain why	I don't know as I have never had Difeikefalin. It would be interesting to learn more about the side effects, but in lay terms plain English would be great.
11. Are there any groups of patients who might benefit more from difelikefalin or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	I think potentially anyone who has itch could potentially benefit from this. Those who may have difficulties or other illness or issues should where possible have someone with them when a medication is introduced so the burden of understanding isn't all put on the patient, if there are issues with other conditions then maybe the two teams involved in the patient's care should communicate and get a fuller broader understanding of what the patient is going through before taking any action.
12. Are there any potential equality issues that should be taken into account when considering chronic kidney disease-associated pruritis? and difelikefalin Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics	If something is going to affect a pregnancy then simply don't do it. Some ethnicities can present skin conditions differently, so although it may not be as visible on some skin then others i would I think it would be all about listening to the experience of what the patient is going through. Religiously are there any ingredients that would be forbidden to certain people, is this explored?

More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	
13. Are there any other issues that you would like the committee to consider?	no

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement

Population in decision problem is narrowed relative to the NICE scope The decision problem is restricted to in-centre haemodialysis. The EAG consider this evidence may not apply to the whole population having haemodialysis for chronic kidney disease	

Patient expert statement

Comparison in evidence base is different to comparison in the NICE scope We consider patient perspectives may particularly help to address this issue The trials compared difelikefalin plus established clinical management with placebo plus established clinical management. The EAG consider this may lead to more optimistic benefits, Which would be the most appropriate comparison?	
Inadequate systematic literature review was carried out for clinical effectiveness The company did not carry out a systematic literature review for	

Patient expert statement

clinical effectiveness. The EAG consider this makes it difficult to identify if all appropriate studies have been included.	
Applicability of trial findings unclear because of lack of anti-itch medication sub-grouping We consider patient perspectives may particularly help to address this issue. The trial population included people who were using anti-itch medication, and those who were not. The EAG consider a sub-group analysis of use of anti-itch medication would help identify any differences between established clinical management in the trials and the target UK population.	

Patient expert statement

In your experience would difelikefalin be likely to be used alongside other anti-itch medications? If so is this likely to impact effectiveness?	
Applicability of findings may be reduced by differences in ethnicity between the trials and the UK target population In the trials there was a larger proportion of people with a black family background and subgrouping suggest they may have better outcomes than the UK target population. The EAG consider this may affect generalisability of the trials.	
Rationale for statistical analysis unclear The EAG consider the company had not provided justification	

Patient expert statement

for the use of its multiple imputation analysis and logistic regression.	
Methods used to pool trials were not appropriate Clinical effectiveness data from the KALM trials was pooled without adjusting for differences between the trials. The EAG consider this may bias estimates of treatment effectiveness.	
Approach estimating transition probabilities We consider patient perspectives may particularly help to address this issue. Based on data from the KALM trials, the company assumed people could move to a better or worse health state (up to three health states at a time)	

Patient expert statement

whereas the EAG preferred to use directly estimated transition probabilities which showed people could improve up to a maximum of two health states but will never switch to a worse health state in terms of itching. In your view how likely are people to transition to mild, moderate severe and very severe health states while using difelikefalin?	
Lack of clarity how multiple imputation was used The EAG would like more details of how imputed data sets were combined to find the estimated transition probabilities	
Insufficient evidence regarding transitions after 64 weeks The company assumed a 5% deterioration per	

Patient expert statement

year for the established clinical management but there would be no waning of treatment effect for difelikefalin. The EAG consider this is uncertain but assumed that people with and without difelikefalin, would remain in their current health state at 64 weeks.	
Insufficient evidence that decreasing level pruritus improved mortality The company used an increased risk of death for people in the moderate, severe and very severe health states of the chronic kidney disease-associated pruritis population. The EAG did not consider the evidence was substantial enough to establish a causal relationship between pruritus and mortality so	

Patient expert statement

it removed this elevated risk of death from the model.	
Are there any important issues that have been missed in the EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is the end of on **16 December 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Kidney Research UK
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
1. The population in the decision problem and the included trials appears narrower than that in the National Institute for Health and Care Excellence (NICE) final scope. The decision problem and trial populations preclude first line treatment and are restricted to people receiving in-centre haemodialysis (ICHD). The NICE scope makes no restrictions in terms of ICHD and treatment line.	**Yes**	Kidney Research UK understands that KALM trials did not directly include any comparator treatments, but that patients using anti-itch medication at baseline were allowed to continue doing so. It is positive that the use of anti-itch medication will not preclude use of difelikefalin, but we would have significant concerns about the possibility of a restricted treatment line. Kidney patients prefer not to be over-medicated, and other anti-itch medication that could be used may not be licensed for treatment of pruritis. [redacted] We already know that patients report being given different advice to cope with the itching – from using certain emollients to changing their diet. Furthermore, as discussed in our previous submission - the DOPPS study has highlighted that 18% of haemodialysis patients are very much or extremely troubled by itching, but up-to 18% receive no treatment for this symptom. In

		addition, 17% had not reported itching to a healthcare professional (Rayner et al., 2017). Therefore, to have off licensed treatments ahead of the licensed treatment in pruritis in a treatment line would not be preferable to patients, nor sensible.
2. The comparison in the included trials is difelikefalin plus established clinical management (ECM) versus placebo plus ECM, whereas the comparison in the NICE final scope and the decision problem is difelikefalin versus ECM. The nature of the treatment comparison in the trials may lead to a more optimistic impression of the study drug's benefits compared with the NICE final scope/decision problem.	**No**	Please provide your response to this key issue, including any new evidence, data or analyses
3. A systematic literature review (SLR) of clinical effectiveness evidence	**No**	Please provide your response to this key issue, including any new evidence, data or analyses

was not carried out. This made it difficult to determine whether all relevant studies were included in the clinical effectiveness part of the submission. The Evidence Assessment Group (EAG) identified two potentially relevant randomised controlled trials (RCTs) that had not been considered within the company submission (CS). Data from these RCTs have been added to the report by the EAG.		
4. Differences between ECM in the included trials and the United Kingdom (UK) target population may limit the generalisability of clinical effectiveness evidence from the trials. The company did not provide the results of sub-group analyses in relation to specific anti-itch	No	Please provide your response to this key issue, including any new evidence, data or analyses

medications other than difelikefalin. This hindered the evaluation of the impact of differences in the use of non-difelikefalin anti-itch medications between the included trials and the UK target population.		
5. The included trials recruited a larger proportion of Black participants relative to those seen in the UK target population. Results from sub-group analyses suggested that Black participants tend to have better difelikefalin outcomes than other ethnic groups. This may further affect the generalisability of the overall trial results.	**Yes**	We are concerned that the key issue relating to health inequalities within this appraisal has been misinterpreted. Kidney disease disproportionally affects people from deprived communities and ethnic minority groups and people in these cohorts' progress faster to end stage renal failure*. Evidence shows that fewer kidney patients from deprived communities are treated with peritoneal dialysis, with more treated with haemodialysis. There are therefore likely to be proportionally more people from these cohorts on haemodialysis, experiencing pruritus and likely to benefit from this treatment. This is not equal to the treatment being more beneficial to those groups exclusively, which in some parts of the technical engagement appears to have been considered as the key equation pertaining to health inequalities. * Kidney Health Inequalities in the UK: Reflecting on the past, reducing in the future. Kidney Research UK 2018
6. The rationale for the statistical analysis in the included trials	No	

(specifically, multiple imputation (MI) and logistic regression) is not transparent. Lack of information about model inputs and outputs in both instances has hindered the EAG's assessment of the quality of the statistical analyses.		
7. Clinical effectiveness data from the KALM-1 and KALM-2 trials were pooled without adjusting for differences between the trials. This may have resulted in biased estimates of treatment effectiveness.	No	
8. The company assumed in the base case and an alternative scenario that transitions can be modelled by only looking	No	

at the probability of shifting between one, two or three health states up or down, regardless of the current health state. This assumption did not seem to be supported by the directly estimated transition probabilities.		
9. In the estimation of the transition probabilities, use has been made of data that was imputed multiple times to account for missing data. It is unclear to the EAG how all transition matrices were derived in light of the multiple imputation.	No	
10. In the base case, treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year	No	Please provide your response to this key issue, including any new evidence, data or analyses

while no waning impact was assumed for the difelikefalin arm. In the absence of further real-world evidence to support the waning impact of ECM and/or the lack of waning over time with difelikefalin, the EAG considers this assumption uncertain.		
11. The company applies an increased risk of death for patients in the moderate, severe and very severe health states of the chronic kidney disease-associated pruritis (CKD-aP) population, based on an observational study. The EAG considers the evidence presented not substantial enough to establish a causal	No	Please provide your response to this key issue, including any new evidence, data or analyses

relationship between pruritus and mortality of these patients.		

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is the end of on **16 December 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Renal Pharmacy Group
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
1. The population in the decision problem and the included trials appears narrower than that in the National Institute for Health and Care Excellence (NICE) final scope. The decision problem and trial populations preclude first line treatment and are restricted to people receiving in-centre haemodialysis (ICHD). The NICE scope makes no restrictions in terms of ICHD and treatment line.	No	Please provide your response to this key issue, including any new evidence, data or analyses The decision problem should restrict to use for in-centre haemodialysis, including satellite dialysis units, due to current licensing and potential for adverse effects such as dizziness and somnolence. As we develop more experience with its use, it may be that use is extended to home haemodialysis patients if they have therapy initiated in hospital or are already established on therapy. Although numbers of patients on home haemodialysis would likely be small. In clinical practice difelikefalin would be expected to be used in addition to established clinical management (ECM), as defined in the decision problem and trial population. This is proposed in figure 3 of the company's evidence.

2. The comparison in the included trials is difelikefalin plus established clinical management (ECM) versus placebo plus ECM, whereas the comparison in the NICE final scope and the decision problem is difelikefalin versus ECM. The nature of the treatment comparison in the trials may lead to a more optimistic impression of the study drug's benefits compared with the NICE final scope/decision problem.	No	Please provide your response to this key issue, including any new evidence, data or analyses In clinical practice difelikefalin would be used in addition to ECM, for the defined population of moderate to severe pruritis. (as per Figure 3 of company's evidence). Patients with moderate to severe pruritis, as per trial population, are likely to have have tried or taken other treatments. It should be considered that in cases where pruritis treatments have been ineffective or led to adverse effects, they may have been discontinued prior to trial enrolment so are not captured. From the pooled analysis of Kalm-1 and 2 there was up to 20.4% participants taking gabapentin or pregabalin for non-itch related indications and there is likely to be benefits of these treatments on pruritis in both groups. If difelikefalin did provide enhanced efficacy when used in addition to ECM this would be beneficial to those with moderate-severe pruritis. There may be people where difelikefailin would be a more appropriate option than some of the ECM treatments, for example in elderly frail patients on dialysis gabapentin and pregabalin may be best avoided due to their adverse effects. Conversely it may be that people who have good responses to difelikefalin will allow reduction of the need for some of their pre-existing clinical management options that were less successful in managing their pruritis. Pill burden is a big problem in patients on dialysis.
3. A systematic literature review (SLR) of clinical effectiveness evidence was not carried out. This made it difficult to	No	Please provide your response to this key issue, including any new evidence, data or analyses

determine whether all relevant studies were included in the clinical effectiveness part of the submission. The Evidence Assessment Group (EAG) identified two potentially relevant randomised controlled trials (RCTs) that had not been considered within the company submission (CS). Data from these RCTs have been added to the report by the EAG.		
4. Differences between ECM in the included trials and the United Kingdom (UK) target population may limit the generalisability of clinical effectiveness evidence from the trials. The company did not provide the results of sub-group analyses in relation to specific anti-itch medications other than difelikefalin. This	No	Please provide your response to this key issue, including any new evidence, data or analyses There is variation in ECM across the UK which is currently being investigated by the SIESMIC survey. These variations relate to the paucity of data for specific therapeutic options and without knowledge of current UK practice it does make it difficult to undertake comparisons of UK practice to ECM in the trials. Topical therapies, antihistamines as a class and pregabalin/gabapentin are likely to be the most common practice alongside practical measures of addressing dialysis (figure 3). These therapeutic options are those that were used in the trials. We have to be mindful that effectiveness of difelikefalin is to be reviewed at 12 weeks, so if not effective therapy will be ceased.

hindered the evaluation of the impact of differences in the use of non-difelikefalin anti-itch medications between the included trials and the UK target population.		The UK Kidney Association supportive care specialist interest group are working on developing a UKKA guideline to include the management of pruritis with a stepwise approach to treatment selection. It is likely to be similar to Figure 3.
5. The included trials recruited a larger proportion of Black participants relative to those seen in the UK target population. Results from sub-group analyses suggested that Black participants tend to have better difelikefalin outcomes than other ethnic groups. This may further affect the generalisability of the overall trial results.	No	Please provide your response to this key issue, including any new evidence, data or analyses Benefit was seen across all ethnicities so this shouldn't limit the option of difelikefalin across ethnic groups.
6. The rationale for the statistical analysis in the included trials (specifically, multiple imputation (MI) and logistic regression) is not		

transparent. Lack of information about model inputs and outputs in both instances has hindered the EAG's assessment of the quality of the statistical analyses.		
7. Clinical effectiveness data from the KALM-1 and KALM-2 trials were pooled without adjusting for differences between the trials. This may have resulted in biased estimates of treatment effectiveness.		
8. The company assumed in the base case and an alternative scenario that transitions can be modelled by only looking at the probability of shifting between one, two or three health states up		

or down, regardless of the current health state. This assumption did not seem to be supported by the directly estimated transition probabilities.		
9. In the estimation of the transition probabilities, use has been made of data that was imputed multiple times to account for missing data. It is unclear to the EAG how all transition matrices were derived in light of the multiple imputation.		
10. In the base case, treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In the	No	Please provide your response to this key issue, including any new evidence, data or analyses

absence of further real-world evidence to support the waning impact of ECM and/or the lack of waning over time with difelikefalin, the EAG considers this assumption uncertain.		
11. The company applies an increased risk of death for patients in the moderate, severe and very severe health states of the chronic kidney disease-associated pruritis (CKD-aP) population, based on an observational study. The EAG considers the evidence presented not substantial enough to establish a causal relationship between pruritus and mortality of these patients.	No	Please provide your response to this key issue, including any new evidence, data or analyses

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Difelikefalin for treating pruritus in people having haemodialysis [ID3890]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is the end of on **16 December 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	NHS England (Specialised Commissioning)
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
1. The population in the decision problem and the included trials appears narrower than that in the National Institute for Health and Care Excellence (NICE) final scope. The decision problem and trial populations preclude first line treatment and are restricted to people receiving in-centre haemodialysis (ICHD). The NICE scope makes no restrictions in terms of ICHD and treatment line.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

2. The comparison in the included trials is difelikefalin plus established clinical management (ECM) versus placebo plus ECM, whereas the comparison in the NICE final scope and the decision problem is difelikefalin versus ECM. The nature of the treatment comparison in the trials may lead to a more optimistic impression of the study drug's benefits compared with the NICE final scope/decision problem.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
3. A systematic literature review (SLR) of clinical effectiveness evidence was not carried out. This made it difficult to determine whether all relevant studies were included in the clinical effectiveness part of the submission. The	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

Evidence Assessment Group (EAG) identified two potentially relevant randomised controlled trials (RCTs) that had not been considered within the company submission (CS). Data from these RCTs have been added to the report by the EAG.		
4. Differences between ECM in the included trials and the United Kingdom (UK) target population may limit the generalisability of clinical effectiveness evidence from the trials. The company did not provide the results of sub-group analyses in relation to specific anti-itch medications other than difelikefalin. This hindered the evaluation of the impact of differences in the use of non-difelikefalin anti-itch medications between the	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

included trials and the UK target population.		
5. The included trials recruited a larger proportion of Black participants relative to those seen in the UK target population. Results from sub-group analyses suggested that Black participants tend to have better difelikefalin outcomes than other ethnic groups. This may further affect the generalisability of the overall trial results.	Yes/No	A therapy that may preferentially benefit ethnic minority groups should not be discounted on the basis of a population-wide evaluation. This will need to be considered by NICE.
6. The rationale for the statistical analysis in the included trials (specifically, multiple imputation (MI) and logistic regression) is not transparent. Lack of information about model inputs and outputs in both instances has		

hindered the EAG's assessment of the quality of the statistical analyses.		
7. Clinical effectiveness data from the KALM-1 and KALM-2 trials were pooled without adjusting for differences between the trials. This may have resulted in biased estimates of treatment effectiveness.		
8. The company assumed in the base case and an alternative scenario that transitions can be modelled by only looking at the probability of shifting between one, two or three health states up or down, regardless of the current health state. This assumption did not seem to be supported by		

the directly estimated transition probabilities.		
9. In the estimation of the transition probabilities, use has been made of data that was imputed multiple times to account for missing data. It is unclear to the EAG how all transition matrices were derived in light of the multiple imputation.		
10. In the base case, treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In the absence of further real-world evidence to support the waning impact of ECM and/or	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

the lack of waning over time with difelikefalin, the EAG considers this assumption uncertain.		
11. The company applies an increased risk of death for patients in the moderate, severe and very severe health states of the chronic kidney disease-associated pruritis (CKD-aP) population, based on an observational study. The EAG considers the evidence presented not substantial enough to establish a causal relationship between pruritus and mortality of these patients.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]

Produced by Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC+)

Authors

Susan O'Meara, Reviews Manager, KSR Ltd, United Kingdom (UK)
Venetia Qendri, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands (NL)
Mark Perry, Reviews Manager, KSR Ltd, UK
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Ingelin Kvamme, Health Economics Researcher, Institute for Medical Technology Assessment (IMTA), EUR, the NL
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Caro Noake, Information Specialist, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Maiwenn Al, Health Economics Researcher, ESHPM, EUR, the NL
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to Susan O'Meara
Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 12/01/2023

Declared competing interests of the authors None

Acknowledgements None

[redacted]

Abbreviations

β	Regression coefficient
AiC	Academic in confidence
CI	Confidence interval
CiC	Commercial in confidence
CKD	Chronic kidney disease
CKD-aP	CKD-associated pruritis
CS	Company submission
DFK	Difelikefalin
EAG	Evidence Assessment Group
EAR	Evidence Assessment Report
EBMR	Evidence-Based Medicine Reviews
ECM	Established clinical management
ESHPM	Erasmus School of Health Policy & Management
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
ICER	Incremental cost-effectiveness ratio
iMTA	Institute for Medical Technology Assessment
Incr.	Incremental
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
KALM	KALM-1 and KALM-2 were randomised trials to study the safety and efficacy of difelikefalin in haemodialysis patients with moderate-to-severe pruritus
KSR	Kleijnen Systematic Reviews Limited
MI	Multiple imputation
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NL	The Netherlands
No.	Number
NRS	Numerical rating scale
OR	Odds ratio
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life year
ROBIS	Risk of Bias in Systematic Reviews
ROW	Rest of world
SAP	Statistical analysis plan
SLR	Systematic literature review
STA	Single Technology Appraisal
TA	Technology Appraisal
TW	Technical Engagement
UK	United Kingdom
UMC+	University Medical Centre
USA	United States of America
WI-NRS	Worst Itching Intensity Numerical Rating Scale

Introduction

This document is the Evidence Assessment Group's (EAG's) response to comments and additional data provided by the company as part of the technical engagement (TE) process for difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis.[1]

Key Issues

Key issue 1:

As part of their TE response, the company provided some statements but no additional evidence, data or analyses.[1]

The EAG wishes to thank the company for the clarification provided. In the light of the additional information received, chiefly that difelikefalin would be given as an adjunct to pre-existing established clinical management (ECM),[1] the EAG accepts that the departures from the NICE scope in terms of population are appropriate.

Key issue 2:

The company provided statements and incorporated subgroup analyses into the updated economic model as part of their TE response.[1]

The EAG thanks the company for the clarification around how difelikefalin would be given in practice – that it would be given as an adjunct to any pre-existing ECM, rather than as a substitute for it.[1] Therefore, the EAG accepts that it does make clinical sense for the comparison to be Difelikefalin + ECM versus placebo + ECM rather than just difelikefalin versus ECM. The EAG still considers that the comparison adopted by the company could lead to an effect of greater magnitude than the latter comparison, requested by NICE. However, it could be argued that this does not matter if the NICE comparison is not relevant to clinical practice: if difelikefalin would only ever be given with existing ECM then it appears unnecessary to compare difelikefalin alone to ECM.

The company's mention of the sub-groups in relation to this topic[1] appears a little irrelevant, because the sub-grouped data do not shed any light on answering the NICE comparison (difelikefalin versus ECM).

Key issue 3:

The company conducted a clinical effectiveness systematic literature review (SLR) as part of their TE response and presented a report of this separately to the TE form.[2]

The EAG considered the SLR question to be appropriate: "*What is the efficacy and safety evidence for intravenous difelikefalin for the treatment of CKD-aP in adults receiving haemodialysis?*". The protocol (shown in the SLR report as the inclusion criteria) specifies the population as adults with chronic kidney disease-associated pruritis (CKD-aP).[2] This is slightly wider than the NICE-specified population which specified that the pruritis must be moderate or severe and that patients must be receiving haemodialysis.[3] Haemodialysis is mentioned in the title of the SLR, but because it is absent in the population definition it is unclear whether trials could be included in the SLR if they did not include participants receiving haemodialysis and/or having moderate/severe disease. The intervention is clearly described. Comparators are not described, which implies that a comparator could comprise any intervention or no intervention. The outcomes are extremely vague, described simply as "*efficacy*

and safety outcomes", implying that any clinical outcome would be eligible for this SLR. The study selection criteria also specified no restrictions on study design, timeframe or language. In summary, it is apparent that the study selection criteria of the SLR[2] are wider than the NICE final scope,[3] particularly in relation to population, comparator and outcome characteristics. Whilst this may have resulted in identification of reports that are less relevant to the NICE final scope, it may also mean that fewer relevant papers were missed.

The EAG noted that whilst the SLR's study selection criteria were very wide the search strategies were more restrictive (e.g., no terms for comparators were included). The Embase, Medline and Evidence-Based Medicine Reviews (EBMR) strategies were structured in 3 facets: CKD/Haemodialysis + Pruritis + Difelikefalin. Given the low number of papers recalled (n=147 across the three databases) and also that a search for the Difelikefalin facet alone retrieved only 252 records, the latter may have been a more preferable, cautious approach to retrieving difelikefalin studies. However, the EAG considers that given the additional searches (including searching of ClinicalTrials.gov and reference list checking) the use of the combined facets is unlikely to have affected the overall recall of results. The EAG noted a limited use of truncation within the strategies, which may have been mitigated by the inclusion of subject headings. There was also a line combination error in the search of EMBR where line #7 containing terms for haemodialysis, appeared to have been omitted from the combination of terms relating to kidney disease in line #8. As with the previous limitations the inclusion of additional searches may have protected against any loss of recall resulting from these errors. Furthermore, the searches appeared to correctly correlate to the retrieved trials.[2] There is potentially a problem with only searching for studies of difelikefalin if an indirect treatment comparison (ITC) might be needed whereby studies of treatments not compared to difelikefalin would be required. If the comparator in the KALM trials, which is ECM, which is mixture of various treatments and no treatment, is identical to UK clinical practice then no ITC would be required but as described under Key Issue 4, there might be a lack of applicability of ECM to UK clinical practice. However, given that ECM is a mixture of interventions it would probably be unfeasible to use it as a common comparator for an ITC with any single one of the mixture of treatments. Therefore, including the difelikefalin facet is unlikely to be a problem and the approach of subgroup analysis of the KALM trials recommended by the EAG in Key Issue 4 is more appropriate.

The study flow was appropriately summarised using PRISMA flow diagrams with accompanying narrative.[2] The EAG considered that the study selection and data extraction processes were rigorous, with two researchers working independently, and a third researcher available to facilitate consensus. The reviewers performed a thorough evaluation of the risk of bias of the seven included studies.[2]

Of note, the SLR document does not report a pre-specified plan for data synthesis. For example, the intended approach for structuring the narrative synthesis or undertaking meta-analysis including methods for investigating statistical heterogeneity (e.g., subgroup analysis) or sensitivity analysis is not described. Furthermore, no completed analyses are presented beyond a tabulated presentation of outcomes from the seven included trials.[2]

It is unclear why a data synthesis was not planned or carried out. The EAG acknowledges the possibility that details of data synthesis were omitted because the results did not support the company's position. The inclusion of a pre-specified analysis plan would have potentially reduced the risk of bias. In light of the lack of any synthesis and conclusion about the effects of difelikefalin across the seven trials it is very difficult to see how the company have reached their conclusion that the results have no impact on cost-effectiveness estimates.

The EAG assessed the risk of bias of the company's SLR using the "*Risk Of Bias In Systematic reviews*" (ROBIS) tool.[4] This returned an overall rating of high risk of bias. For domain 1 (study eligibility criteria) a rating of high risk of bias was given because of the lack of information on the review protocol, including how data would be synthesised. Domain 2 (identification and selection of studies) and domain 3 (data collection and study appraisal) were both given low risk of bias. Domain 4 (synthesis of findings) was given high risk of bias as no meaningful synthesis had been carried out.

Key issue 4:

As part of their TE response, the company provided some statements but no additional evidence, data or analyses.[1]

The company's statement that "*…there is no clinical basis to suggest that 1) use of anti-itch medication is consistent across UK clinical practice, and 2) one medication out of a sub-group of treatments (e.g., antihistamines) is preferred over another*" is not backed up by supporting evidence.[1] The classes of anti-itch medication that are most frequently used in the UK might be an issue that could be answered by the clinical members of the Appraisal Committee.

The EAG accepts that the sizes of sub-groups would be small in any evaluation of the different classes of anti-itch medication, but this cannot be used as a valid reason to not undertake the sub-group analysis. It is only once the sub-group analyses are performed, and the relative effect sizes in the different strata are observed, that any conclusions can be made about whether the class of anti-itch medication is a potential source of inapplicability of trial results to the UK target population. It is acknowledged that it may not be possible to observe statistically significant differences in efficacy between classes, but large differences between classes may still be informative, indicating potential effects that should be considered. On the other hand, small differences may allow the matter to be put to rest. However, until the company carries out the requested sub-group analysis the questions about applicability will remain.

Key issue 5:

Again, the company provided some statements but no additional evidence, data or analyses in relation to this Key Issue.[1]

The EAG has noted two important features of the objective data, which have been previously stated, but which will be reiterated here:

1. The trials had more black participants than the UK target population (29.2% compared to 12.8%; these data from Table 3.8 in the EAG report[5]were provided by the company's response to clarification).[6]

2. Sub-group analyses suggest a trend for a better relative response to difelikefalin in black participants, in terms of the number with a >3 point improvement in the Worst Itching Intensity Numerical Rating Scale (WI-NRS) (Table 3.9, EAG report).[5] Since the analysis was not powered for sub-group analyses, the fact that this is probably not a significant difference does not mean that no difference exists in the population – it means that one may exist but was probably not detected. Whilst the EAG accept the uncertainty in the sub-group estimates in Table 3.9 of the EAG report[5] and realise that it cannot be definitively concluded that race is an effect modifier, the EAG believe there is enough evidence to suggest consideration of this point by the Appraisal Committee. The company's statement that the confidence intervals overlapping indicates definitive evidence of a lack of significant difference between the effects in the different race categories[1] is incorrect. It is quite possible for two estimates to be significantly different even though their individual 95% confidence intervals (CIs) overlap; the

correct test is whether the confidence interval around the relative effect crosses the null line.[7] Thus, observation of the overlap of the confidence intervals in Figure 12[1] is not helpful. In addition, the EAG would like to take issue with a statement made by the company that "*The nonsignificant difference on odds ratios was due to a lower placebo response in black participants*".[1] This statement misses the point of the placebo group. In a properly conducted RCT, the intervention and control groups should be so similar that the effects of any intervening variables (such as the placebo effect) should be extremely similar across groups. This allows the researcher to assume that the level of these intervening variables is equivalent across groups. This then means that taking the difference in effect across groups will allow the effect of the intervening variables to cancel out, leaving only the treatment effect. Following this argument, any level of 'placebo' response by the black participants in the placebo group would have been accompanied by a very similar 'placebo' response by the black participants in the intervention group. Therefore, this effect would have cancelled out.

Given these two features, it can be concluded that there is a realistic possibility that the effect sizes from the trial may not be applicable to the UK target population.

Key issue 6:

As part of their TE response, the company provided some statements but no additional evidence, data or analyses.[1]

In attempting to clarify the choice of statistical methods used in the trials and the company submission (CS), the EAG has posed specific questions to the company throughout the Single Technology Appraisal (STA) process. The use of multiple imputation (MI) was built into the statistical analysis plan (SAP) for handling missing numerical rating scale (NRS) data as we note in the EAG report.[5] As part of their response to the clarification questions, the company maintained that:

"*The choice of multiple imputation for the treatment of missing data was suggested by the Food and Drug Administration (FDA) during a meeting held on September 6th 2017 to discuss the Phase 3 clinical development program for IV difelikefalin. Specifically, the FDA stated that "The efficacy analyses should be based on the intent-to-treat (ITT) population (i.e., all randomized subjects)" and added that the protocol "should pre-specify a scientifically sound primary imputation method (e.g. multiple imputation) to handle missing data. Multiple imputation is also one of the analytical methods recommended by the National Research Council Committee on National Statistics in their 2010 report on the prevention and treatment of missing data in clinical trials. Based on these regulatory and technical recommendations, Cara decided to use multiple imputation as the primary method for the treatment of missing data in the pivotal studies KALM-1 and KALM-2*." [6]

In their response to TE, [1] the company states for the first time that the MI method was chosen based on previous experience derived from data generated in the Phase 2 studies CLIN2005 and CLIN2101. These were Phase 2, randomized, double-blind, placebo-controlled studies. In CLIN2005 patients were administered either IV CR845 1.0 mcg/kg (n=33) or placebo (n=32) 3 times per week for 2 weeks, after each haemodialysis session; while in CLIN2101, three different doses were tested: 0.5 mcg/kg (n=44), 1 mcg/kg (n=41) and 1.5 mcg/kg (n=44), or placebo (n=45) over an 8-week treatment period in haemodialysis patients.[8] The relevance of the Phase 2 studies to the altered dose and regime of the Phase 3 studies used for the CS is not immediately obvious, especially regarding missing data. Detailed information is not available at this stage to the EAG to provide an informed critique.

Nevertheless, the main issues around the statistical methods, for both MI and logistic regression, is the lack of a justification on why specific covariates were used in the models and others were not considered

or excluded e.g., race, as detailed in the EAG report. [5] The company has provided no further evidence in relation to this. In addition, the company was asked to formally report the design and results of the logistic regression as key aspects were missing to which they did not respond.[6]

Finally, the data presented in "*Key issue 6 additional table of results*"[1] have previously been provided in the response to clarification letter.[6]

Key issue 7:

As part of their TE response, the company provided statements plus an updated scenario analysis.[1]

The company stated that there was a typo in their response to clarification questions, now reporting that the pooled analysis of clinical effectiveness data from the KALM-1 and KALM-2 trials were additionally adjusted for a combined region/study covariate. The EAG notes that the same typo is in the published study presenting these results in Topf et al (2022). [9] The company specified that:

"*As all sites in CLIN3102 were in the United States, a single model parameter will be used to account for the effects of region and study that will have 3 levels: CLIN3102 United States, CLIN3103 United States, and CLIN3103 ROW (Western Europe/ Western European Origin [Canada, United Kingdom, Germany, Australia, New Zealand], Eastern Europe [Poland, Hungary, Romania, Czech Republic], Asia [Taiwan, South Korea])*" (page 12 of the TE response).[1]

There is a series of concerns that the EAG would like to raise. Regarding this newly reported model parameter:

- KALM-1 was carried out in one region (USA), while KALM-2 in four (USA, Eastern Europe, Western Europe/ European origin and Asia). If the purpose of the covariate was to account for differences in region it should contain four levels, one for each region. It is not clear why only two geographical regions were considered, effectively USA and non-USA.
- The company states the combined covariate adjusted for differences between the studies and regions.[1] This assumption introduces a potential two-fold systematic error in the analysis. Firstly, in regression analysis when two covariates are known to correlate highly with each-other (conceptually or statistically) only one should be included, as this multicollinearity could distort the results. Nevertheless, this must be justified with a rationale and a sensitivity analysis using one of the covariates at a time. No such evidence has been provided.
- Secondly, the company assumes that by essentially 'dividing' the pooled KALM-1 and KALM-2 population using one combined characteristic in 3 parts (one making up the entire population of KALM-1 and the rest two dividing the KALM-2 population to USA and non-USA region) can adjust for differences between studies and regions at the same time. This simplification assumes that the potential effect of the characteristic 'study' (coefficient estimate) and the effect of the characteristic 'region' are equal. There is no evidence provided to that end i.e., sensitivity analysis.
- The company reports that MI was executed separately for the two studies. On the other hand, the consequent logistic regression was executed for the two studies pooled.[10] For consistency in the methodological approaches the MI analysis should have also been executed for the two studies pooled, preferably using the same parameters.

Nonetheless, the EAG maintains that the standard practice to combine the results of two separate trials is to execute a meta-analysis. The differences between studies will not only come from them being undertaken in different regions. Pooling the data before the initial statistical analysis is carried out can

introduce systematic errors. At the same time, it is misleading as the two trials are effectively presented and handled as one. A meta-analysis acknowledges that results from different trials are brought together using statistical methods, with the limitations that this creates.

In light of the EAG's meta-analysis[5] generating a larger effect size compared with the estimate as reported by Topf et al., (2022),[9] the company concludes that any potential bias in their analysis underestimates (rather than overestimates) the treatment effect.[1] However, the EAG maintains that proper critique of methods should apply irrespective of the results generated.

The company concludes that "*The cost-effectiveness results are consistent when trial data are used individually and pooled*",[1] with which the EAG does not agree.

Key issue 8:

As part of their TE response, the company provided statements but no new data or new analyses.[1]

The EAG has indicated in their report that they favour the use of the patient-level data to derive transition probabilities, rather than the base case approach favoured by the company.[5] The company's approach basically uses only aggregate data (percentage of patients in each itch score at base line, and the mean change from baseline in itch score after 4, 8, 12 and 64 weeks) to simulate how the average patient will move through the health states.

In their TE response,[1] the company draws attention to the probabilistic results of the EAG preferred base case which were materially higher than the EAG deterministic base case results (£41,157 versus £35,048) and the explanation given in the EAG report that "*this is due to the skewness in the distribution around the transition probabilities whenever these are very close to zero, i.e., a 0% transition in the deterministic analysis will become a small but non-0% transition in the PSA*".[5] The company explains that they raised this concern at submission wherein creating a matrix which calculates the probability of moving from any one state to each of the other states can result in small observation numbers estimating a single probability value, which may lead to unrealistic outcomes.[1]

The EAG however, does not consider the probabilistic sensitivity analysis (PSA) outcomes that follow from using the observed data to be unrealistic, but an appropriate reflection of the uncertainty around the estimates of some transition probabilities. The fact that the probabilistic incremental cost-effectiveness ratio (ICER) is larger than the deterministic ICER does not point to any shortcomings of using the observed data without additional assumptions, instead it correctly reflects the fact that current zeros in the transition matrices could well be small non-zero values and that the uncertainty about these values can only be in one direction as small negative values are by definition not possible.

In their response,[1] the company further emphasises the information provided at clarification about the distribution of patients in the difelikefalin arm at weeks 12 and week 64 when using the "*change in state*" transition estimates (base case) and the directly observed transitions as estimated from the pooled patient level data from the KALM trials (figures in company TE response form[1]). The EAG agrees with the company that these distributions are quite similar, and this is reflected in the modest increase in ICER when using patient-level data instead of aggregate data, as can be seen in Table 6.6 of the EAG report (company base case £23,277 per QALY, EAG approach direct estimation £25,792 per QALY).[5]

Despite the statements presented by the company, the EAG maintains their position that transition probabilities estimated directly from the observations is preferable to estimated and simulated from aggregate data.

Key issue 9:

For this key issue, the company provided explanation and a scenario analysis in their response to the TE.[1] However, the explanation provided focusses on the step of imputing the missing data, and though the added details are of interest, it does not address the issue raised by the EAG. In Table 1.10 of the EAG report, the problem relating to Key issue 9 was summarised as *"Information is required how analyses per complete dataset were combined in order to find the final estimates, and how uncertainty was estimated (which should be a function of within dataset variation and between dataset variation)"*.[5]

The company has explained during clarification that through multiple imputation, 20 complete datasets were generated.[6] It is not clear at all how, based on these 20 datasets, the pooled estimates of the transition matrices were derived and how the uncertainty around these pooled estimates was derived and used in the PSA. And as mentioned in the EAG report, it is important and informative to know what the relative importance is of the within-dataset variation versus the between-dataset variation in the estimation of the overall uncertainty.[5]

Furthermore, the company did provide a scenario analysis where the unimputed data were used to derive the transition matrices. This showed that for the EAG approach of estimating transition probabilities based on patient-level data the ICER would increase from £35,048 to £37,830 whereas the revised company base case ICER would decrease very slightly from £24,552 to £24,516.[1, 5]

Key issue 10:

For this key issue, the company provided additional information and scenario analyses in their response to the TE.[1]

The issue described by the EAG is that in the company base case analysis, a treatment waning was modelled for the ECM arm equal to a 5% probability of deteriorating per year while no waning impact was assumed for the difelikefalin arm. In absence of further real-world evidence to support the waning impact of the ECM and/or the lack of waning throughout the years with difelikefalin, the EAG considers this assumption uncertain.[5]

In response to this, the company has set out to identify other assessments in which waning of treatment effect might play a role.[1] To this end, the company identified three NICE Technology Appraisals (TAs) (TA534,[11] TA681,[12] and TA814[13]) regarding atopic dermatitis, in which itching was assumed to play an important part. In these three TAs, treatment waning was discussed extensively and the committees accepted for these TAs the assumption that for the comparator group, treatment would wane over five years to the level observed at baseline or close to it.

As a result of these findings, the company has explored two scenarios with regards to waning of the ECM treatment effect.[1] A waning effect is applied separately to the none, mild, moderate, and severe CKD-aP health states so that the proportion of people in each health state at year 5 (sensitivity analysis 1) and at year 10 (sensitivity analysis 2) equals the proportion of people in each health state at baseline. This approach assumes that the rate of waning may vary depending on CKD-aP severity.

When estimating the waning rates per health state for each of these two scenarios, it became clear that these rates are much higher than the 5% waning per year for each health state that was assumed in the company base case. Consequently, the company has increased that rate from 5% to 10% for a revised company base case.[1]

The EAG considers the approach taken by the company to find external information for the waning rates reasonable and agrees that the waning rates used in the scenario analyses (which were seen as acceptable in previous TAs) are much higher than the 5% rate that was used in the previous company base case. As such, the EAG accepts the choice of 10% waning per year in the ECM group.

However, it is clear that similarly there is currently no evidence that no waning at all will occur in the difelikefalin group.

Key issue 11:

For this key issue the company has decided to adopt the position of the EAG that there is not enough evidence (yet) to assume that mortality rates are higher for patients with worse itching, and that a reduction in itching will thus lead to a decrease in mortality.

Additional issue from the EAR (Table 3):

The company tabulated details of two further issues arising from the EAG report (Table 3 in the TE response).[1]

The first issue related to the rationale and validity of the variables used in the regression analysis whilst the second focused on whether other variables considered. In their response the company does not offer any additional evidence or explanation to these inquiries.

In addition, the company stated that "*While the demographic subgroups defined by age, sex and race were not identified pre-hoc as potential effect modifiers, the robustness of the treatment effects across subgroups were assessed in the ISE/SCE*" (page 21 of the TE response).[1] Indeed in the study by Topf et al. 2022,[9] a subgroup analysis based on demographic variables is presented for the two studies pooled where a notable difference between the three race subgroups was observed: odds ratio (OR) (95% CI) 1.84 (1.27, 2.67) (white); 2.82 (1.63, 4.87) (black/African American); and 0.75 (0.29, 1.92) (other). The same trends were observed in the subgroup analysis that was presented for the two trials separately after the request of the EAG.[6] These issues have been reported in detail in section 3.2.1.1; 2 (Applicability – race, gender and age) of the EAG report[5] and have not been addressed by the company.

The company provided an additional table reporting the mean and 95% CI of the baseline WI-NRS by subgroup (intention-to-treat population) stating that "*that baseline itch levels were homogeneous for all subgroups*".[1] The similarity between the reported baseline WI-NRS speaks to the validity of the concerns raised by the EAG in relation to the results of the subgroup analysis.

New model results and scenario analyses

The company made changes to their base case model, one regarding the rate of waning of treatment effect in the ECM group, and one regarding the impact of itching on mortality. Table 1 presents the original base case ICER, the impact of each of the changes separately and the combined impact in a revised base case ICER. Since one change decreases the ICER whilst the other change increases the ICER, the net difference between the original and revised base case ICER is fairly small.

Table 1 Changes in company base case after TE

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Company's base case ICER			£23,277
Key issue 10	The company applied a 5% treatment waning effect for the ECM arm only.	The company applied a 10% treatment waning effect for the ECM arm only.	£21,072
Key issue 11	The company included a mortality adjustment for the moderate, severe, and very severe CKD-aP health state, and excluded dialysis costs.	The company has removed this mortality adjustment and included dialysis costs.	£31,865
Company's updated base case	Incremental QALYs: [redacted]	Incremental costs: [redacted]	£24,552

Source: Table 4 Company response to TE[1]
CKD-aP = chronic kidney disease-associated pruritis; EAR = Evidence Assessment Report; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life year; TE = Technical Engagement.

The EAG was convinced by the information provided by the company regarding treatment waning, and thus adopted the same rate of waning as the company. Thus, the only remaining difference between the company base case and the EAG base case relates to the approach to estimate the transition probabilities for the model.

Table 2 Changes in EAG base case after TE

Key issue(s) in the EAR that the change relates to	EAG's base case before technical engagement	Change(s) made in response to company response technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
EAG's base case ICER			£35,048
Key issue 10	The EAG assumed no treatment waning effect for the ECM arm, in line with the difelikefalin arm.	The EAG applied a 10% treatment waning effect for the ECM arm only.	£26,646
EAG's updated base case	Incremental QALYs: [redacted]	Incremental costs: [redacted]	£26,646

EAG = Evidence Assessment Group; EAR = Evidence Assessment Report; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life year; TE = Technical Engagement

Table 3 presents the results of the various subgroup and scenario analyses that were mentioned under Key issues 2, 7, 9 and 10.

Table 3 Results of various scenario and subgroup analyses for key issues

	Company			EAG		
Subgroup/Scenario	**Incr. Costs (£)**	**Incr. QALYs**	**ICER (£/QALY)**	**Incr. Costs (£)**	**Incr. QALYs**	**ICER (£/QALY)**
Revised Base case	[redacted]	[redacted]	£24,552	[redacted]	[redacted]	£26,646
Subgroup anti-inch medication at baseline	[redacted]	[redacted]	£25,339a	[redacted]	[redacted]	£26,461
Subgroup no anti-itch medication at baseline	[redacted]	[redacted]	£26,956a	[redacted]	[redacted]	£27,455
Subgroup KALM-1 only	[redacted]	[redacted]	£30,389			-b
Subgroup KALM-2 only	[redacted]	[redacted]	£23,115			-b
No MI for missing data	[redacted]	[redacted]	£24,516	[redacted]	[redacted]	£28,366
ECM waning effect applied to match baseline at year 5	[redacted]	[redacted]	£18,613	[redacted]	[redacted]	£19,248
ECM waning effect applied to match baseline at year 10	[redacted]	[redacted]	£20,668	[redacted]	[redacted]	£21,625

a: These estimates from the company could not be reproduced by the EAG, the Excel file for these analyses presented the base case ICER
b: the EAG does not have the transition matrices in order to perform these subgroup analyses
EAG = Evidence Assessment Group; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; MI = multiple imputation; QALY = quality-adjusted life year

References

[1] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Technical engagement response form*, 2022 [accessed 19.12.22]

[2] Vifor Pharma UK, Initiate Consultancy Ltd. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Technical engagement response. Clinical SLR: clinical efficacy and safety of difelikefalin for chronic kidney disease-associated pruritus* 2022 [accessed 19.12.22]

[3] National Institute for Health and Care Excellence. *Single Technology Appraisal: Difelikefalin for treating pruritus in people having haemodialysis: final scope [Internet]*. London: NICE, 2022 [accessed 20.6.22] Available from: https://www.nice.org.uk/guidance/gid-ta10793/documents/final-scope

[4] Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016; 69:225-34

[5] O'Meara S, Qendri V, Perry M, Danopoulos E, Kvamme I, McDermott K, et al. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: a Single Technology Assessment*. York: Kleijnen Systematic Reviews Ltd, 2022

[6] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Response to request for clarification from the EAG*, 2022 [accessed 20.9.22]. 82p.

[7] Wolfe R, Hanley J. If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2. CMAJ 2002; 166(1):65-6

[8] Cara Therapeutics Inc. *CR845-CLIN3102 Clinical study report – double-blind phase (AIC): a multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of intravenous CR845 in hemodialysis patients with moderate-to-severe pruritus, with a 52-week open-label extension [PDF provided by company]*, 2020 [accessed 11.8.22]

[9] Topf J, Wooldridge T, McCafferty K, Schomig M, Csiky B, Zwiech R, et al. Efficacy of difelikefalin for the treatment of moderate to severe pruritus in hemodialysis patients: pooled analysis of KALM-1 and KALM-2 phase 3 studies. Kidney Med 2022; 4(8):100512

[10] Vifor Pharma UK. *Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]: Document B [revised after clarification]. Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 20.9.22]

[11] National Institute for Health and Care Excellence. *Dupilumab for treating moderate to severe atopic dermatitis: NICE technology appraisal guidance 534 [Internet]*. London: NICE, 2018 [accessed 12.1.23] Available from: https://www.nice.org.uk/guidance/ta534/

[12] National Institute for Health and Care Excellence. *Baricitinib for treating moderate to severe atopic dermatitis: NICE technology appraisal guidance 681 [Internet]*. London: NICE, 2021 [accessed 12.1.23] Available from: https://www.nice.org.uk/guidance/ta681/

[13] National Institute for Health and Care Excellence. *Abrocitinib, tralokinumab or upadacitinib for treating moderate to severe atopic dermatitis: NICE technology appraisal guidance 814 [Internet]*. London: NICE, 2022 [accessed 12.1.23] Available from: https://www.nice.org.uk/guidance/ta814/

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

[

Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]

Addendum - EAG critique of the company's response to Technical Engagement

Produced by Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC+)

Authors

Susan O'Meara, Reviews Manager, KSR Ltd, United Kingdom (UK)
Venetia Qendri, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, the Netherlands (NL)
Mark Perry, Reviews Manager, KSR Ltd, UK
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Ingelin Kvamme, Health Economics Researcher, Institute for Medical Technology Assessment (IMTA), EUR, the NL
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Caro Noake, Information Specialist, KSR Ltd, UK
Nigel Armstrong, Health Economist, KSR Ltd, UK
Maiwenn Al, Health Economics Researcher, ESHPM, EUR, the NL
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to

Susan O'Meara, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 25/01/2023

Introduction

This appendix accompanies the addendum dated 12 January 2023 that contains the EAGs critique of the company's updated analyses and base-case assumptions, provided in the company's response to technical engagement. In this appendix, the probabilistic sensitivity analyses (PSA) are presented for both the company's and the EAG's updated base case.

Company new base case and PSA

Table 1 and 2 presents the company's deterministic and probabilistic revised base case results, respectively, and Figure 1 shows the results of 1000 iterations of the PSA. The plot clearly shows that the probabilistic ICER is the same as the deterministic ICER.

The acceptability curve in Figure 2 shows that the probability that difelikefalin combined with ECM is cost effective at thresholds of £20,000 and £30,000 per QALY gained is 22% and 75%, respectively.

Table 1 Updated deterministic company base case after technical engagement, with PAS

Technologies	Total costs	Total LYG	Total QALYs	Inc. costs	Inc. LYG	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	4.78	[redacted]				
ECM	[redacted]	4.78	[redacted]	[redacted]	0.00	[redacted]	£24,552

Based on updated company model 12-12-2022
DFK = difelikefalin; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; LYG = life years gained; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

Table 2 Updated probabilistic company base case after technical engagement, with PAS

Technologies	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	[redacted]			
ECM	[redacted]	[redacted]	[redacted]	[redacted]	£24,299

Figure 1 Company PSA results presented on the CE plane (1,000 iterations) after technical engagement

[redacted]

Figure 2 Company acceptability curve after technical engagement

EAG new base case and PSA

Table 3 and 4 presents the deterministic and probabilistic EAG's revised base case ICERs, respectively. Figure 3 shows the results of 1000 iterations of the PSA. In this plot it can be seen that the probabilistic ICER is higher than the deterministic ICER, as was the case for the original EAG preferred base case. This is due to the skewness in the distribution around the transition probabilities whenever these are very close to zero, i.e., a 0% transition in the deterministic analysis will become a small but non-0% transition in the PSA.

Figure 4 presents the acceptability curve, which shows that the probability that difelikefalin combined with ECM is cost effective at thresholds of £20,000 and £30,000 per QALY gained is 4% and 56%, respectively, using the EAG base case assumptions.

Table 3 Updated deterministic EAG base case after technical engagement, with PAS

Technologies	Total costs	Total LYG	Total QALYs	Inc. costs	Inc. LYG	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	4.78	[redacted]				
ECM	[redacted]	4.78	[redacted]	[redacted]	0.00	[redacted]	£26,646
Based on updated company model 12-12-2022 DFK = difelikefalin; ECM = established clinical management; ICER = incremental cost-effectiveness ratio; LYG = life years gained; PAS = Patient Access Scheme; QALYs = quality-adjusted life years							

Table 4 Updated probabilistic EAG base case after technical engagement, with PAS

Technologies	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (£/QALY)
DFK plus ECM	[redacted]	[redacted]			

Technologies	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (£/QALY)
ECM	[redacted]	[redacted]	[redacted]	[redacted]	£29,121

Figure 3 EAG PSA results presented as CE plane (1,000 iterations) after technical engagement

Figure 4 EAG acceptability curve after technical engagement

Difelikefalin for the treatment of moderate-to-severe pruritus associated with chronic kidney disease in adult patients receiving in-centre haemodialysis [ID3890]

Answers to questions from the PMB on 27 January 2023

PREPARED BY: MAIWENN AL – ERASMUS UNIVERSITY, PART OF THE KSR TEAM

1. **Why did the severe and very severe group get assigned the same utility and costs?**

In the model, states were defined based on total 5-D Itch Scale score.

Table 42 of CS: Health state definitions by outcome measure scores

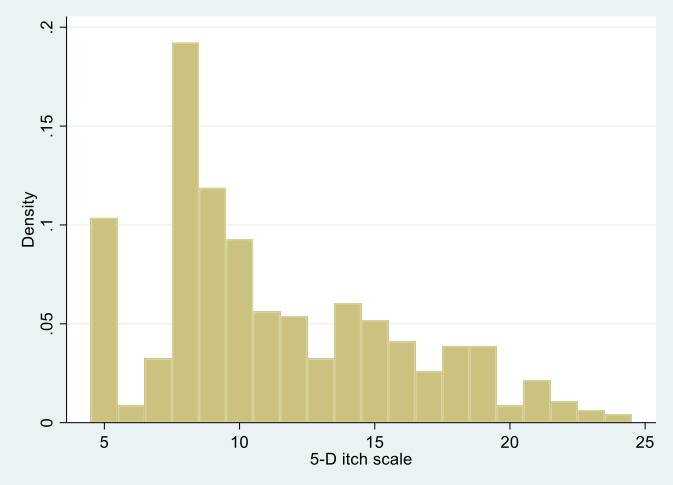
Health state	**WI-NRS score**	**5-D Itch Scale (total score)**
None	0	5-8
Mild	1-3	9-11
Moderate	4-6	12-17
Severe	***7-8***	***18-21***
Very severe	***9-10***	***22-25***
Abbreviations: WI-NRS: Worst Itch Numeric Rating Scale		

Treatment is intended for patients in the states Moderate, Severe and Very severe, and from the KALM1 and KALM2 studies the distribution of patients over these states at t=0 was 55.28%, 34.17% and 10.55%, respectively.

From the company submission: "*In the mapping study, the severe and very severe (unbearable) populations were merged, given the small numbers of observations in each group.*"

From the mapping study we find that 71 patients (19%) were classified as severe and unbearable. Below is the figure showing the distribution of the patients in the mapping study, and we can see that the group of 71 patients with severe and very severe is mostly comprised of patients in the severe group.

Figure 3 of Mapping study report: Distribution of 5-D itch scale (range 5-25)



This small number of patients in the unbearable group in the mapping study is indeed a reasonable reason to estimate resource use and utility for both the unbearable and the severe group combined. This does mean that from a practical point of view there is no reason for the model to distinguish between severe and very severe. However, it is very likely that the model was developed before the results of the mapping study were available. Also, in the initial model that was submitted, it was assumed that higher itching scores would lead to increased mortality, so separating the 2 severe states had an impact on the LYs gained with difelikefalin. In the current model version of the company, this impact of itching on mortality is no longer included.

2. **Scenario analysis alternative source of utilities**

To address the uncertainty around the utility values used in the model, a scenario analysis has been done using the utility values for the 5 CKD-aP health states derived from the SHAREHD study. Below we can see that using these alternative values for the utilities has a small impact in the ICER, with the ICER increasing slightly for the company base case, in which transitions are based on aggregated estimates of efficacy, whereas the EAG ICER, based on the patient-level transition data, remains the same.

Note that the number of QALYs per treatment arm increases by using the SHAREHD utility values, by about 0.35, but this impact becomes mostly invisible when looking at the incremental QALYs.

	Company			**EAG**		
Subgroup/Scenario	**Incr. Costs (£)**	**Incr. QALYs**	**ICER (£/QALY)**	**Incr. Costs (£)**	**Incr. QALYs**	**ICER (£/QALY)**
Revised Base case	[redacted]	[redacted]	£24,552	[redacted]	[redacted]	£26,646
SHAREHD as source utility values	[redacted]	[redacted]	£26,473	[redacted]	[redacted]	£26,504
No MI for missing data	xxxxxx	xxxx	£24,516	xxxxxx	xxxx	£28,366

3. **Interpretation last bullet point slide 33**

The last bullet point on slide 33 reads: "*Company scenario analysis showed that estimating transition probabilities based on patient-level data increases the EAG ICER whereas the revised company base case ICER would decrease slightly"*

In their response to TE, the company wrote: "*Additional scenario analyses have been conducted without MI to account for missing data and presented below in the company's updated base case and EAG preferred base case. …. In the company's updated base case, the cost-effectiveness results are consistent. In the EAG base case the cost-effectiveness estimates have a more material increase (approx. £2,800/QALY). The company believe this is due to the method of transition matrix estimation as the reliance on a small number of observations informing the Cycle 4 transition is more heavily weighted, increasing the uncertainty of the outcomes."*

So, a revised wording at the bullet point could be: "*Company scenario analysis showed that estimating transition probabilities using non-imputed data increases the EAG ICER slightly whereas the revised company base case ICER would remain the same".*

Note that I have included the results of that scenario analysis in the table above for reference, based on the updated base cases for the company and the EAG.