Single Technology Appraisal

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Contents:

The following documents are made available to stakeholders:

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

1. **Company submission** from Boehringer Ingelheim Ltd

2. **Clarification questions and company response**

3. **Patient group, professional group and NHS organisation submissions** from:
   a. UK Clinical Pharmacy Association – Heart Failure Committee

4. **Expert personal perspectives** from:
   a. Sarah Worsnop – patient expert, nominated by Pumping Marvellous Foundation
   b. Nick Hartshorne-Evans – patient expert, nominated by Pumping Marvellous Foundation
   c. Lisa Anderson – clinical expert, nominated by British Society for Heart Failure
   d. Andrew Ludman – clinical expert, nominated by British Cardiovascular Society

5. **External Assessment Report** prepared by BMJ-TAG

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

7. **External Assessment Report first addendum** prepared by BMJ-TAG

8. **External Assessment Group summary of direct and indirect treatment effects** prepared by BMJ-TAG

9. **External Assessment Report second addendum** prepared by BMJ-TAG

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

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Document B

Company evidence submission

14th September 2022

File name	Version	Contains confidential information	Date
ID3945_Empagliflozin_Document B_v8.0_AIC_CIC.docx	**8.0**	**Yes**	**14th September 2022**

Contents

Tables and figures

List of abbreviations

Acronym	Definition
ACEI	Angiotensin-converting enzyme inhibitor
AE	Adverse event
AESI	Adverse event of special interest
AF	Atrial fibrillation or atrial flutter
AICc	Akaike Information Criterion with a correction for finite sample size
ARBs	Angiotensin receptor blockers
ARNI	Angiotensin receptor-neprilysin inhibitor
BMI	Body mass index
CEA	Cost-effectiveness analysis
CKD	Chronic kidney disease
CKD-EPI	Chronic kidney disease - Epidemiology collaboration equation
COPD	Chronic obstructive pulmonary disease
CPRD	Clinical Practice Research Datalink
CRM	Cardio-renal-metabolic
CRS	Cardio-renal syndrome
CSR	Clinical study report
CT	Computerised tomography
CV	Cardiovascular
DM	Diabetes Mellitus
ECG	Electrocardiogram
eGFR	Estimated glomerular filtration rate
EF	Ejection fraction
ESC	European Society of Cardiology
EMA	European Medicines Agency
GP	General practitioner
HbA1c	Glycated haemoglobin
HES	Hospital Episode Statistics
HF	Heart failure
HFmrEF	Heart failure with mildly reduced ejection fraction
HFpEF	Heart failure with preserved ejection fraction
HFrEF	Heart failure with reduced ejection fraction
HHF	Hospitalisation for heart failure
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
IHD	Ischaemic heart disease
KCCQ-CSS	Kansas City Cardiomyopathy Questionnaire - clinical summary score
KCCQ-TSS	Kansas City Cardiomyopathy Questionnaire - total symptom score

Acronym	Definition
KCCQ-OSS	Kansas City Cardiomyopathy Questionnaire - overall summary score
EF	Left ventricular ejection fraction
MA	Marketing authorisation
MACE	Major adverse cardiovascular events
MI	Myocardial infarction
MRA	Mineralocorticoid receptor antagonists
MRI	Magnetic resonance imaging
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NYHA	New York heart association
RCT	Randomised controlled trial
RS	Randomised set
RWE	Real-world evidence
SAE	Serious adverse event
SD	Standard Deviation
SBP	Systolic blood pressure
SGLT1	Sodium-glucose co-transporter-1
SGLT2i	Sodium-glucose co-transporter-2 inhibitor
SLR	Systematic literature review
SmPC	Summary of medicinal product characteristics
SoC	Standard of care
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
TS	Treated set
UK	United Kingdom

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

B.1.1 Decision problem

This submission addresses the cost-effectiveness, comparative effectiveness, clinical efficacy and safety of empagliflozin versus standard of care (SoC) in adult patients with chronic HF (EF >40%).

Empagliflozin is currently recommended for the treatment of the National Health Service (NHS) England patients with heart failure with reduced ejection fraction (HFrEF), i.e., left ventricular ejection fraction (EF) ≤40% (National Institute for Health and Care Excellence [NICE] appraisal TA773) (1). To achieve a recommendation that is consistent with the proposed MHRA marketing authorisation (MA) extension for empagliflozin; that is, "*adult patients with symptomatic chronic heart failure*" across the EF spectrum; this company submission considers evidence that reflects the remaining patient population not included in TA773, those chronic HF patients with EF >40% (Table 1**Error! Reference source not found.**) (1) (2). A MHRA Marketing Authorisation to cover the whole EF spectrum of chronic HF was issued in June 2022 (Table 2). A positive recommendation in the remaining chronic HF patient population (i.e. EF >40%) would result in the inclusion of empagliflozin in the NICE clinical guideline, NG106, as a recommended treatment option for all adults with chronic symptomatic HF, regardless of EF (3).

Table 1. MA and NICE recommendation status of empagliflozin

Indication	Date of MA received	NICE recommended/ date of NICE recommendation
Treatment of insufficiently controlled T2DM as an adjunct to diet and exercise in adults: • as monotherapy when metformin is considered inappropriate due to intolerance • in addition to other medicinal products for the treatment of diabetes	22 May 2014	Yes
Treatment of symptomatic chronic HF and EF ≤40% (HFrEF) in adults	30 July 2021	9 March 2022 (Final Guidance)
Jardiance is indicated in adults for the treatment of symptomatic chronic heart failure.	13th June 2022	The focus of this company submission ID3945

Abbreviations: HFrEF, heart failure with reduced ejection fraction; EF, left ventricular ejection fraction; MA, marketing authorisation; NICE, National Institute for Health and Care Excellence; T2DM, type 2 diabetes mellitus.
Reference: (1, 2, 7, 8)

This company submission represents a low risk to the committee in making a recommendation for routine commissioning for HF patients with an EF ≥40%. In TA773, the committee were satisfied that most uncertainties in the evidence were addressed during technical engagement and the committee did not explore these further during the committee meeting in December 2021. The only outstanding uncertainty discussed at the committee meeting was a comparison of the benefits of empagliflozin vs dapagliflozin, which is not relevant for this appraisal. In TA773, the BMJ Evidence Assessment Group (EAG) provided a thorough critique of the evidence, based on the EMPEROR-Reduced and PULSE (CPRD) study. The same cost utility and budget impact models will be presented in this submission, however utilizing the pivotal EMPEROR-Preserved trial instead. The main distinction between the EMPEROR-Reduced and EMPEROR-Preserved trials is in the inclusion criteria, with the former having enrolled patients with a baseline EF <40%) and the latter having patients with EF ≥40% (5, 6). Both trials were phase III international, multicentre, randomised, double-blind, parallel-group, placebo-controlled trials (4-6).

To provide the committee with confidence that empagliflozin represents value for money, and thus is suitable for routine commissioning, the company has proactively pressure tested the key uncertainties in the economic evidence identified by the BMJ EAG during Technical Engagement in TA773 (Section. B.3.10.3). We hope that this will accelerate the evidence appraisal process. Broadly, these uncertainties focused on whether the cost utility model accurately predicted the rate of deaths, hospitalisations and treatment discontinuation compared to the rate observed in

EMPEROR-Reduced. Mortality and hospitalisation rates were also compared to PULSE to ascertain the generalisability of the model compared to UK clinical practice. The BMJ also queried whether treatment waning could be expected over time, and whether this impacted the incremental cost effectiveness ration (ICER). PULSE will be used again to validate the outcomes predicted from the economic models. The company show in Section B.3.10.3 that the ICER in HF patients with an EF≥40% is stable across a broad range of scenarios and assumptions, providing certainty for the committee.

Table 2. The decision problem

Parameter	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults for the treatment of symptomatic chronic heart failure (EF >40%)	Same	Not applicable
Intervention	Empagliflozin	Same	Not applicable
Comparator(s)	Established clinical management without empagliflozin, including but not limited to loop diuretics, calcium-channel blockers, amiodarone and anticoagulants	The evidence case is for empagliflozin as an add-oin to established clinical management. Empagliflozin does not replace established clinical management.	Not applicable
Outcomes	The outcome measures to be considered include: • symptoms of heart failure • hospitalisation for heart failure • all-cause hospitalisation • mortality • cardiovascular mortality • kidney function • adverse effects of treatment • health-related quality of life	Same	Not applicable
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of	Same	Not applicable

Parameter	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	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. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.		
Subgroups to be considered	Not included in the draft scope	No subgroups were considered separately in the economic analysis	Not applicable

Parameter	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Special considerations including issues related to equity or equality	Not included in the draft scope	Broad prescribing of SGLT2i in primary and secondary care for HF, regardless of EF, could reduce the inequality in terms of access to heart failure care in the UK	The socio-economic inequalities in CV disease present a major and persistent UK public health challenge. The UK-based population studies demonstrate that socio-economic deprivation is a strong risk factor for the development of HF and adverse HF outcomes (9, 10). Individuals in the lowest socio-economic group are 1.61 times more likely to experience incident HF than the most affluent individuals and on average, at a 3.5 years younger age with a greater comorbidity burden at the time of HF symptom onset (9). Findings from Conrad et al. (2018) report socio-economic inequalities among all age bands and by sex in the most deprived region, which were twice as high in younger adults (IRR 2.56; 95% CI, 2.30-2.85 in the 45-54 years age group vs. 1.17; 95% CI, 1.13-1.22 in the >85 years age group) (9).

Parameter	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			Socio-economic status has an impact on access to secondary care in the UK, and subsequently access to HF treatments. Moscelli et al. (2018) reported a statistically significant difference in waiting times across socio-economic groups for patients who attend the same hospital: patients living in more income-deprived areas waited longer (35% difference, or 43 days) than patients who lived in areas of a lesser deprivation (11). In addition to waiting longer, coronary heart disease patients in a lower socio-economic class were admitted to hospital less often than those in a higher class (11). McCartney et al. (2013) reported on a prospective study of 7,049 men and 8,353 women in the west of Scotland who were followed up for 37 years; the likelihood of a hospital admission for CV disease was 21% higher for female patients in highest socio-economic class than patients in lowest class. Those patients in social class IV and V (partly skilled and unskilled occupations) also stayed 25% longer in hospital than social class I and II (professional, managerial and technical occupations) (736 vs. 589 bed day/1,000 person-years, respectively) (12). These studies indicate that if patients in lower socio-economic classes utilise secondary care less often, their opportunity to access HF medications would also be lower, if they are solely prescribed in secondary care.

Parameter	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			BI support the UK Government's and NICE's commitment to the reduction of health inequalities, reiterated in the recent NICE five-year strategy publication (13, 14). Principle 9 of NICE's Social Value Judgments states that due regard must be given to reducing inequalities. It states that equality should be considered in relation to the nine protected characteristics in the Equality Act 2010 (age, disability, gender reassignment, race, religion or belief, sex, sexual orientation, marriage and civil partnership, pregnancy and maternity) and socio-demographic factors (14). Further, the COVID-19 Marmot review aims to reduce the widened gap in health inequalities and build a fairer society post pandemic (15). Broad prescribing of SGLT2i across primary and secondary care can support the reduction in disparity in terms of access to HF care across socio-economic groups within the UK. Together with TA773, this appraisal further supports this objective by providing a treatment option for those patients regardless of EF.

Abbreviations: CV, cardiovascular; HF, heart failure; IRR, incidence rate ratio; EF, left ventricular ejection fraction; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; SGLT2i, sodium-glucose co-transporter 2 inhibitors; UK, United Kingdom.

B.1.2 Description of the technology being appraised

- Empagliflozin is an orally bioavailable, selective sodium-glucose co-transporter-2 inhibitor (SGLT2i) which has cardioprotective effects and improves HF-related outcomes (16, 17).

Empagliflozin's mechanism of action, MA, indication, mode of administration and list price are summarised in Table 3. Appendix C includes the draft summary of medicinal product characteristics (SmPC) for empagliflozin.

Table 3. Technology being appraised

UK approved name and brand name	Empagliflozin (JARDIANCE®)
Mechanism of action	Empagliflozin is an orally bioavailable, reversible, highly potent and selective inhibitor of SGLT2 (16). Through SGLT2 inhibition, empagliflozin simultaneously reduces renal reabsorption of glucose and sodium in the proximal tubules of the kidney and leads to increased urinary excretion of glucose and moderate natriuresis. The molecular bases of empagliflozin's cardioprotective and nephroprotective effects are unknown; however, accumulating evidence suggests several distinct mechanisms are involved, including: • osmotic diuresis and natriuresis resulting in lowering of arterial pressure and stiffness and improvement in ventricular loading • improved myocardial and renal metabolism via switch to ketone bodies as the energy source • prevention of adverse cardiac remodelling through inhibition of inflammation, fibrosis and cardiomyocyte cell death • direct inhibition of the Na+/H+ exchanger in myocardium, leading to reduction or reversal of cardiac injury, fibrosis and systolic dysfunction • prevention of ischaemia/reperfusion injury through decrease in calmodulin kinase II activity (17, 18).
Marketing authorisation/ CE mark status	Empagliflozin currently holds the EMA MA and is recommended by NICE for the treatment of T2DM as a monotherapy (25 May 2016) or as a combination therapy with insulin or other antidiabetic drugs (25 March 2015) (2, 7, 8). For treatment of chronic HFrEF, empagliflozin received the EMA Marketing Authorisation on 30 July 2021 and the NICE recommendation was published on 09 March 2022 (1, 2). On 27 January 2022, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion, recommending a change to the terms of the MA for the medicinal product Jardiance to include all adult patients with chronic symptomatic heart failure (2). The EMA MA was granted on 3 March 2022, and the UK MHRA MA was granted

	on 13th June 2022. The draft SmPC is provided in Appendix C (19).
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	***Indication relevant to this submission***: Empagliflozin is indicated in adults for the treatment of symptomatic chronic heart failure (pending indication expansion from the EMA and the MHRA). ***Other indications:*** • Empagliflozin is indicated for the treatment of adults with insufficiently controlled type 2 diabetes mellitus as an adjunct to diet and exercise o as monotherapy when metformin is considered inappropriate due to intolerance o in addition to other medicinal products for the treatment of diabetes. • Empagliflozin is indicated in adults for the treatment of symptomatic chronic HFrEF.
Method of administration and dosage	10 mg oral empagliflozin once daily
Additional tests or investigations	None
List price and average cost of a course of treatment	List price of a pack of 28 tablets (10 mg) is £36.59. This equates to a cost of £1.31 per tablet per day for each patient.
Patient access scheme (if applicable)	None

Abbreviations: CHMP, Committee for Medicinal Products for Human use; EMA, European Medicines Agency; HF, heart failure; HFrEF, heart failure with reduced ejection fraction; EF, left ventricular ejection fraction; MA, marketing authorisation; MHRA, Medicines and Healthcare products Regulatory Agency; NICE, National Institute for Health and Care Excellence; SGLT2, sodium-glucose co-transporter 2; T2DM, type 2 diabetes mellitus; UK, United Kingdom.

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

- Heart failure is a complex clinical syndrome caused by structural and/or functional abnormalities of the myocardium resulting in the impairment of ventricular filling and ejection of blood (20, 21).
- Heart failure presents as either acute or chronic HF. Patients who are acutely decompensated might actually have chronic HF (20).
- There are limitations to the New York Heart Association (NYHA) classification as it is dependent on clinician's interpretation and poor agreement has been identified between cardiologists (20, 22).
- The recent European Society of Cardiology (ESC) guideline (2021) classifies HF based on EF; however, there are inconsistencies regarding the definition of different classes of HF observed in clinical trials and among clinical experts (20, 23).

- Given the limitations in NYHA classification and the definition of HFrEF, HFmEF and HFpEF, the best available way to define patients in this company submission is by EF.
- This company submission provides evidence for all patients with chronic HF and EF >40%, with the overall preferred outcome to have a broad recommendation for empagliflozin for all chronic HF patients in the NICE guideline NG106.

B.1.3.1 Overview of the disease for which the technology is indicated

B.1.3.1.1 Disease overview

Clinical presentation and aetiology of heart failure

Heart failure is a complex clinical syndrome caused by structural and/or functional cardiac abnormality that results in reduced cardiac output and/or elevated intracardiac pressure, impairing the ability of the heart to function adequately and act as a pump to support physiological circulation (20, 21). Heart failure is characterised by a range of symptoms including breathlessness, fatigue, poor exercise tolerance, ankle swelling and peripheral oedema; however, none are specific to HF (21). Signs of congestion, such as jugular venous distention, gallop rhythm and displaced apical impulse, are more specific to HF and indicative of higher risk of adverse outcome although harder to detect (20, 21). Heart failure results from myocardium injury caused by a wide range of pathologies including ischaemic heart disease (IHD), congenital heart defects, hypertension and non-cardiovascular (non-CV) systemic diseases such as diabetes and severe lung disease (24). More than two-thirds of all cases of HF can be attributed to IHD, hypertension (25), obesity, chronic obstructive pulmonary disease (COPD) and rheumatic heart disease (24, 25). Less common aetiologies include cardiomyopathies, valvular disease, myocarditis, infections, systemic toxins and cardiotoxic drugs (24).

Classification

Heart failure can be classified into acute and chronic in nature (20, 26). Acute HF is a life-threatening condition, with a rapid onset of HF symptoms, typically leading to urgent hospital admissions (20, 27) and patients who are acutely decompensated may have chronic HF (20). Chronic HF refers to patients who have had HF diagnosis for at least three months and can be categorised into left or right ventricular failure (28). Most

patients with chronic HF have a systolic dysfunction caused by the left ventricle failing to pump blood efficiently (29). In some patients, diastolic dysfunction is observed due to increased left ventricular stiffness or abnormal left ventricular relaxation, resulting in reduced left ventricular filling (30). Failure of the left ventricle may lead to right ventricular dysfunction by multiple mechanisms including myocardial ischaemia involving both ventricles, increased pulmonary venous and arterial pressure, and reduced right ventricular coronary perfusion due to decreased systolic blood pressure (SBP) (29).

The NYHA classification differentiates patients based on severity of HF (20) and it is commonly used for functional classification in patients with HF in clinical practice and as an entry criterion and/or outcome measure in clinical trials (Table 4). However, there are limitations to the NYHA classification, despite it reflecting the natural course of HF. Patients rely on the clinician's subjective interpretation of their functional capacity (which can fluctuate over time) to determine the NYHA classification; however, the functional classification was found to be a poor prognostic indicator in HF (22). The validity of NYHA class as an entrance criterion or an outcome measure in clinical trials is disputed since the concordance between cardiologists assigning NYHA classes can be as low as 54% (31). This suggests that there is a poor agreement between cardiologists in differentiating patients between class II and class III (31). This is a key reason why the economic model based on health states defined by NYHA is not presented in this submission.

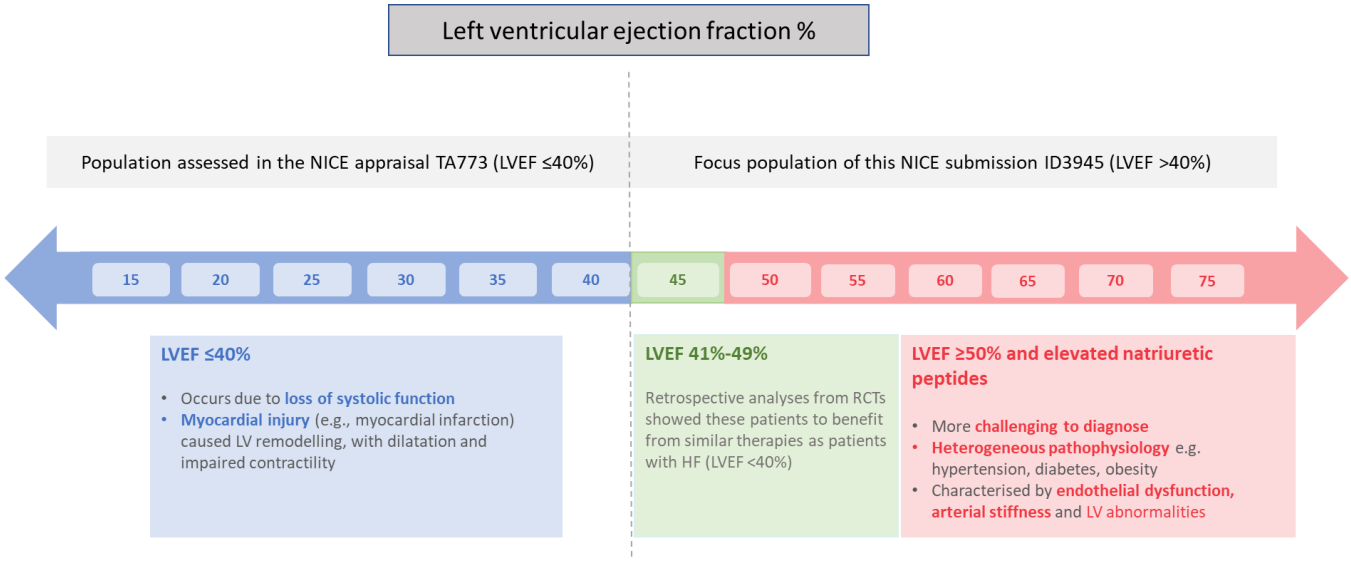
Table 4. NYHA functional classification based on severity of symptoms and physical activity

Classification	Description
Class I	No limitation of physical activity. Ordinary physical activity does not cause undue breathlessness, fatigue or palpitations.
Class II	Slight limitation of physical activity. Comfortable at rest, but ordinary physical activity results in undue breathlessness, fatigue or palpitations.
Class III	Marked limitation of physical activity. Comfortable at rest, but less than ordinary activity results undue breathlessness, fatigue or palpitations.
Class IV	Unable to carry on any physical activity without discomfort. Symptoms at rest can be present. If any physical activity is undertaken, discomfort is increased.

Abbreviation: NYHA, New York Heart Association.Reference: (20).

Alternatively, HF is categorised by EF; however, there is ambiguity across the clinical community regarding the range of EF that should be classified as HFrEF, heart failure with mildly reduced ejection fraction (HFmrEF) or heart failure with preserved ejection fraction (HFpEF). According to the 2021 European Society of Cardiology (ESC) HF guidelines, HFrEF includes HF with EF ≤40%, HFmrEF includes HF with EF 41% to 49% and HFpEF includes HF with EF ≥50% (20). The definition of HFpEF has differed in clinical trials ranging from EF >40%, ≥40% and ≥45% to ≥50% (32-38). Furthermore, clinicians have agreed that it is not well-defined or understood in the real-world practice (39, 20, 40). Given this ambiguity in nomenclature, for simplicity, this submission has defined the target population-based on EF, (those with symptomatic chronic HF with EF >40%) (Figure 1), which focuses on the population not already assessed by NICE in TA773 (chronic HF with EF ≤40%). Therefore, the preferred outcome of this submission is to have a broad recommendation for empagliflozin for all chronic HF patients in the NICE guideline NG106, regardless of EF.

Figure 1. Left ventricular chronic HF



Abbreviations: HF, heart failure; LV, left ventricular; EF, left ventricular ejection fraction; RCTs, randomised controlled trials.
Reference: (20).

B.1.3.1.2 Epidemiology

- The size of the challenge of improving outcomes for patients with HF is substantial:
  - Population growth, aging and the rising burden of T2DM, hypertension and obesity are driving the increasing global prevalence of HF (9).
  - In the UK, 920,000 people are estimated to live with HF and 200,000 people are newly diagnosed with HF every year (3, 41-43).
  - The proportion of HF patients with EF ≥40% is increasing each year with approximately half of all HF patients estimated to have HF with EF ≥40% (43).
  - The burden of HF is just as high as other chronic conditions such as some types of cancer or COPD (9, 44).
  - Heart failure is associated with a high prevalence of comorbidities, with coronary heart disease (47.8%-61.1%) and hypertension (45.7%-54.6%) being the most common (17, 45-47).

Prevalence and incidence

Heart failure is a growing public health problem driven by an increase in population size and age (9). The increasing prevalence of chronic conditions such as T2DM, hypertension and obesity also contributes to increasing HF burden (9). Approximately 64.3 million people worldwide are estimated to have HF (43). Based on 2014 data, there are more than 920,000 people with HF in the UK (9); of those, approximately 650,000 are on their general practitioner (GP)'s HF register (41). From 2002 to 2014, the prevalence of HF in the UK increased by 23% (9). In 2017, the prevalence of HF diagnosed in primary care was 2% and 5.9% among those aged 65 to 74 years and those older than 75 years, respectively, with a higher estimated prevalence in men (7.5%) compared to women (4.8%) in those older than 75 years (41). The proportion of HF patients with EF >40% is increasing each year with approximately half of all HF patients estimated to have HF with EF >40% (43). A real-world evidence (RWE) study conducted in the UK (PULSE) reported that 8.7% of all HF patients had EF >40% in 2015 which increased to 10.4% in 2019; however, this data should be interpreted with caution since a large proportion of patients had unknown EF in this RWE study (48).

The number of newly diagnosed HF cases in the UK has increased by 12% between 2002 and 2014 and there is no indication that the trend is slowing down (9). Around 176,000 to 200,000 people are newly diagnosed with HF each year in the UK, with the average age of diagnosis between 72 and 77 years (9, 41, 48). There has been a year-on-year increase in the incidence of HF since 2015. The RWE PULSE study reported that the incidence of diagnosed HF increased from 4.10 per 1,000 person-years in 2015 to 4.85 per 1,000 person-years in 2019 (48).

Prioritising the improvement of outcomes for HF patients is just as important as for other common conditions with a high burden of disease. The prevalence and incidence of HF in the UK is similar to the four most common causes of cancer combined (breast, prostate, lung and bowel) or COPD (9, 44). Between 2015 and 2017, Cancer Research UK reported the number of aggregated new cases for the four aforementioned cancers to be over 183,000 (49). Similarly, around 1.2 million people in the UK have COPD and approximately 115,000 people are newly diagnosed with COPD each year (44). The burden of HF is therefore similar to that of cancer or COPD, indicating an urgent need to improve outcomes for HF patients at scale.

Comorbidities

The cardio-renal syndrome (CRS) encompasses a spectrum of disorders of the heart and kidneys whereby the physiological interdependence of the two organs leads to their simultaneous, accelerated decline in a negative feedback cycle (50). Metabolic disturbances associated with diabetes can also lead to the pathogenesis of the CRS by causing biochemical, functional and morphological abnormalities of the heart and kidney (51). HF patients therefore often suffer from renal or metabolic comorbidities due to the overlapping risk factors for these conditions (52).

The prevalence of comorbidities is high among HF patients across the entire spectrum of EF (53). Nearly half of all HF patients have moderate to severe kidney dysfunction which increases the risk of hospitalisation or death compared to HF alone (10, 54, 55). Furthermore, nearly one-third have comorbid T2DM, also known to increase the risk of hospital admissions and cardiovascular (CV) death (10, 56). The onset of T2DM increases the risk of HF by two-fold in men and five-fold in women (57). Other common comorbidities related to HF are atrial fibrillation, valvular heart disease, IHD,

hypertension and stroke (9, 20). Some non-CV comorbidities are thyroid disorder, obesity, anaemia and COPD (9, 20). A UK population-based cohort study showed that the patients with incident HF had high comorbidity burden, with 79% patients having at least three comorbidities (9). The prevalence of common comorbidities found in the study are presented in Table 5 (9).

Table 5. Common comorbidities in patients with heart failure

Medical History	HF (%)
Hypertension	67
IHD	49
Osteoarthritis	43
Atrial fibrillation	40
Dyslipidaemia	28
CKD	24
Diabetes	22
COPD	19

Abbreviations: CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; IHD, ischaemic heart disease.
Reference: Conrad et. al. (2018) (9).

The burden of comorbidities is much higher for HF compared to other common conditions, such as cancer. A retrospective Scottish study conducted between 2002 and 2011 on adults with HF and four of the most common cancers showed that 94.5% of HF patients had comorbidities compared to 62%-80% of patients with a cancer diagnosis (Table 6 and Table 7) (45). The data reported in the tables below further demonstrate the significant burden of HF disease to patients and the NHS.

Table 6. Baseline characteristics in men from Scotland with cancer, heart failure and comorbidities

	Prostate cancer	Lung cancer	Colorectal cancer	Bladder cancer	Heart failure
Cases, n	6,795	4,693	4,239	2,082	10,309
Heart failure, n (%)	95 (1.4%)	97 (2.1%)	81 (1.9%)	41 (2.0%)	-
Cancer, n (%)	-	-	-	-	226 (2.2%)

	Prostate cancer	Lung cancer	Colorectal cancer	Bladder cancer	Heart failure
No comorbidity, n (%)	1,949 (28.7%)	1,116 (23.8%)	1,278 (30.1%)	499 (24.6%)	562 (5.5%)
Hypertension, n (%)	2,614 (38.5%)	1,515 (32.3%)	1,596 (37.7%)	801 (39.5%)	4,711 (45.7%)
Asthma, n (%)	491 (7.2%)	355 (7.6%)	286 (6.7%)	124 (6.1%)	788 (7.6%)
Coronary heart disease, n (%)	1,303 (19.2%)	1,091 (23.2%)	817 (19.3%)	488 (24.1%)	6,295 (61.1%)
Diabetes, n (%)	688 (10.1%)	562 (12.0%)	611 (14.4%)	314 (15.5%)	2,234 (21.7%)
COPD, n (%)	611 (9.0%)	1,241 (26.4%)	390 (9.2%)	237 (11.7%)	1,707 (16.6%)
Stroke or TIA, n (%)	321 (4.7%)	445 (9.5%)	245 (5.8%)	112 (5.5%)	754 (7.3%)
Previous MI, n (%)	657 (9.7%)	563 (12.0%)	442 (10.4%)	261 (12.9%)	4,448 (43.1%)
Chronic kidney disease, n (%)	550 (8.1%)	473 (10.1%)	381 (9.0%)	220 (10.8%)	1,560 (15.1%)
Atrial fibrillation, n (%)	238 (3.5%)	168 (3.6%)	162 (3.8%)	106 (5.2%)	552 (5.4%)
PVD, n (%)	388 (5.7%)	285 (6.1%)	250 (5.9%)	115 (5.7%)	2,519 (24.4%)

Abbreviations: COPD, chronic obstructive pulmonary disease; MI, myocardial infarction; PVD, peripheral vascular disease; TIA, transient ischaemic attack.
Reference: (45)

Table 7. Baseline characteristics in women from Scotland with cancer, heart failure and comorbidities

	Breast cancer	Colorectal cancer	Lung cancer	Ovarian cancer	Heart failure
Cases, n	10,760	3,610	3,859	1,234	9,131
Heart failure, n (%)	85 (0.8%)	43 (1.2%)	61 (1.6%)	15 (1.2%)	-
Cancer, n (%)	-	-	-	-	364 (4.0%)
No comorbidity, n (%)	4,115 (38.2%)	10,24 (28.4%)	769 (19.9%)	465 (37.7%)	500 (5.5%)
Hypertension, n (%)	3,259 (30.3%)	1,450 (40.2%)	1,451 (37.6%)	364 (29.5%)	4,984 (54.6%)

	Breast cancer	Colorectal cancer	Lung cancer	Ovarian cancer	Heart failure
Asthma, n (%)	945 (8.8%)	296 (8.2%)	386 (10.0%)	95 (7.7%)	925 (10.1%)
Coronary heart disease, n (%)	839 (7.8%)	499 (13.8%)	718 (18.6%)	108 (8.8%)	4,367 (47.8%)
Diabetes, n (%)	786 (7.3%)	425 (11.8%)	421 (10.9%)	89 (7.2%)	1,708 (18.7%)
COPD, n (%)	583 (5.4%)	275 (7.6%)	1,118 (29.0%)	74 (6.0%)	1,455 (15.9%)
Stroke or TIA, n (%)	445 (4.1%)	237 (6.6%)	382 (9.9%)	58 (4.7%)	1,404 (15.4%)
Previous MI, n (%)	305 (2.8%)	207 (5.7%)	292 (7.6%)	48 (3.9%)	2,665 (29.2%)
Chronic kidney disease, n (%)	265 (2.5%)	179 (5.0%)	228 (5.9%)	37 (3.0%)	722 (7.9%)
Atrial fibrillation, n (%)	316 (2.9%)	158 (4.4%)	161 (4.2%)	25 (2.0%)	2,370 (26.0%)
PVD, n (%)	238 (2.2%)	130 (3.6%)	274 (7.1%)	30 (2.4%)	740 (8.1%)

Abbreviations: COPD, chronic obstructive pulmonary disease; MI, myocardial infarction; PVD, peripheral vascular disease; TIA, transient ischaemic attack.
Reference: (45)

Risk factors for disease

Risk factors associated with CHF can be modifiable (e.g., diet and exercise) or non-modifiable (e.g., age, gender and comorbidities). Coronary heart disease, diabetes and age are strongly associated with an increased risk of HF (10). Hypertension, smoking, elevated body mass index (BMI), diet and poor physical activity are also contributing to the pathogenesis of HF (58-60).

B.1.3.1.3 Disease burden

The extent to which HF impacts patients' lives is substantial:

- Heart failure is a debilitating condition; the cardio-renal-metabolic (CRM) system-related comorbidities increases the symptom burden in HF patients (20, 61, 62).
- In the UK, HF mortality is variable and ranges between 14.4% and 26% at one year and between 48.5% and 68.1% at five years (45, 47, 63, 64).
- Comorbidities, such as chronic kidney disease (CKD), T2DM and lung disease lead to an increased number of hospitalisations and in turn are associated with an increased risk of mortality (45, 47, 63, 64).
- Heart failure is associated with a high rate of hospitalisation, especially in elder patients (45, 65-68). There is an unmet need to lower the hospitalisation rates and reduce the risk of mortality for chronic HF patients (68).

Symptomatic burden

Heart failure patients experience debilitating symptoms including breathlessness, orthopnoea, paroxysmal nocturnal dyspnoea, reduced exercise tolerance, fatigue and ankle swelling (20, 61). The interdependencies within the CRM system lead to accelerated progression of CKD and HF and this increase the symptomatic burden on HF patients (62). Around a quarter of HF patients develop T2DM and up to 50% of these patients develop CKD (62). Furthermore, there are challenges in diagnosing chronic HF in terms of availability and use of echocardiography services (69, 70). Access to echocardiography, shortage of technically trained staff and complexity of symptoms usually cause delay in commencement of treatment in HF patients (69, 70). This means that HF patients are at a high risk of decompensation or cardiac event such as breathlessness, severe peripheral oedema and chest pain (42).

Morbidity and mortality

There remains a high unmet need to reduce the risk of mortality in all chronic HF patients. The prognosis of HF remains poor and the burden of HF in the UK is similar in magnitude to that of the four most common cancers (breast, prostate, lung and bowel) combined (9, 45, 47); further evidencing the need to improve outcomes across the UK population. Estimates for 1- and 5-year HF mortality in the UK are variable, but range between 14.4% and 26% for 1-year and between 48.5% and 68.1% for 5-year

post-diagnosis (45, 47, 63, 64). A population-based cohort study in the UK estimated the 10-year mortality for HF patients to be 75.5% (63). A UK retrospective study of 241 people (41 with chronic HF [EF >40%]) indicated that 27% of patients with chronic HF (EF >40%) died within 1 year of hospital admission (71). IHD was a significant predictor of mortality among these patients (hazard ratio [HR] 7.14; 95% confidence interval [CI], 1.51 to 33.85; p=0.01) (71).

The reasons for CV mortality and non-CV mortality in patients with HF are varied, and the risk of death is significant across the EF spectrum, although some differences have been reported. Published literatures have reported that the majority of deaths in patients with chronic HF (EF >40%) are CV-related (60%-70%), and the proportion of non-CV deaths are higher in patients with higher EF (72, 73). Frequently documented CV mortalities in chronic HF (EF >40%) are sudden cardiac death (around 40%), worsening HF (20%-30%), and myocardial infarction (MI) and stroke (5%-15%) (73). Among the non-CV deaths, cancer (30%-40%) followed by infection/sepsis (around 25%) are most commonly reported (73).

Compared to other European countries, outcomes for HF patients in the UK are far worse. Significantly higher mortality rates have been observed compared to a European and global RWE studies (74-76). These studies reported that in other countries, the 1-year mortality rate ranged from 6.4% to 20.0% in chronic HF patients and rate 5-year mortality rate was 45.0% in chronic HF patients (48, 74-76).

The overall prognosis of HF patients is exacerbated when patients have other comorbidities including CKD, T2DM, atrial fibrillation and obesity (77-79). Presence of diabetes and CKD in HF patients is associated with increased mortality and hospitalisation (54, 62). A UK study reported that in patients who have both HF and CKD, hospitalisation and mortality rates increased by 11% and 17%, respectively, compared to HF patients who do not have CKD (80). Furthermore, HF patients with T2DM showed a higher mortality rate of 34% compared to those without T2DM with a mortality rate of 22% from either a CV death or hospitalisation for heart failure (HHF) (78).

Healthcare system burden

There is an unmet need to lower the hospitalisation rates and in turn reduce the risk of mortality in chronic HF patients (68) as HF is the most common cause of hospitalisation in patients over 65 years of age (45, 65-68). A 2014 study from the UK suggested that approximately 20% of patients hospitalised with HF have an EF ≥50% (81). A global study, including UK patients, found that over a median follow-up of 4.1 years, the all-cause hospitalisation rate was 56.5% among all patients with chronic HF (EF ≥45%) (78). A hospital readmission rate of 20% was also reported in the 12 months after discharge for patients with chronic HF (EF >40%) in the UK (71).

There are a number of factors contributing to the increased risk of hospitalisation and rehospitalisation. A recent Clinical Practice Research Datalink (CPRD) study identified that 80% of HF cases in England are diagnosed after emergency hospital admission for acute HF symptoms (82). The burden of hospitalisation is significant across the EF spectrum; however, some differences have been observed. Higher rates of HHF are observed in HF patients with diabetes, where the readmission rate is nearly double compared to those without diabetes (83-85). Furthermore, in patients with chronic HF (EF >40%), the challenges in diagnosis and limitation of available treatment for its management contribute to the increased risk of hospitalisations and mortality (20, 42).

B.1.3.1.4 Economic burden

- Heart failure accounts for 2% of the total NHS budget annually (86).
- The economic burden increases even further in those patients with HF and comorbidities (87, 88).
- In the UK in 2012, the direct and indirect costs of HF amounted to £2.0 billion and £888 million, respectively (69, 86, 89, 90).
- There is a need to reduce HHF costs, considering its major contribution towards total HF costs in the UK.

There is a substantial economic burden of HF in the UK, where it is estimated to annually account for 2% of the NHS budget, with 60% to 70% of the costs related to hospitalisations (89, 90). Heart failure patients accounted for 1 million inpatient bed days (representing 2% of all NHS inpatient bed days and 5% of all emergency medical

hospital admissions), with an average length of stay of 6 to 9 days and a 3-month readmission rate of 25% (86, 91).

The economic burden increases even further in HF patients with comorbidities. Notably, in HF patients, the burden of CRM-related conditions is substantial in terms of the cost burden and all-cause hospital admissions and this is further amplified in the T2DM population (87). Patients with chronic HF (EF >40%) and T2DM are reported to have higher hospitalisation rates and longer length of stays compared to those without T2DM (88).

The cost associated with hospitalisation is the main driver of UK healthcare spending in HF patients; hence, the reduction of hospitalisation frequency and duration would significantly lower the overall economic burden of HF to the NHS. In 2012, it was estimated that the direct and indirect costs of HF amounted to ~£2.0 billion and £888 million, respectively (86, 89, 90). During the last three months of a HF patient's life, the inpatient care or critical care account for more than 90% of healthcare costs (92). Although not a direct cost to the NHS, the indirect cost of informal care cost has also shown to rise with increasing rates of hospitalisation (93-95). It is important to note that the broader societal costs of hospitalisation may be even higher as informal care has further shown to significantly impact both the caregiver's leisure time and productivity; evidence suggest that caregiving responsibilities result in an average of 28 hours per week of time commitment (93-95).

B.1.3.1.5 Humanistic burden

- Heart failure has a substantial impact on patients' quality of life, affecting their physical, social, emotional and psychological well-being (42).
- The impact of HF on the quality of life of carers is also significant (95).
- This impact further evidences the need to improve outcomes for HF patients across the spectrum of EF.

Heart failure has a significant impact on patients' physical well-being across the EF spectrum. Some differences have been reported in quality of life across the EF; however, there is no consensus in the published literature. The physical well-being of HF patients was reported in several UK studies, where patients experienced a range

of symptoms including breathlessness, reduced sleep quality, frailty, cognitive/psychomotor impairment, respiratory symptoms and chest pain (20, 61, 96). One UK study reported a continuous quality of life difference in chronic HF patients compared to those without HF, where on an average, a 16% reduction in physical activity was observed (96). Another UK study showed that a higher proportion of patients with chronic HF (EF >40%) experienced a reduction in daily activities compared to those without HF (52.2% *versus* 36.8%) (61).

The impact of HF on patients' emotional well-being is substantial across the EF spectrum. Patients have often reported feeling overwhelmed, frustrated, limited and worried, particularly around the caring for their children/spouses and the impact it has on their self-confidence (42). The substantial reduction in patient's physical and emotional well-being are even associated with a higher risk of mortality (61, 96).

Similarly, the impact of HF on the quality of life of carers is also significant. Carer's health as a result of carer's responsibilities were also shown to be significantly impacted by stress (35%), moderate to severe anxiety/depression (32%), emotional strain (33%), physical (33%) or mental (31%) tiredness and pain/discomfort (29%) (95).

B.1.3.2 Clinical pathway of care

Although several therapies are recommended for the treatment of patients with chronic HF (EF ≤ 40%) (3, 20), there remains an unmet need for an effective treatment for patients with HF (EF >40%) that improves disease-related outcomes.

- Diuretics, calcium-channel blockers, amiodarone (in consultation with a specialist) and anticoagulants are recommended for the management of all patients with chronic HF, (i.e EF <40%, >40%)
- The NICE guideline for chronic HF in adults (NG106) does not recommend any targeted pharmaceutical treatment for chronic HF (LVEF >40%). (3).
- For patients with chronic HF (EF >40%), treatment is focused on the management of comorbidities such as hypertension, atrial fibrillation, IHD and diabetes in line with NICE guidance (3).

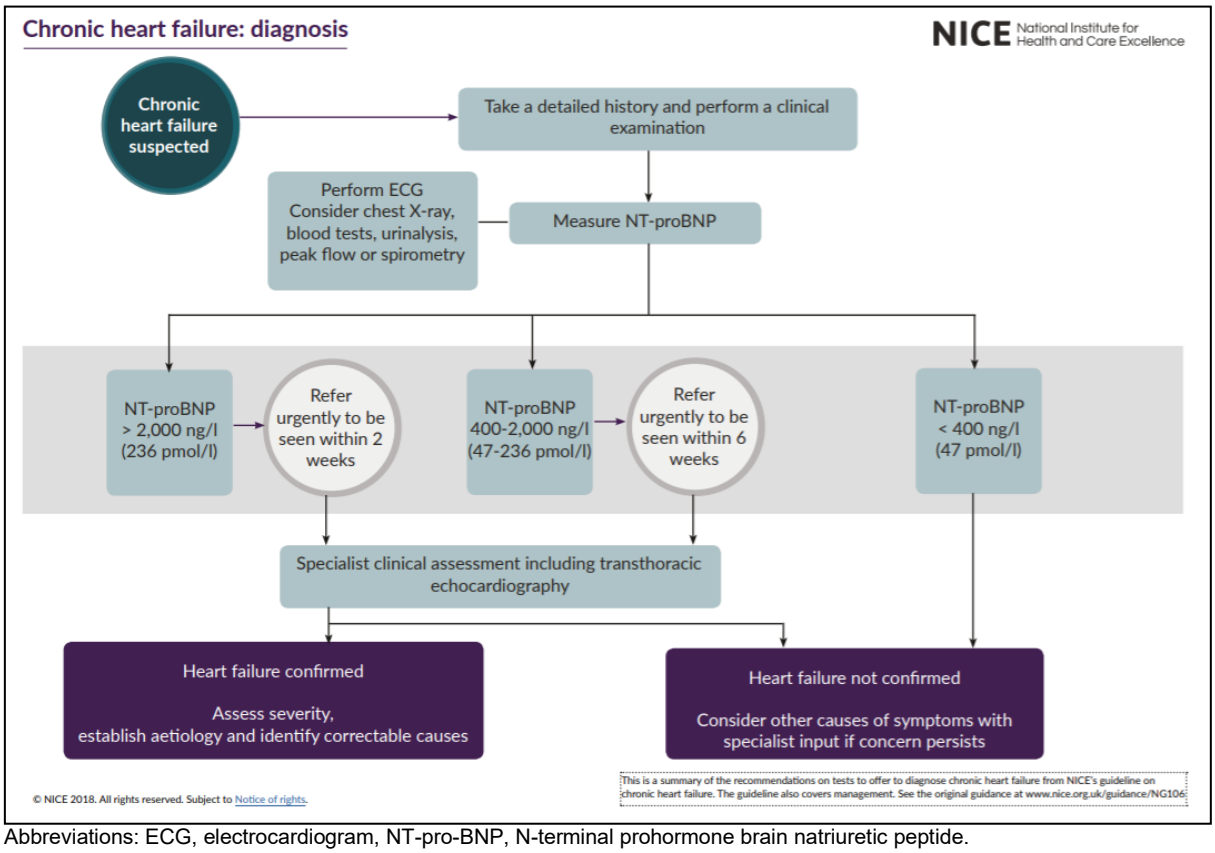
- The 2021 ESC guideline does not recommend any targeted pharmaceutical treatment for chronic HF (LVEF >50%). Additionally, although angiotensin-converting enzyme inhibitors (ACEI), angiotensin receptor blockers (ARBs), angiotensin receptor-neprilysin inhibitors (ARNI), beta-blockers and mineralocorticoid receptor antagonists (MRA) have been recommended to be used in patients with chronic HF (EF 41% to 49%), the strength of recommendations are low and are not well established by evidence/expert opinion (20).
- In the clinical practice, the implementation of NG106 for HF patients is highly variable (3, 47). This is observed more acutely in those patients with a higher EF, due to challenges in diagnosis (39, 97).

B.1.3.2.1 Current standard of care

NICE clinical treatment pathway

The diagnosis of HF is multifactorial and encompasses detailed clinical history, physical examinations, electrocardiograms (ECGs), stress tests, chest x-rays, coronary angiograms, cardiac computerised tomography (CT) scans, magnetic resonance imaging (MRI), myocardial biopsies and laboratory tests. Given the uncertainties that are intrinsic to a clear diagnosis of HF on physical examination alone, and the outcome for patients left undiagnosed, the NICE and ESC guidelines recommend testing of serum N-terminal pro-B-type natriuretic peptide (NT-pro-BNP) in people with suspected HF as an essential diagnostic tool (Figure 2) (3, 20). However, the NT-pro-BNP level cannot differentiate between chronic HF EF ≤40% and EF >40% (3). Transthoracic echocardiography is required for confirmatory diagnosis and to inform classification of HF, which in turn guides the management of the condition (3, 20).

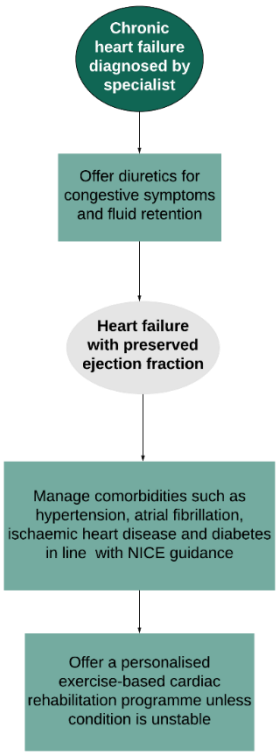
Figure 2. Chronic heart failure diagnostic pathway



Abbreviations: ECG, electrocardiogram, NT-pro-BNP, N-terminal prohormone brain natriuretic peptide.
Source: NICE guideline NG106 (3)

Following chronic HF (EF >40%) diagnosis, the treatment focuses on the management of comorbidities and to alleviate symptoms and improve well-being (74, 98). Currently, the NICE guidelines do not recommend any specific therapy for the treatment of chronic HF (EF >40%) (Figure 3) (3).

Figure 3. NICE treatment pathway for chronic HF (EF >40%)



Abbreviation: HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction.
Note: As per NICE guideline, all patients with EF <40% are classified as HFrEF and remaining other HF patients are classified as HFpEF.
Source: Adapted from NICE guideline NG106, 2018 (3)

Calcium-channel blockers, amiodarone (in consultation with a specialist), anticoagulants and diuretics are used for the management of all patients with chronic HF (3). Diuretics are used routinely to provide symptomatic relief, particularly in the presence of oedema, but without direct evidence of survival benefit (99). Additionally, the efficacy benefit of diuretics across the EF spectrum of HF is not equal (3, 99). Patients with chronic HF (EF >40%) are usually offered a low to medium dose of loop diuretics such as furosemide (<80 mg per day) (3). Patients who do not respond to diuretics are then referred to a specialist who can optimise comorbidity management and can advise patients to use other services, including cardiac rehabilitation, services for older people and palliative care services, as needed (3).

Clinical practice and heart failure services

The recent 2021 ESC guideline has recommended the use of ACEI, ARB, ARNI, beta-blockers, MRA and diuretics for management of patients with chronic HF (EF 41% to 49%) (Table 8) (20). However, the strength of the recommendations are low and are not supported by evidence as no substantial prospective randomised controlled trials

(RCTs) have been exclusively conducted in patients with EF 41% to 49% (20). For patients with chronic HF (EF ≥50%), the guideline still does not recommend any specific medications (20). The management of these patients are limited to screening and treatment of CV and non-CV comorbidities (Table 8) (20). In clinical practice, the prescription of pharmacological treatments in patients with chronic HF (EF >40%) was similar to that with EF ≤40% since there are no evidence-based guidelines for these patients (100, 101). This finding may reflect the use of concomitant treatment for comorbidities such as loop diuretics to manage fluid overload and congestion, and ACEI/ARBs for hypertension, which are presented in chronic HF patients irrespective of EF (20).

Table 8. ESC recommendations or pharmacological treatments to be considered in patients with chronic HF

EF ≤40%	EF 41%-49%	EF ≥50%
All patients: • An ACEI • A beta-blocker • An MRA • Dapagliflozin or empagliflozin • ARNI **Selected patients:** <u>In patients with congestion</u> • Diuretics <u>In patients who failed on an ACEI (or ARNI), a beta-blocker and an MRA</u> • An ARB • Ivabradine • Vericiguat • Hydralazine and isosorbide dinitrate • Digoxin	• Diuretics • An ACEI • A beta-blocker • An MRA • ARNI	• Screening for, and treatment of, aetiologies, and CV and non-CV comorbidities • Diuretics

Abbreviations: ACEIs, angiotensin-converting enzyme inhibitors; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; CV, cardiovascular; ESC, European Society of Cardiology; HF, heart failure; EF, left ventricular ejection fraction; MRAs, mineralocorticoid receptor antagonists.
Reference: (20).

Many challenges are faced by GPs in managing patients with chronic HF in primary care. GPs reported limited understanding of different types of HF, mostly due to a lack of clear consensus on its definition and diagnosis (97). Patients with chronic HF,

regardless of their EF, present in clinical practice in the same way (102). Both associated CV and non-CV comorbidities make the diagnosis very complex (102). Primary care often test for elevated NT-proBNP, and if HF is suspected, the patient is referred to specialist care where the diagnosis and type of HF is confirmed (39). However, echocardiograms are unreliable for some of the arbitrary definitions for HF subtypes (e.g., distinguishing between an EF of 40% to 45%) and do not often report an exact value for EF >55% in clinical practice as it would be classed as “normal EF” (39). There are limitations in access to echocardiography and technically trained staff as well. Variation has been observed in access to natriuretic peptide testing for diagnosis and monitoring and use of validated tools to quantify the severity of symptoms (47, 70). Combined, these factors lead to delays in diagnosis and subsequently, in the commencement of treatment.

There are inconsistencies between the guidelines and clinical practice in HF service settings (e.g. hospital-based, community-based, hospital- and community-based or hospital with community work) (3, 47). Optimal management of chronic HF requires optimisation of pharmacological treatment, nursing support and treatment of comorbidities, and should be delivered by a multidisciplinary team (MDT) (3, 47). In UK clinical practice, however, the structure and provision of HF care varies and is not always in accordance with the current guidelines (3, 47). There is a lack of availability of specialist services for patients with chronic HF with higher EFs (103). Around 60% to 80% of specialist HF practices reported patients with EF >50% and only 53% of community services reported these patients of EF >50% (47, 103). Thus, the patients with EF >50% are usually discharged to primary care after diagnosis, who then take the lead in managing these patients (70, 104).

The COVID-19 pandemic has further negatively impacted the availability of HF services, including diagnostic and outpatient specialist care services. Patients are now also less likely to seek medical care for any HF symptoms they experience (Section B.1.4) (15). A UK-based study reported that in primary and secondary care, inpatient ECGs reduced by 44% and NT-proBNP tests reduced by 75% during the COVID-19 lockdown period (105). This led to a reduction in patients presenting with signs and symptoms of HF and a 34% decrease in the number of new patients referred to

community HF service (105). Many HF specialist clinicians had to be reallocated to acute or medical wards in order to accommodate COVID-19 patients (105-107). Home visits and in-clinic appointments were postponed for around 65% of HF patients, and telephone or video consultation services increased by 66% (105, 108).

COVID-19 has exacerbated pre-existing health inequalities, as patients in a lower socio-economic group were already less likely to seek medical attention in secondary care before the pandemic. The impact is significant. HF is also a risk factor for worse outcomes with COVID-19 (109, 110), and patients with chronic HF were 17% more likely to die of COVID-19 than those who did not (111). Additionally, HF patients with a lower socio-economic status were already more likely to have worse CV outcomes than those with a higher socio-economic status (9).

B.1.3.2.2 Unmet need

- Patients with chronic HF continue to experience high mortality and morbidity, high symptom burden, reduced functional capacity and poor quality of life (9, 20, 47, 61, 82).
- No therapy so far has demonstrated efficacy for HF across the broad spectrum of EF (86).
- There is no evidence that the currently prescribed drugs for HF across the broad spectrum of EF (e.g. ACEI/ARB, beta-blockers, or MRAs), which are generally used to control CV comorbidities, significantly reduce symptom burden (20).
- As there are no licensed efficacious SoC treatments for patients with HF (EF >40%), the usual therapies (i.e., treatments used to treat CV comorbidities) are the current SoC for these patients (20).
- There is a need for targeted therapies that are available across primary and secondary care that have an immediate impact on patient prognosis and quality of life without dose-limiting side effects. This will be important for post–COVID-19 recovery.

As discussed in earlier sections, the mortality and morbidity of HF remain high due to a number of factors, including late diagnosis that most often occurs after emergency admission for acute HF symptoms, lack of pharmacological therapy and widening socio-economic inequalities (9, 47, 82). Further, with each subsequent hospitalisation, the risk of an unplanned death increases, reducing the opportunity for a patient to choose how they want to die (112, 113). During 2018-2019, there were more than

100,000 hospital admissions for HF in the UK, an increase of almost a third compared to 2013-2014 (114, 115). The challenges in diagnosis and limitation of available treatment for its management contribute to the increased risk of HF hospitalisations and mortality (20, 42, 70). The quality of life of patients is also markedly reduced especially in patients with chronic HF (EF >40%) since there are no licensed efficacious SoC treatments targeting this population (20, 61).

Many different classes of therapies such as ACEI, ARBs, beta-blockers, MRA and ARNI have been tested in international (including the UK) clinical trials but did not show a statistically significant improvement in patients with HF across a broad spectrum of EF. ACEI (perindopril), in the PEP-CHF trial, did not show any statistically significant differences in outcomes compared with placebo in all-cause mortality or CV hospitalisation that included elderly patients with HF (EF 40%-50%) due to left ventricular diastolic dysfunction; thus, it is non-efficacious across a broad spectrum of EF (116). Similarly, angiotensin receptor blocker (irbesartan) failed to show benefit across a broad spectrum of EF. As assessed in the I-PRESERVE trial including patients with EF >45%, irbesartan did not improve the primary composite outcome of death from any cause or hospitalisation for a CV cause (HF, MI, unstable angina, arrhythmia or stroke) (35). Angiotensin receptor-neprilysin inhibitor (sacubitril/valsartan) assessed in the PARAGON-HF trial, including patients aged >50 years and EF >45%, was the largest clinical trial to date (N=4,822). After a median follow-up of 35 months, sacubitril/valsartan did not show a statistically significant reduction in CV death and HHF among patients with HF and EF >45% (37).

Global studies of other drugs also had similar results. Another ARB, candesartan in the CHARM-Preserved trial, showed no difference in CV deaths but a non-significant reduction in HHF compared to placebo (HR, 0.84; 95% CI, 0.70 to 1.00) (38). Nebivolol, a beta-blocker, significantly reduced the combined primary endpoint of all-cause mortality or CV hospitalisation in the SENIORS trial; however, this trial included very few (15%) patients with an EF >50% (117, 118). Furthermore, this trial had a high likelihood for type II errors which reduced its power to detect the reported differences among the EF subgroups (118). A meta-analysis including 15 non-interventional and two interventional studies found that beta-blockers significantly reduced HHF but not

mortality in patients with HF (EF >40%) (119). The primary composite endpoint (death from CV causes, aborted cardiac arrest or HHF) for the MRA (spironolactone) studied in the TOPCAT trial was not met (HR, 0.89; 95% CI, 0.77-1.04); however, it showed a significantly lower incidence for HHF (HR, 0.83; 95% CI, 0.69-0.99; p=0.04) (36). Thus, to date, there are no approved disease-modifying treatments for management of HF patients across a broad spectrum of EF. Furthermore, in clinical practice, the management of patients with chronic HF (EF >40%) focuses on identification and treatment of the underlying risk factors, aetiology and coexisting comorbidities; however, improved outcomes have not been observed for these patients (20).

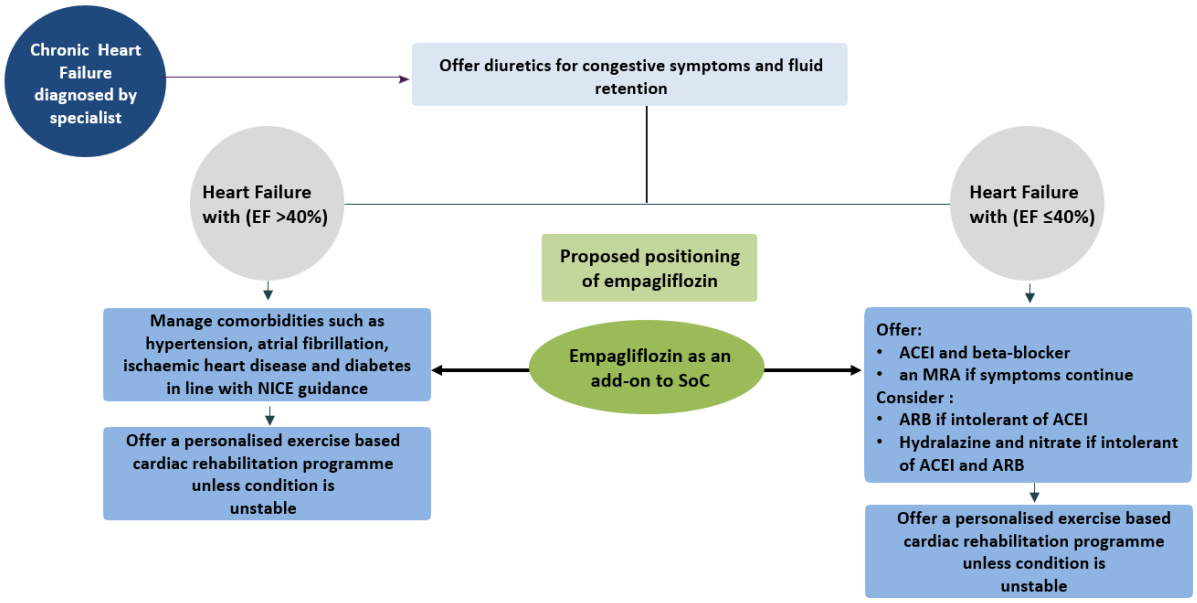
There is a need for evidence-based disease-modifying therapies for patients with chronic HF across a broad spectrum of EF that are available across primary and secondary care and not reliant on echocardiography to establish EF. This will have an immediate impact on patient prognosis and quality of life which will be important for post-COVID-19 recovery.

B.1.3.2.3 Positioning of empagliflozin in the UK treatment pathway

- Empagliflozin is the first therapy to demonstrate efficacy in a broad range of chronic HF patients across the full spectrum of EF (5, 6).
- Empagliflozin should be broadly positioned as an add-on to background therapy regardless of LVEF in all patients with chronic symptomatic heart failure.
- Empagliflozin is already recommended as an add-on in patients with an EF≤40%, based on EMPEROR-Reduced (TA733).
- Based on EMPEROR-Preserved, the optimal positioning in patients with an EF>40% is as an add-on to background therapy for the management of co-morbidities and symptomatic relief (5).
- With broad prescribing of empagliflozin across primary and secondary care in chronic HF regardless of EF, there is an opportunity to maximise outcomes for these patients immediately; a key objective of the NHS Long Term Plan
- Broad prescribing of empagliflozin across primary and secondary care reduces the risk of widening the gap in health inequalities seen in HF patients as a result of COVID-19 (Section B.1.4).

Based on the population studied in the pivotal phase III study EMPEROR-Preserved, the optimal positioning for empagliflozin in the NICE pathway is as an add-on to background therapy for symptomatic relief and comorbidity management for chronic HF (EF >40%) patients. Throughout this company submission, background therapy is referred to as standard of care (SoC). As it will be the first approved treatment for this population, offering it as an add-on will reduce CV death or HHF, decrease the rate of renal function decline and improve quality of life (5).

Figure 4. Proposed positioning of empagliflozin in NICE treatment pathway for chronic HF



Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blockers; EF, ejection fraction; MRA, mineralocorticoid receptor antagonist; NICE, National Institute for Health and Care Excellence; SoC; standard of care.
The preferred positioning for empagliflozin (green box) is in diagnosed symptomatic patients with chronic HF.
Source: Adapted from NICE guideline NG106, 2018 (3)

Since there are currently no efficacious therapies approved for use in patients with chronic HF across the broad spectrum of EF, the following drugs recommended for all HF patients are considered relevant comparators for empagliflozin in this positioning:

- Diuretics – for relief of congestive symptoms and fluid retention
- ACEI/ARB
- Beta-blockers
- MRAs
- ARNI (sacubitril/valsartan)

This is also consistent with the perspective of UK-based clinical advisers BI have consulted (39), and the results from a CPRD study, PULSE, which looked at the prevalence, incidence, and expected outcomes for patients in the UK with HF (48). Both reported that these treatments are frequently used to treat HF patients with an EF>40%.

Table 9. Background treatments for HF in EMPEROR-Preserved vs PULSE (a CPRD study)

Treatment Arm	EMPEROR-Preserved (ITT; Combined Groups)	PULSE (unknown group)	PULSE (EF>40%)
N	9718 (100.0)	[redacted]	[redacted]
HF medication [(N), %]			
ACEi/ARB	7305 (75.2)	[redacted]	[redacted]
Beta blocker	8700 (89.5)	[redacted]	[redacted]
Diuretic	8708 (89.6)	[redacted]	[redacted]
MRA	4905 (50.5)	[redacted]	[redacted]
Sacubitril/valsartan	861 (8.9)	[redacted]	[redacted]
Ivabradine	331 (3.4)	[redacted]	[redacted]
Digoxin	NR	[redacted]	[redacted]
Hydralazine/nitrate	282 (2.9)	[redacted]	[redacted]

Footnote: a. The baseline characteristics are for the EF>40% and unknown groups at a 2019 prevalent cross section. The unknown group included patients where the EF was not recorded.
Abbreviations: NR, not reported; ARB, angiotensin receptor blocker; ACEi; angiotensin converting enzyme inhibitors
Source: PULSE Study Report (48), CTR, Table 1.3.15

In a combined HF analysis stratified by EF, empagliflozin reduced the risk of CV death or HHF, mainly by reducing HF hospitalisations in chronic HF patients (120). The magnitude of the effect of empagliflozin on HF outcomes was similar in all HF patients irrespective of EF (120). The EMPEROR-Preserved trial showed that empagliflozin has the potential to provide additional efficacy in combination with any given background therapy in patients with chronic HF (EF >40%) (Section B.2) (5, 121). The composite primary outcome in EMPEROR-Preserved study showed that 13.8% of patients receiving empagliflozin plus SoC *versus* 17.1% receiving SoC alone experienced either a HHF or CV death event (HR 0.79; 95% CI, 0.69 to 0.90; p=0.0003). The decline in renal function, evaluated based on change in eGFR slope from baseline, was significantly slower in the empagliflozin group, with an estimated difference in slope of about 1.36 mL/min/1.73 m^2 per year *versus* placebo (95% CI, 1.06 to 1.66; p<0.0001). The results from the adverse events (AE) and safety

laboratory analyses in the EMPEROR-Preserved study were similar to the known safety profile of empagliflozin and no new occurrences were identified. Empagliflozin has an established safety profile and does not show any relevant drug-drug interactions (5). Further benefits of empagliflozin are that it is a once-daily dose without the need of any dose titration and thus, no additional clinical time is needed to optimise a patient's treatment (5); it does not have any effect on patient's potassium levels and blood pressure; and it does not need additional renal monitoring compared to the usual therapy (5). Empagliflozin has demonstrated improvement in HF-related outcomes across a broad range of chronic HF (EF >40%) populations including the presence or absence of T2DM and/or CKD, and baseline health status as measured by KCCQ (5). Empagliflozin has shown similar efficacy results among patients with chronic HF (EF ≤40%) in EMPEROR-Reduced trial which has been summarised in a previous NICE submission TA773 (1).

A NICE recommendation for empagliflozin in chronic HF across a broad range of EF will have a positive impact on the existing pathway. The experience of GPs in prescribing SGLT2i for T2DM and there being no requirement for dose adjustment should facilitate initiation of empagliflozin in chronic HF patients within primary care (122). Currently, the HF patients are treated based on their EF and diagnosed using echocardiograms (20). It has been seen that echocardiograms can be unreliable and thus, depending on the result can make some patients not be referred or receive the guideline-directed treatment (39, 42). With empagliflozin, all patients diagnosed with HF have the opportunity to receive an evidence-based targeted treatment, regardless of EF. Additionally, prescribing empagliflozin does not require any additional monitoring beyond usual care and what is expected for other drugs impacting renal function (19).

B.1.4 Equality considerations

Socio-economic inequalities in CV disease present a major and persistent UK public health challenge. The UK-based population studies demonstrate that socio-economic deprivation is a strong risk factor for the development of HF and adverse HF outcomes (9, 10). Individuals in the lowest socio-economic group are 1.61 times more likely to experience incident HF than the most affluent individuals and on an average, at a 3.5

years younger age with a greater comorbidity burden at time of HF symptom onset (9). Furthermore, the socio-economic status is associated with a diverging trend of HF outcomes in England, whereby patients from the most deprived group have a significantly higher risk of all-cause (HR, 1.17; 95% CI,1.14 to 1.21) and CV mortality (HR, 1.18; 95% CI, 1.14 to 1.23) than the most affluent ones (114).

Since the early 2000s, the socio-economic gradient in HF incidence and outcomes have been widening:

- The mean age at diagnosis increased by 2.45 years (95% CI, 1.58 to 3.32) among the most affluent but tended to decrease among the most deprived (9).
- The annual risk in HHF has increased by 1.6% (95% CI, 0.6 to 2.6) for the most deprived compared to a stable risk for the most affluent group (114).

The inequality in access to specialist care in the UK may be one of the drivers of the observed trends in HF. England-based socio-economic studies have shown that after controlling for need, richer individuals tend to consume more public and private specialist visits, but not family physician visits, than those from a lower socio-economic class, and experience significantly shorter waiting times for a coronary revascularisation procedure at the same public hospital (11, 123). The prominent role of a secondary care specialist in all aspects of HF care (including diagnosis, management and initiation of new medicines) that can only be accessed upon referral from a GP, could therefore be contributing to the observed socio-economic disparities in clinical characteristics and outcomes of HF (3).

The choice of setting for empagliflozin initiation in primary care or under specialist supervision is thus a highly pertinent public health issue. Broad prescribing of empagliflozin across primary and secondary care can support the reduction in disparity in terms of access to HF care across socio-economic groups within the UK, given that empagliflozin significantly improves CV and slows renal decline in chronic HF (EF >40%) patients in an early, sustained manner and prevents HHF compared to SoC (5). Empagliflozin significantly reduced worsening of an HF event (CV death, HHF or an emergency or urgent HF visit requiring intravenous treatment) 18 days after randomisation and maintained significance thereafter (124). It is the first treatment that can simultaneously provide cardiac and renal benefits to chronic HF patients across

broad range of EF with comorbid diabetes and/or severe renal impairment (eGFR 20 to 30 mL/min/1.73 m^2), which are more likely to coexist in the most deprived patients (10). Additionally, patients treated with empagliflozin do not show any relevant drug-drug interactions, have a once-daily dose and do not need additional renal monitoring compared to the usual therapy (5). Limiting initiation of empagliflozin to secondary care specialists could lead to a delayed and/or lower uptake of empagliflozin among the most socioeconomically disadvantaged groups as they consume fewer specialist visits and present to healthcare providers at a later stage of illness (123). Delayed exposure to the benefits of SGLT2 inhibition may in turn widen the existing divide in HF outcomes between socio-economic classes in England.

The COVID-19 pandemic has further reinforced this trend through significant disruption in the provision of all types of cardiology services including outpatient and community HF services (105). Patterns of past care suggest that the elderly and those living in deprived areas are most likely to be disproportionately affected by increased waiting times for cardiology appointments (11, 125). With a condition that has a 1-year mortality of approximately 24% and being the leading cause of hospital readmissions, a long wait for a HF specialist appointment may have grave consequences for socioeconomically disadvantaged HF patients in England (63).

In this company submission, equity of access to empagliflozin among chronic HF patients from all socio-demographic groups is an important consideration. A recommendation by NICE that facilitates broad prescribing of empagliflozin across primary and secondary care and its classification as “green” on local/regional formularies would support this objective. This in turn will support the overarching goal of reducing inequity in access to care for HF patients, in line with NICE’s Social Value Judgments, pillar 3 of NICE’s new 5-year strategy (13, 14) and the conclusions from the Marmot COVID-19 Build Back Fairer review (15).

B.2 Clinical effectiveness

- Empagliflozin is efficacious in chronic HF patients across a broad spectrum of EF, (i.e., EF ≤40% and EF >40% in a combined HF trial analysis) stratified by EF (120).
- The results of the EMPEROR-Reduced trial are described in the previous NICE submission TA773 corresponding to an appraisal of empagliflozin for treating chronic HF patients with EF ≤40% (6).
- This company submission presents the clinical effectiveness of empagliflozin in the EMPEROR-Preserved trial conducted among patients with EF >40%, which completes the evidence package demonstrating the benefits of empagliflozin across a broad spectrum of EF for chronic HF patients (5).
- The EMPEROR-Preserved trial was similar in study design to EMPEROR-Reduced as an event-driven, double-blind RCT which enrolled 5,988 patients with moderate to severe chronic HF (EF >40%, NYHA II-IV) randomly assigned to receive empagliflozin (N=2,997) or placebo (N=2,991).
- After a median follow-up of 26.2 months, empagliflozin significantly reduced the risk of death from CV causes or HHF compared to placebo (HR, 0.79; 95% CI, 0.69-0.90; p=0.0003).
- Empagliflozin was superior to placebo with respect to key secondary endpoints:
  - It led to a significant reduction in the total number of adjudicated HHF (first and recurrent) (HR, 0.73; 95.03% CI, 0.61 to 0.88; p=0.0009) vs placebo.
  - The rate of the decline in eGFR was slower in empagliflozin group compared to placebo group over the duration of the double-blind treatment period (between-group difference, 1.36 mL/min/1.73 m^2 per year; 95% CI, 1.06 to 1.66; p<0.0001).
- Empagliflozin was also superior to placebo in other secondary endpoints:
  - It reduced risk of adjudicated HHF (first event) (HR, 0.71; 95% CI, 0.60 to 0.83; nominal p<0.0001).
  - It led to improvement in the HRQoL score on KCCQ at 52 weeks (placebo-corrected adjusted mean change from baseline 1.32; 95% CI, 0.45 to 2.19; nominal p=0.0028).
- Empagliflozin appears to reduce the CV mortality (HR, 0.91; 95% CI, 0.76 to 1.09), all-cause hospitalisation (HR, 0.93; 95% CI, 0.85 to 1.01) and composite renal endpoint (HR, 0.95; 95% CI, 0.73 to 1.24); however, these results need to be interpreted with caution as were not statistically significant.

- The CV and renal benefits of empagliflozin were consistent across subgroups of chronic HF patients (EF >40%) defined by demographics, baseline characteristics, and baseline medications.
- In chronic HF patients (EF >40%) with diabetes, there was a greater mean reduction in glycated haemoglobin (HbA1c) from baseline in the empagliflozin group than in the placebo group.
- Empagliflozin improves CV and renal outcomes of chronic HF patients (EF >40%) including those with an eGFR down to 20 mL/min/1.73 m^2.

B.2.1 Identification and selection of relevant studies

A systematic literature review (SLR) was conducted to identify from the published literature RCT evidence on the efficacy and safety of empagliflozin and relevant comparators in patients with chronic HF (EF >40%, NYHA II-IV). Full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised are described in Appendix D.

Searches of Medline® and Embase® (via Embase.com), Medline in-Process® (via PubMed.com) and the Cochrane Library were performed on 14 May 2020, and subsequently updated on 8 October 2020, 8 July 2021, 16 February 2022 and 07 July 2022. The search of electronic databases was supplemented with a desk search of conference proceedings, last conducted on 7 July 2022.

The eligible studies encompassed all RCTs evaluating efficacy of pharmacological interventions used in the treatment of adults (age ≥18 years) with chronic HF (EF >40%). The search strategy was designed to be broad and to encompass all interventions that are generally used for the management of chronic HF (eligibility criteria are shown in Table 15 in Appendix D). All studies meeting the pre-specified population, intervention, comparator, and outcomes (PICOS) eligibility criteria were retained and were extracted.

The EMPEROR-Preserved trial compared empagliflozin with usual therapy, and is therefore the primary source of clinical evidence in the economic model for the chronic HF patient population with EF >40% (5). A combined HF analysis stratified by EF was

also identified from the SLR as relevant to the decision problem for empagliflozin for chronic HF with EF >40% and is therefore described in this submission (120). The analysis provided by Butler et al. (2022) demonstrates important differences in demographic characteristics across EF intervals, with older women more likely to present with higher EF levels (120). The results further show consistency in benefits of empagliflozin across a broad spectrum of EF, which is irrespective of demographic characteristics (120). More details regarding this analysis are provided in Section **Error! Reference source not found.** below. A full list of studies that were included and excluded during the SLR is provided in Appendix D.

B.2.1.1 Clinical trials with empagliflozin 10 mg (Jardiance®)

Empagliflozin is being investigated in the EMPOWER clinical trial programme. The most comprehensive development programme for an SGLT2i to date, EMPOWER, is comprised of nine clinical trials and a RWE study that have been designed to evaluate the impact of empagliflozin on CV and renal outcomes of patients across the spectrum of CRM disorders (Table 10). Furthermore, the aim of the programme is to advance the scientific understanding of the pathophysiology of cardio-renal interactions and enable a holistic management of the interconnected CRM organ system.

Table 10. Overview of the studies comprising the EMPOWER clinical trial programme for empagliflozin

Study name	Study identifier	Main objective	Status	Relevant for this appraisal & reason
EMPEROR-Preserved	NCT03057951 (126)	Efficacy & safety of empagliflozin in prevention of CV death and HHF in adults with chronic HF patients (EF >40%) with or without T2DM	Completed	Yes, meets the PICO criteria as defined in the decision problem

Study name	Study identifier	Main objective	Status	Relevant for this appraisal & reason
EMPEROR-Reduced	NCT03057977 (127)	Efficacy & safety of empagliflozin in prevention of CV death and HHF in adults with chronic HFrEF with or without T2DM	Completed	No; population is not relevant for the decision problem
EMPERIAL-Preserved	NCT03448406 (128)	Effect of empagliflozin on functional ability and PROs in adults with chronic HFpEF with or without T2DM	Completed	No; primary outcome is not relevant for the decision problem; QoL secondary endpoint measured using PROs is not recommended by the NICE reference case (93)
EMPERIAL-Reduced	NCT03448419 (129)	Effect of empagliflozin on functional ability and PROs in adults with chronic HFrEF with or without T2DM	Completed	No; population is not relevant for the decision problem
EMPA-REG OUTCOME	NCT01131676 (130)	Efficacy & safety of empagliflozin in prevention of major adverse CV events, including CV death, in adults with T2DM and established CV disease	Completed	No; population is not relevant for the decision problem
EMPULSE	NCT04157751 (131)	Efficacy of empagliflozin in improving clinical and PRO outcomes in adults hospitalised for acute HF	Completed	No; population is not relevant for the decision problem

Study name	Study identifier	Main objective	Status	Relevant for this appraisal & reason
EMPA-KIDNEY	NCT03594110 (132)	Effect of empagliflozin on progression of kidney disease and the occurrence of CV death in patients with pre-existing CKD	Ongoing	No; population is not relevant for the decision problem
EMPA-VISION	NCT03332212 (133)	Effects on cardiac physiology and metabolism in patients with HF	Completed	No; the study outcomes are not relevant for the decision problem
EMPACT-MI	NCT04509674 (134)	Efficacy of empagliflozin in improving outcomes and preventing HF in adults hospitalised with an acute MI	Ongoing	No; population is not relevant for the decision problem
EMPRISE	NCT03363464 (135) EUPAS20677 (136)	Real-world comparative effectiveness, safety, healthcare resource utilisation and costs of empagliflozin *versus* DPP-4 inhibitors in T2DM in routine clinical care	Ongoing	No; population is not relevant for the decision problem

Abbreviations: CKD, chronic kidney disease; CV, cardiovascular; DPP-4, dipeptidyl peptidase-4; HF, heart failure; HFrEF, heart failure with reduced ejection fraction; HFpEF, heart failure with preserved ejection fraction; HHF, hospitalisation for heart failure; MI, myocardial infarction; PICO, patient intervention comparator outcome; PROs, patient-reported outcomes; T2DM, type 2 diabetes mellitus.

Of the studies listed in Table 10, the EMPEROR-Preserved trial provides the main evidence base for clinical efficacy and safety of empagliflozin in the population of HF patients with EF >40%. In the trial, randomisation was stratified by EF (<50%, ≥50%) and >66% of enrolled patients had EF ≥50% (5). It should be noted that the clinical

effectiveness of empagliflozin for the treatment of patients with EF ≤40% as studied by the EMPEROR-Reduced trial was appraised by NICE in appraisal TA773 (1, 6).

The EMPEROR-Preserved (NCT03057951) trial was an international phase III trial from the EMPOWER programme that investigated the effect of empagliflozin versus placebo, in addition to usual therapy on the combined risk of CV death and HHF in 5,988 patients with chronic HF (EF >40%), with or without diabetes. It also evaluated the effects of empagliflozin on recurrent hospitalisation events, renal function, CV death, all-cause mortality, and change in the Kansas City Cardiomyopathy Questionnaire - clinical summary score (KCCQ-CSS) (126). The rationale for the design of the EMPEROR trials; EMPEROR-Preserved and EMPEROR-Reduced, was that the data from the EMPA-REG OUTCOME trial was not sufficient to demonstrate efficacy of empagliflozin in patients with chronic HF, especially those at increased risk of an outcome event. The EMPEROR-Preserved trial enrolled patients from eight UK sites, increasing its relevance to the NHS clinical practice. External validity of the trial is strengthened by the protocol requirement for patients to receive usual therapy for chronic HF (EF >40%). The generalisability of the trial results to the NHS clinical practice is discussed further in Sections B.2.7, B.2.13 and B.3.2. Its outcomes provide the key clinical and quality of life inputs for the economic model of empagliflozin in chronic HF (EF >40%). The design and methodology of the EMPEROR-Preserved trial is similar to the EMPEROR-Reduced trial, and this is described in Section B.2.3.1.

B.2.1.2 Non-randomised clinical effectiveness studies

Evidence from PULSE, a large retrospective observational study of the burden of chronic HF, in England, was used to characterise patients seen in the NHS clinical practice and validate the long-term outcome predictions of the chronic HF (EF >40%) cost-utility model for patients treated with the SoC against the real-world outcomes (48). Patients with a diagnosis of HF recorded in the UK CPRD or Hospital Episode Statistics (HES) database between 1 January 2015 and 31 December 2019 were eligible for inclusion in the PULSE study (48). Based on the availability of evidence of EF classification in CPRD records, the cohort was split into EF measure ≤40%, EF measure >40% and “unknown ejection fraction” subpopulations. The study objectives were to determine the incidence and prevalence of HF and those with EF >40% in

England, estimate rates of HF- and those with EF >40%- related hospitalisation, CV and all-cause mortality, and evaluate resource utilisation over the study period. One of the limitations of using the PULSE study data was that the majority of patients were not coded as per EF subtype at or prior to their HF index date (48). Therefore, other non-randomised clinical effectiveness studies were explored; however, RWE for the population with HF and EF >40% are limited. A study by Oo et al. (2021) was considered and the authors conducted a single-centre study to determine clinical characteristics and outcomes of patients defined as preserved EF according to recent guidelines and outcome trials (137). Another study by Uijl et al. (2021) was considered which included patients with EF ≥50% and clustered them according to their clinical characteristics (138). However, these two studies (Oo et al. [2021] and Uijl et al. [2021]) were not considered relevant for this decision problem to appropriately support the evidence package because of missing data and lack of transparency with endpoints (137, 138). Due to lack of available published evidence, and the inappropriateness of the alternative studies considered, the outcomes of the PULSE study are therefore considered the best available evidence to support this decision problem (Section B.3.2) (48).

B.2.2 List of relevant clinical effectiveness evidence

The clinical evidence on empagliflozin as an addition to SoC (usual therapy i.e., treatments used to treat CV comorbidities) in the treatment of chronic HF (EF >40%) consists of one phase III trial, EMPEROR-Preserved (Table 11). This pivotal trial was the main source of clinical efficacy evidence in the cost-utility model described in Section B.3.

Table 11. Clinical effectiveness evidence: EMPEROR-Preserved trial

Study	EMPEROR-Preserved (NCT03057951) (126)
Primary sources	Anker et al 2021 (5)
Additional sources	EMPEROR-Preserved CSR (139)
Study design	• Phase III, randomised, double-blind, placebo-controlled trial with parallel assignment

Study	EMPEROR-Preserved (NCT03057951) (126)
	• The trial was event-driven and all randomised patients remained in the trial until the defined number of adjudicated primary endpoint events had been reached
Population	Adults with chronic HF NYHA class II-IV and EF >40% who have been diagnosed at least 3 months before screening, with or without DM • N=5,988 • Age ≥18 years • Baseline natriuretic peptide levels >300 pg/mL for patients without atrial fibrillation or atrial flutter (AF); >900 pg/mL for patients with AF at screening (see Section B.2.3.1.2) • Oral diuretics, if prescribed to patient according to local guidelines and discretion of the investigator, should have been stable for at least 1 week prior to (randomisation)
Intervention(s)	Empagliflozin PO 10 mg once daily in addition to SoC (usual therapy i.e., treatments used to treat CV comorbidities)
Comparator(s)	Placebo plus SoC
Does trial support application for marketing authorisation?	Yes
Is trial used in the economic model?	Yes
Reported outcomes specified in the decision problem	The outcomes relevant for the decision problem include: • Time to first event of adjudicated CV death or adjudicated HHF • Occurrence of adjudicated HHF (first and recurrent) • Decline in renal function • Time to first occurrence of chronic dialysis, renal transplant or sustained reduction of eGFR • Time to first adjudicated HHF • Time to adjudicated CV death • Time to all-cause mortality • Occurrence of all-cause hospitalisation • Adverse effects of treatment • PRO measured by KCCQ • Health-related quality of life measured by EQ-5D-5L
All other reported outcomes	• Other clinical outcome events o 3-point MACE (adjudicated CV death, adjudicated non-fatal MI, or adjudicated non-fatal stroke) o Adjudicated MI (fatal or non-fatal)

Study	EMPEROR-Preserved (NCT03057951) (126)
	○ Composite of adjudicated CV death or adjudicated non-fatal MI ○ Adjudicated stroke (fatal or non-fatal) ○ Composite of adjudicated CV death or adjudicated non-fatal stroke ○ Adjudicated TIA ○ Time to new onset of atrial fibrillation • NYHA class change from baseline • Body weight change from baseline • Blood pressure change from baseline • Pulse rate change from baseline • NT-proBNP change from baseline • eGFR change from baseline • Albuminuria • Health economic analysis by HCRU

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin II receptor blocker; CSR, clinical study report; CV, cardiovascular; DM, diabetes mellitus; EF, ejection fraction; EQ-5D, EuroQol five dimensions questionnaire; HbA1c, glycated haemoglobin; HCRU, Health care resource utilisation; HF, heart failure; HHF, hospitalisation for heart failure; eGFR, estimated glomerular filtration rate; KCCQ, Kansas City Cardiomyopathy Questionnaire; EF, left ventricular ejection fraction; MACE, Major adverse cardiovascular events; MI, myocardial infarction; NT-proBNP, N-terminal pro-brain natriuretic peptide; NYHA, New York Heart Association classification; PO, per os; PRO, patient-reported outcome; SoC, standard of care; TIA, transient ischaemic attack.

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

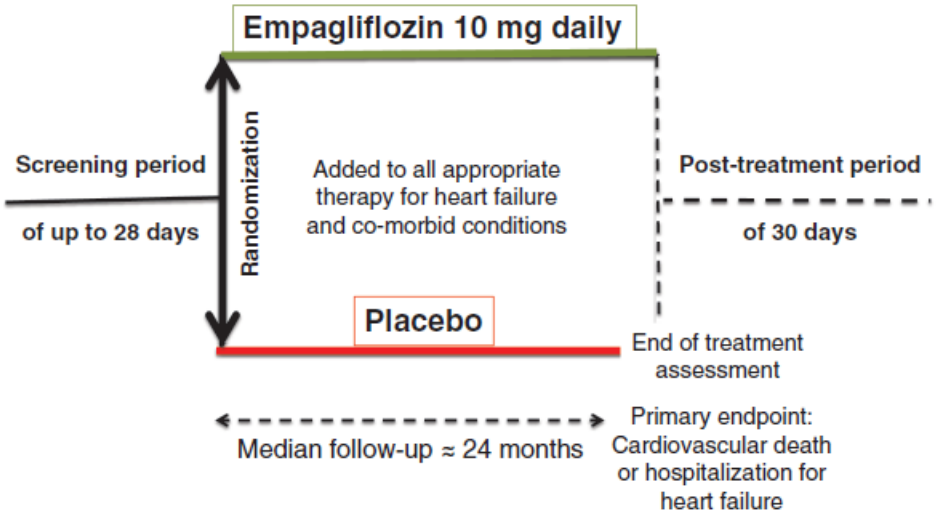
B.2.3.1 Summary of methodology of the EMPEROR-Preserved trial (NCT03057951)

EMPEROR-Preserved was an international phase III study designed to evaluate the long-term efficacy and safety of empagliflozin *versus* placebo in addition to guideline-directed medical therapy in patients with symptomatic chronic HF (EF >40%) (5). The trial had a double-blind, placebo-controlled, randomised design with parallel assignment of participants in 1:1 ratio to one of the two treatment arms:

- Empagliflozin, 10 mg PO once daily in addition to SoC (usual therapy i.e., treatments used to treat CV comorbidities which could include treatment with a low to medium dose of loop diuretic, ACEI, ARB, mineralocorticoid receptor antagonist, beta-blocker and/or sacubitril/valsartan), *or*
- Placebo PO once daily in addition to the SoC.

Following a screening period lasting 4–28 days, patients who fulfilled all eligibility criteria were randomised to receive placebo or empagliflozin daily in addition to their usual therapy for HF. The EMPEROR-Preserved trial design is illustrated in Figure 5 (121).

Figure 5. Design of EMPEROR-Preserved trial



Source: Adapted from Anker et al, 2019 (121).

Randomisation was performed using a permuted block design with a computer pseudo-random number generator and was stratified by:

- geographical region (North America, Latin America, Europe, Asia or “Other”),
- history of diabetes (diabetes, pre-diabetes and no diabetes),
- eGFR (by the Chronic Kidney Disease - Epidemiology Collaboration Equation [CKD-EPI] equation) at screening <60 or ≥60 mL/min/1.73 m², and
- EF (<50%, ≥50%).

Following randomisation, all appropriate treatments for HF or other medical conditions were initiated and individualised at the discretion of each subject’s physician. Patients were evaluated periodically at pre-specified study visits.

The primary objective of the EMPEROR-Preserved was to compare the time to first event of adjudicated CV death or adjudicated HHF among patients taking empagliflozin relative to those taking placebo in addition to their standard CRM

therapy. The trial also evaluated the effects of empagliflozin on recurrent HHF, renal function, CV death, all-cause mortality, and quality of life.

EMPEROR-Preserved was an event-driven trial and all randomised patients remained in the study until the defined number of adjudicated primary endpoint events were reached. As such, EMPEROR-Preserved was appropriately designed to determine if the addition of empagliflozin can improve outcomes of chronic HF (EF >40%) since the current treatment options show limited benefit for patients with chronic HF and a preserved EF. Aspects of the trial methodology are described in more detail below in accordance with the Consolidated Standards of Reporting Trials (CONSORT) statement (140).

As mentioned earlier, the trial design of EMPEROR-Preserved was similar to that of EMPEROR-Reduced, i.e., both were phase III international, multicentre, randomised, double-blind, parallel-group, placebo-controlled trials, but the EMPEROR-Reduced trial enrolled patients with a baseline EF ≤40%, and the EMPEROR-Preserved trial enrolled those who had EF >40% (4). The EMPEROR-Reduced trial design has been described in the NICE appraisal TA773 (1).

B.2.3.1.1 Changes to trial design

The description of the EMPEROR-Preserved methodology outlined in this submission is based on the revised study protocol number c03946327-04 which was issued on 20 November 2019 and incorporates Global Amendment 03.

B.2.3.1.2 Eligibility criteria for study participants

The intent of the EMPEROR-Preserved trial was to recruit chronic HF (EF >40%) patients on various HF background therapies to evaluate the long-term effect of empagliflozin on CV death and HHF in a -real life clinical setting. The trial, therefore, included adult patients with chronic HF with EF >40% diagnosed at least 3 months before screening and in the functional NYHA class II-IV. Details of inclusion and exclusion criteria are presented in Table 12.

Table 12. Inclusion and exclusion criteria of the EMPEROR-Preserved trial

Inclusion criteria	• Males and females aged ≥18 years; for Japan only: age ≥20 years • Patients with chronic HF diagnosed for at least 3 months before screening, and currently in HF NYHA class II-IV • Chronic HF with preserved EF defined as EF >40% per local reading • In addition to EF >40%, patient must have NT-proBNP >300 pg/mL for patients without AF, or >900 pg/mL for patients with AF (analysed at the Central Laboratory at screening) • Patients with either documented structural heart disease (left atrial enlargement and/or left ventricular hypertrophy) within 6 months or HHF documented within 12 months prior to screening • Oral diuretics, if prescribed to patients according to local guidelines and discretion of the investigator, must be stable for at least one week prior to randomisation • BMI <45 kg/m^2 at screening
Exclusion criteria	• Myocardial infarction (increase in cardiac enzymes in combination with symptoms of ischaemia or new ischaemic ECG changes), CABG or other major CV surgery, stroke or transient ischaemic attack in past 90 days • Heart transplant recipient or listed for heart transplant • Cardiomyopathy based on infiltrative diseases (amyloidosis), accumulation diseases (haemochromatosis, Fabry disease), muscular dystrophies, cardiomyopathy with reversible causes (e.g., stress cardiomyopathy), hypertrophic obstructive cardiomyopathy or known pericardial constriction • Any severe (obstructive or regurgitant) valvular heart disease expected to lead to surgery during the trial period • Acute decompensated HF requiring intravenous diuretics, vasodilators, inotropic agents or mechanical support within 1 week of screening and during the screening period prior to randomisation • Atrial fibrillation or atrial flutter with a resting heart rate >110 bpm, documented by ECG at screening • SBP ≥180 mmHg at randomisation. If SBP is 151–179 mmHg, the patient should be receiving ≥3 anti-hypertensive drugs • Symptomatic hypotension and/or a SBP <100mmHg at screening or at randomisation • Chronic PD requiring home oxygen, oral corticosteroid therapy or hospitalisation for exacerbation within 12 months, significant chronic PD or primary pulmonary arterial hypertension • Indication of liver disease, defined by serum levels of either ALT (SGPT), AST (SGOT), or alkaline phosphatase above 3x ULN as determined at randomisation • Impaired renal function, defined as eGFR <20 mL/min/1.73 m^2 (CKD-EPI) or requiring dialysis at the time of screening • Haemoglobin <9 g/dL at screening • History of ketoacidosis • Major surgery performed within 90 days prior to screening or major scheduled elective surgery (e.g., hip replacement) within 90 days after screening

	• GI surgery or GI disorder that could interfere with medication absorption • Any documented active or suspected malignancy or history of malignancy within 2 years prior to screening, except appropriately treated basal cell carcinoma of the skin, *in situ* carcinoma of uterine cervix or low risk prostate cancer • Presence of any other disease with a life expectancy of <1 year in the opinion of the investigator) • Current use or prior use of a SGLT2i or combined inhibitor of SGLT1 and SGLT2 within 12 weeks prior to screening or randomisation • Currently enrolled in another investigational device or drug study or are less than 30 days since the completion of a trial of another investigational device or drug. Any patient receiving any investigational treatment other than the study medications for this trial • Known allergy or hypersensitivity to any SGLT2i • Chronic alcohol or drug abuse or any condition that, in the investigator's opinion, will make the patient unlikely to fulfil the trial requirements or complete the trial • Women who are pregnant or are nursing or who plan to become pregnant while in the trial • Any other clinical condition that would jeopardise patient safety while participating in this trial or may prevent the subject from adhering to the trial protocol

Abbreviations: AF, atrial fibrillation or atrial flutter; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; bpm, beats per minute; CABG, coronary artery bypass grafting; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration equation; CV, cardiovascular; ECG electrocardiogram; EF, ejection fraction; eGFR, estimated glomerular filtration rate; GI, gastrointestinal; ICD, implantable cardioverter-defibrillator; HF, heart failure; EF, left ventricular ejection fraction; NYHA, New York Heart Association classification; NT-proBNP, N-terminal prohormone B-type natriuretic peptide; PD, pulmonary disease; SGLT, sodium-glucose co-transporter; SGLT2i, sodium-glucose co-transporter 2 inhibitor; SGOT, serum glutamic-oxaloacetic transaminase; SGPT, serum glutamic-pyruvic transaminase; ULN, upper limit of normal.

B.2.3.1.3 Study locations

Patient enrolment (N=5,988) started on 27 March 2017 in university hospitals, specialist CV clinics and clinical research centres across 622 locations in 23 countries (Poland, Czech Republic, Hungary, Germany, the Netherlands, Italy, Romania, Spain, Belgium, UK, Brazil, Argentina, Colombia, Mexico, US, Canada, Japan, China, Korea, Singapore, South Africa, Australia and India). From the UK, 53 patients were enrolled of which 25 patients were randomised and treated. The study completion date was 26 April 2021.

B.2.3.1.4 Trial drugs and concomitant medications

Study interventions are summarised in Table 13. The use of medication for the treatment of HF was at the discretion of the investigator and was to be in accordance with local and international guidelines. Disallowed concomitant medications included any SGLT2i or combined SGLT1 and 2 inhibitors, except the blinded trial medication.

Table 13. EMPEROR-Preserved trial drugs

Drug	Dose	Frequency of administration	Route of administration	Duration
Empagliflozin, film coated tablet	10 mg	Once daily	Oral	Until the necessary number of events were observed to evaluate efficacy for the primary composite endpoint
Placebo matching empagliflozin, film coated tablet	-			

B.2.3.1.5 Pre-specified primary and secondary outcomes of EMPEROR-Preserved

The endpoints relevant for the decision problem are summarised in Table 14. The definitions of adjudicated CV endpoints are summarised in Table 15.

Table 14. Pre-specified primary and secondary outcomes

Primary endpoint	Definition	NICE scope/ economic model?
Combined risk of CV death or HHF	A composite of adjudicated CV death or HHF, analysed as the time to the first event	Per NICE scope, not included in the economic model as a composite outcome
Key secondary endpoints	**Definition**	**NICE scope/economic model?**
Total HHF (first and recurrent)	Occurrence of adjudicated HHF (first and recurrent)	Per NICE scope, included in the economic model
Rate of renal function decline	eGFR (CKD-EPI)cr slope of change from baseline	Per NICE scope, not included in the economic model
Other secondary endpoints	**Definition**	**NICE scope/economic model?**
Risk of composite renal endpoint	Time to first event in the composite renal endpoint: occurrence of chronic dialysis[a] or	Per NICE scope, included in the economic model

(chronic dialysis, renal transplant or renal insufficiency)	renal transplant or sustained[b] reduction in eGFR (CKD-EPI)cr[c]	
Risk of first HHF	Time to first adjudicated HHF	Per NICE scope, not included in the economic model
Risk of CV death	Time to adjudicated CV death	Per NICE scope, included in the economic model
Risk of death	Time to all-cause mortality	Per NICE scope, included in the economic model
Risk of diabetes mellitus	Time to onset of DM defined as HbA1c ≥6.5% or as diagnosed by the Investigator in patients with pre-DM (defined as no history of DM and no HbA1c ≥6.5% before treatment, and a pre-treatment HbA1c value of ≥5.7% and <6.5%)	Not in scope, not included in the economic analysis
Change in KCCQ clinical summary score	Change from baseline in the KCCQ clinical summary score (HF symptoms and physical limitations domains) at week 52	Per NICE scope, included in the economic model
Risk of all-cause hospitalisation	Occurrence of all-cause hospitalisation (first and recurrent)	Per NICE scope, not included in the economic model
Further endpoints	**Definition**	**NICE scope/economic model?**
Risk of atrial fibrillation	New onset of atrial fibrillation	Not in scope, not included in the economic model
Risk of myocardial infarction	Adjudicated myocardial infarction (fatal or non-fatal)	Not in scope, not included in the economic model
Risk of stroke	Adjudicated stroke (fatal or non-fatal)	Not in scope, not included in the economic model
Safety	Adverse events, adverse events of special interest and specific adverse events	Per NICE scope, included in the economic model

Abbreviations: CV, cardiovascular; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration equation; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; HbA1c, haemoglobin A1c; HF, heart failure; HHF, hospitalisation for heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire.

[a]Chronic dialysis was defined as dialysis with a frequency of twice per week or more for at least 90 days

[b]Sustained was determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (the first to last of the consecutive eGFR values)

[c]Reduction in eGFR (CKD-EPI)cr was defined as reduction in eGFR from baseline of ≥40%, eGFR <15 mL/min/1.73 m^2 for patients with baseline eGFR ≥30 mL/min/1.73 m^2, or eGFR <10 mL/min/1.73 m^2 for patients with baseline eGFR <30 mL/min/1.73 m^2

Table 15. Definitions of adjudicated endpoints

Endpoint	Definition[a]
HHF	HHF endpoint must meet the following criteria: • Adjudicated primary diagnosis is admission to hospital for HF • Length of stay in hospital extends for ≥12 hours (emergency room visit for ≥12 hours with IV therapy is considered equivalent to admission to hospital) • The patient exhibits documented new or worsening symptoms due to HF on presentation, including at least one of the following: o Dyspnoea (dyspnoea with exertion, dyspnoea at rest, orthopnoea, paroxysmal nocturnal dyspnoea) o Decreased exercise tolerance o Fatigue o Other symptoms of worsened end-organ perfusion (dizziness, confusion, or volume overload such as weight gain or lower extremity swelling) • Objective evidence of new or worsening HF consisting of at least two physical examination findings or one physical examination finding and at least one laboratory criterion, including: o Physical examination findings considered to be due to HF: - Peripheral oedema - Increasing abdominal distension or ascites - Pulmonary rales/crackles/crepitations - Increased jugular venous pressure and/or hepatojugular reflux - S3 gallop - Clinically significant rapid weight gain related to fluid retention o Laboratory evidence of new or worsening HF, if obtained within 24 hours of presentation, including: - Increased BNP/NT pro-BNP concentrations consistent with decompensation of HF - Radiological evidence of pulmonary congestion - Non-invasive evidence of clinically significant left- or right- sided ventricular filling pressure or low cardiac output, *or* - Invasive diagnostic evidence with right heart catheterisation showing a pulmonary capillary wedge pressure ≥18 mmHg, central venous pressure ≥12 mmHg, or a cardiac index <2.2 L/min/L[2] • The patient receives initiation or intensification of treatment for HF, including at least one of the following: o Augmentation in oral diuretic therapy o IV diuretic or vasoactive agent (e.g., inotrope, vasopressor, or vasodilator) o Mechanical or surgical intervention (circulatory support with intra-aortic balloon pump, ventricular assist device, extracorporeal membrane oxygenation, total artificial heart or fluid removal with ultrafiltration, hemofiltration, dialysis)

Endpoint	Definition[a]
CV death	CV death includes the following categories: • Death due to MI, a procedure to treat MI or elective coronary procedure to treat myocardial ischaemia • Death due to clinically worsening signs and symptoms of HF including cardiogenic shock and pulmonary oedema • Death due to stroke, CV procedures, CV haemorrhage or other CV causes (e.g., pulmonary embolism or peripheral arterial disease) • Sudden cardiac death, including: o Death witnessed and occurring without new or worsening symptoms o Death witnessed within 60 minutes of the onset of new or worsening cardiac symptoms o Death witnessed and attributed to an identified arrhythmia or unwitnessed but found on implantable cardioverter-defibrillator review o Death after unsuccessful resuscitation from cardiac arrest or successfully resuscitated from cardiac arrest without identification of a specific cardiac or non-cardiac aetiology o Unwitnessed death in a subject seen alive and clinically stable ≤ 72 hours prior to being found dead without any evidence supporting a specific non-CV cause of death

Abbreviations: CV, cardiovascular; HF, heart failure; HHF, hospitalisation for heart failure; BNP, B-type natriuretic peptide; NT pro-BNP, N-terminal pro-B- type natriuretic peptide; MI, myocardial infarction.

[a]All CV endpoint definitions were modifications of the guideline recommendations by Hicks et al 2014 (141).

B.2.3.2 Demographics and baseline characteristics

Patients in the empagliflozin and the placebo group were well balanced with respect to demographic and clinical characteristics at baseline (Table 16). About a third of patients were in each of the pre-defined EF categories (EF <50%, 50 to <60%, and ≥60%). An eGFR of <60 mL per minute per 1.73 m^2, a history of AF or flutter and T2DM were reported for around 50% of the patients. Majority of patients were in HF NYHA class II (81.5%) while a third of patients had HF diagnosis for at least 5 years before the trial.

Table 16. Demographic and baseline characteristics of randomised participants in EMPEROR-Preserved trial

Baseline characteristic[a]	Empagliflozin 10 mg	Placebo
Number of subjects	2,997	2,991
Age (years), mean (SD)	71.8 (9.3)	71.9 (9.6)

Baseline characteristic[a]	Empagliflozin 10 mg	Placebo
Female sex, N (%)	1,338 (44.6)	1,338 (44.7)
Race, N (%)[b]		
White	2,286 (76.3)	2,256 (75.4)
Black	133 (4.4)	125 (4.2)
Asian	413 (13.8)	411 (13.7)
Other including mixed race	164 (5.5)	198 (6.6)
Region, N (%)		
North America	360 (12.0)	359 (12.0)
Latin America	758 (25.3)	757 (25.3)
Europe	1,346 (44.9)	1,343 (44.9)
Asia	343 (11.4)	343 (11.5)
Other	190 (6.3)	189 (6.3)
NYHA functional class, N (%)		
I	3 (0.1)	1 (<0.1)
II	2,432 (81.1)	2,451 (81.9)
III	552 (18.4)	531 (17.8)
IV	10 (0.3)	8 (0.3)
Body mass index[c] (kg/m^2), mean	29.8+/-5.8	29.9+/-5.9
Heart rate (beats/min), mean	70.4+/-12.0	70.3+/-11.8
SBP (mm Hg), mean	131.8+/-15.6	131.9+/-15.7
DBP (mm Hg), mean	[redacted] +/-[redacted]	[redacted] +/-[redacted]
Left ventricular ejection fraction (EF)		
Mean	54.3+/-8.8	54.3+/-8.8
Value of >50%, N (%)	995 (33.2)	988 (33.0)
Value of 50% to >60, N (%)	1,028 (34.3)	1,030 (34.4)

Baseline characteristic[a]	Empagliflozin 10 mg	Placebo
Value of ≥60%, N (%)	974 (32.5)	973 (32.5)
NT-proBNP (pg/mL)		
Median (IQR)	994 (501-1740)	946 (498-1725)
Time since diagnosis of HF (years)		
Mean	4.5+/-5.2	4.3+/-5.0
≤1 year, N (%)	730 (24.4)	782 (26.1)
>1 to 5 years, N (%)	1,368 (45.6)	1,325 (44.3)
>5 to 10 years, N (%)	550 (18.4)	553 (18.5)
>10 years, N (%)	349 (11.6)	331 (11.1)
Cause of HF, N (%)		
Ischaemic	1,079 (36.0)	1,038 (34.7)
Nonischaemic	1,917 (63.9)	1,953 (65.3)
CV history, N (%)		
Hospitalisation for HF in ≤12 months	699 (23.3)	670 (22.4)
Atrial fibrillation	1,543 (51.5)	1,514 (50.6)
Hypertension	2,721 (90.8)	2,703 (90.4)
Estimated glomerular filtration rate		
Mean (mL/min/1.73 m^2)	60.6+/-19.8	60.6+/-19.9
Value of <60 mL/min/1.73 m^2, No (%)	[redacted] ([redacted]2)	[redacted] ([redacted])
UACR (mg/mL)		
Normal (<30), N (%)	[redacted] ([redacted])	[redacted] ([redacted])
Microalbuminuria (30 to ≤300), N (%)	[redacted] ([redacted])	[redacted] ([redacted])
Macroalbuminuria (>300), N (%)	[redacted] ([redacted])	[redacted] ([redacted])

Baseline characteristic[a]	Empagliflozin 10 mg	Placebo
HF medication, N (%)		
ACEI/ARB/ARNI	[redacted]	[redacted]
ARNI	[redacted]	[redacted]
Beta-blocker	[redacted]	[redacted]
Diuretics	[redacted]	[redacted]
Lipid-lowering drugs	[redacted]	[redacted]
Anti-thrombotic drugs	[redacted]	[redacted]
Diabetes status		
Without diabetes, N (%)	[redacted]	[redacted]
Without diabetes or pre-diabetes, N (%)	[redacted]	[redacted]
With pre-diabetes, N (%)	[redacted]	[redacted]
With diabetes, N (%)	1,466 (48.9)	1,472 (49.2)
T2DM, N (%)	[redacted]	[redacted]
T1DM, N (%)	[redacted]	[redacted]

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; DBP, diastolic blood pressure; HF, heart failure; IQR, interquartile range; No, number; NT pro-BNP, N-terminal prohormone of brain natriuretic peptide; NYHA, New York Heart Association; SBP, systolic blood pressure; SD, standard deviation; T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus; UACR, urine albumin-to-creatinine ratio.
Notes: Patients with information missing are not shown.
[a]Plus-minus values are means ± SD. Percentages may not total 100 because of rounding.
[b]Race was reported by the patients. Those who identified with more than one race or with no race were classified as "other".
[c]The body mass index is the weight in kilograms divided by the square of the height in metres.
Source: EMPEROR-Preserved CSR Table 10.4.1: 1, 10.4.2: 1, 10.4.3: 1; Table 10.4.4.1 (139).

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

The statistical analysis methods and definitions of study groups used in the pivotal EMPEROR-Preserved trial are described in Table 17. The method of statistical analysis were similar to the EMPEROR-reduced trial, which has been assessed in the TA773 appraisal (1).

B.2.4.1 Statistical methods and analysis sets

Table 17. Summary of statistical analysis in the EMPEROR-Preserved RCT

Study name (number)	EMPEROR-Preserved (NCT03057951)
Research hypothesis	There is no difference between the efficacy of empagliflozin and efficacy of placebo in reducing the combined risk of CV death and HHF.

Analysis sets	• **Randomised set (RS):** All randomised patients, whether treated or not
	  ○ OC-AD: Observed case including data after treatment discontinuation
	  ○ OC-OT: Observed case on-treatment
	• **Treated set (TS):** All patients who were dispensed study medication and were documented to have taken at least one dose of investigational treatment.
	• **Treated Set-Follow-up (TS-FU):** All patients in the TS who performed the follow-up visit. Patients who took open-label SGLT-2 inhibitor between their end of treatment and follow-up visit were excluded from TS-FU.
	• **Pharmacokinetic set (PKS):** All patients in the TS who provided at least one observation for any empagliflozin PK parameter.

Statistical analysis for primary endpoint	The primary and key secondary endpoints were tested in the following hierarchical order: • Time to first event of adjudicated CV death or adjudicated HHF • Occurrence of adjudicated HHF (first and recurrent) • eGFR (CKD-EPI)cr slope of change from baseline For each of these confirmatory endpoints, superiority of empagliflozin over placebo was evaluated with a two-sided test. The overall type I error rate for the trial was preserved at α = 0.05. Due to the amount of α spent on the interim analysis, the remaining two-sided α level for the final analysis was 0.0497. The primary analysis was a Cox PH regression with factors treatment, geographical region, diabetes status at baseline, age, gender, EF, and baseline eGFR (CKD-EPI)cr. Following the ITT principle, the primary analysis was based on RS using all data up to the end of the planned treatment period (including the data after end of treatment for patients not completing the treatment phase as planned). Patients without a specific endpoint event were censored at the last date the patient was known to be event free or at the end of the planned treatment period, whichever was earlier. When violation of the PH assumption was observed, groups of patients for which the proportionality assumption held were identified, and a stratified Cox regression was performed.

Statistical analysis for key secondary endpoints	• **Occurrence of adjudicated HHF (first and recurrent)** was analysed by a joint frailty model that accounted for the dependence between recurrent HHF and CV death. The primary analysis included all data until completion of the planned treatment phase, including the data after end of treatment for patients not completing the treatment phase as planned. The model included the same covariates used for the analysis of the primary endpoint. The evaluation was assigned an alpha level of 0.0497. The joint model provided two distinct HRs:
	  ○ HR_{HHF} associated with the effect of treatment on the recurrent event rate of HHF
	  ○ HR_{CVD}, the HR for CV death.

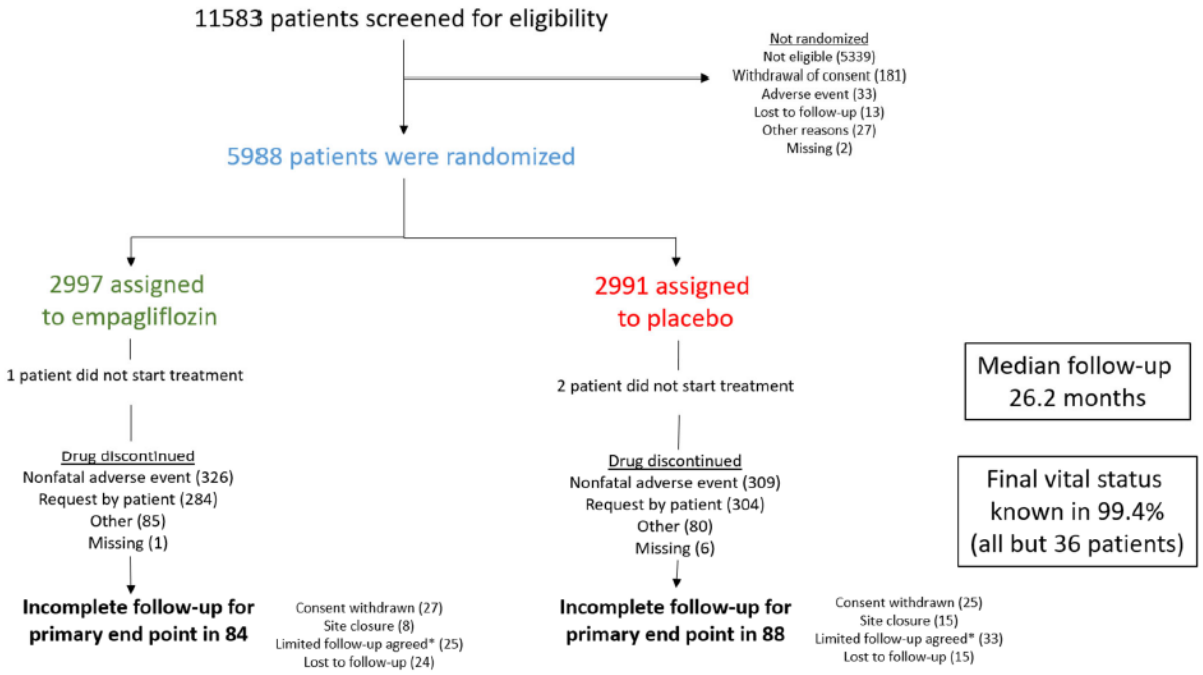
Study name (number)	EMPEROR-Preserved (NCT03057951)
	• **Slope in change from baseline of eGFR (CKD-EPI)cr** was analysed by a random coefficient model allowing for random intercept and random slope per patient, with the same factors used for the primary endpoint and the additional factors time, treatment-by-time and baseline eGFR (CKD-EPI)cr-by-time interaction as linear covariates. The model included all on-treatment change from baseline. This endpoint was tested with a two-sided α of 0.001.
Statistical analysis of exploratory endpoints	• **Time to event endpoints:** as analysis of primary endpoint • **Recurrent event endpoints:** as analysis of the first key secondary endpoint • **Continuous endpoints:** mixed model repeated measure (MMRM) analysis • **Categorial endpoints:** descriptive
Sample size & power calculation	Sample size calculation was based on the number of events needed to achieve power of 90% for a two-sided test with α=0.05 and HR 0.80. Achieving that treatment effect size required 841 primary endpoint events. Assuming a 10% event rate per year in the placebo arm, a recruitment period of 18 months and a follow-up period of 20 months, 4,126 patients needed to be randomised to receive empagliflozin or placebo in 1:1 manner.
Data management, patient withdrawals	• **Handling of drop-outs or missing data:** o For patients without primary event and lost to follow-up before trial completion, the treatment specific incidence rates for empagliflozin and placebo for retrieved drop-outs were used to impute the primary events in a multiple imputations framework. The primary model was applied to the imputed datasets. o There was no imputation of data for safety analyses or for time to event endpoints. o For endpoints of KCCQ scores in case of patients who die, a score of 0 was imputed at all subsequent scheduled visits where the score would have been assessed. o Missing covariates in multivariate Cox regression models and for recurrent event analyses were imputed using the overall population median of the corresponding variable for continuous covariates and the most frequent category for categorical covariates. No imputation was done for covariates included in treatment by subgroup interaction terms. • **Subjects could have been instructed to permanently discontinue study drug only after discussion with investigator if:** o eligibility criteria were violated o in the case of an AE o if the patient failed to comply with the protocol o if any restricted treatment was given during the trial

Abbreviations: AE, adverse event; CI, confidence interval; CV, cardiovascular; CVD, cardiovascular death; (CKD-EPI)cr, Chronic Kidney Disease Epidemiology Collaboration equation based on creatinine measurement; CPRD, Clinical Practice Research Datalink; DM, diabetes mellitus; EF, ejection fraction; EOT, end of treatment; EQ-5D, EuroQol 5 dimensions instrument; FU, follow-up; HES, Hospital Episode Statistics; HF, heart failure; HHF, hospitalisation for heart failure; HFrEF, heart failure with reduced ejection fraction; HR, hazard ratio; eGFR, estimated glomerular filtration rate; ITT, intention to treat; KCCQ, Kansas City Cardiomyopathy Questionnaire; EF, left ventricular ejection fraction; MMRM, mixed model repeated measure analysis; PH, proportional hazards; RS, randomised set; SCR, screened set; SGLT2, sodium-glucose co-transporter 2; SoC, standard of care; TS, treated set; TS-FU, treated set with follow-up.
Source: Anker et al 2021 (5); EMPEROR-Preserved CSR (139).

B.2.4.2 Participant flow in the relevant randomised controlled trials

Participant flow in EMPRERO-Preserved is shown in Figure 6.

Figure 6. CONSORT diagram of patient flow in each stage of EMPEROR-Preserved RCT



Notes: Incomplete follow-up for the primary end point refers to incomplete information on either vital status or hospitalisation until the planned end of the treatment period for those patients who had not experienced an adjudicated primary outcome. The 36 patients with unknown vital status at the end of the trial included 17 on empagliflozin and 19 on placebo. Five patients with missing vital status at the end of the trial experienced an adjudicated HHF and are not considered to have incomplete follow-up for the primary endpoint. Asterisk denotes patients who discontinued study medication before the trial end but agreed to collection of vital status data at trial completion.
Source: Anker et al 2021 (5).

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

A summary of the quality assessment of the EMPEROR-Preserved, a parallel-group RCT, is shown in Table 18. The complete quality assessment is provided in Appendix D. The quality appraisal was similar to the EMPEROR-reduced trial, which has been assessed in the TA773 appraisal (1).

Table 18. Results of the quality assessment of EMPEROR-Preserved trial

Quality assessment questions	EMPEROR-Preserved (NCT03057951)
Was randomisation carried out appropriately?	Yes. Randomisation was performed by using a permuted block design with a computer pseudo-random number generator.
Was the concealment of treatment allocation adequate?	Yes. An Interactive Response Technology System (voice response or web response) was used to determine treatment assignment.
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes. Demographic and patient characteristics were well balanced between the two treatment groups at baseline, and randomisation was stratified by geographical region, diabetes status and eGFR at screening.
Were the care providers, participants and outcome assessors blind to the treatment allocation?	Yes. This was a double-blind study. Independent external clinical event committees evaluated all reported and potential clinical events in a manner blinded to the treatment assignment.
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No. Proportion of patients who discontinued study treatment was low and well balanced between the two treatment groups.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No. All outcomes specified in the study protocol were reported in the clinical study report.
Did the analysis include an intention to treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes. Efficacy analysis were performed in the randomised set.

B.2.6 Clinical effectiveness results of the relevant trial: EMPEROR-Preserved

As described in sections that follow, the null hypotheses for the primary and the two key secondary endpoints of the EMPEROR-Preserved trial were rejected in a hierarchical testing procedure. Results of the trial demonstrate that empagliflozin is superior to placebo in improving HF outcomes in patients with symptomatically stable chronic HF (EF >40%). Empagliflozin treatment demonstrated a clinically meaningful and statistically significant reduction in the risk of CV death or HHF (primary endpoint), compared with placebo. The benefit was consistent across pre-specified EF

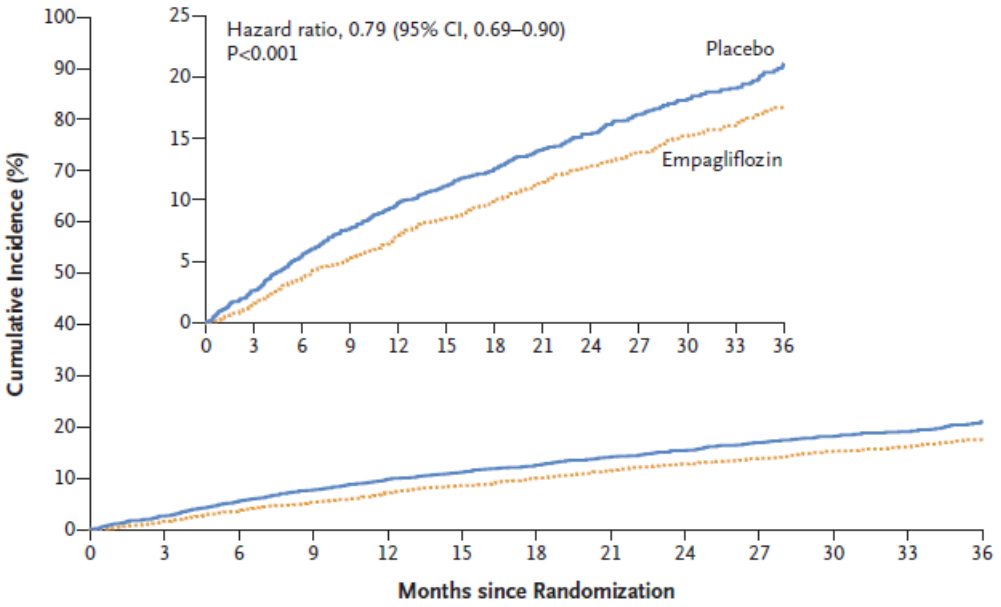
subgroups and irrespective of diabetes status at baseline. In comparison to placebo, addition of empagliflozin to SoC is also associated with a slower decline of renal function as assessed by eGFR slope of change (5). It should be noted that the primary and secondary outcomes (except time to composite renal outcome) of the EMPEROR-Preserved trial were similar to EMPEROR-Reduced trial which has been assessed in the NICE appraisal TA773 (1).

Results of the pre-specified efficacy outcomes that are within the scope of the decision problem are described in Sections B.2.6.1 to B.2.6.2. The results of a combined HF analysis stratified by EF will also be described in Section **Error! Reference source not found.**. Pre-specified subgroup analysis of the trial data used in the economic model is described in Appendix E.

B.2.6.1 Primary outcome: combined risk of CV death or HHF

The primary composite outcome of CV death or HHF occurred in a lower proportion of patients in the empagliflozin group (415 of 2,997 patients, 13.8%) than in the placebo group (511 of 2,991 patients, 17.1%). The separation of the estimated cumulative incidence of CV death or first HHF curves, considering non-CV death as a competing risk, started shortly after randomisation and was maintained throughout the trial period (Figure 7). Cox regression of data for all randomised patients adjusted for age, baseline eGFR (CKD-EPI)cr, region, gender, treatment, baseline diabetes status and EF, revealed that the risk of CV death or HHF was significantly reduced with empagliflozin compared with placebo (HR, 0.79; 95% CI, 0.69 to 0.90).

Figure 7. Primary outcome, a composite of CV death or HHF



No. at Risk													
Placebo	2991	2888	2786	2706	2627	2424	2066	1821	1534	1278	961	681	400
Empagliflozin	2997	2928	2843	2780	2708	2491	2134	1858	1578	1332	1005	709	402

Abbreviation: CI, confidence interval; CV, cardiovascular; HHF, hospitalisation for heart failure
The estimated cumulative incidence of the primary outcome in the two groups is shown. The inset shows the same data on an expanded y axis.
Source: Figure 1, Anker et al 2021 (5).

During a median trial period of 26 months, the number of patients treated with empagliflozin needed to prevent one primary outcome event was 31 (95% CI, 20 to 69).

Several sensitivity analyses of the primary endpoint were performed to consider competing risks and to account for missing follow-up data in 172 patients who discontinued trial prematurely. The results of sensitivity analysis were generally consistent with the results of the primary analysis, with similar HRs (Table 19).

Table 19. Sensitivity analyses for the primary endpoint: time to the first event of adjudicated CV death or HHF

Sensitivity analyses in RS	Hazard ratio (95% confidence interval)
Multiple imputation analysis addressing incomplete data for primary endpoint[a], RS	[redacted]
Results unadjusted for covariates, RS	[redacted]
Sub-distribution hazard ratio adjusted for non-CV death as a competing risk in RS (Fine-Gray model)[b]	[redacted]

Abbreviations: CV, cardiovascular; HHF, hospitalisation for heart failure; RS, randomised set.
[a]Imputations were performed for 172 patients with incomplete data (84 empagliflozin, 88 placebo). Treatment specific incidence rates for empagliflozin and placebo for patients who discontinued study medication with available follow-up data were used to impute the primary events in a multiple imputations framework via sampling from an exponential distribution. One hundred imputations were performed and evaluated by the primary model. Log hazard ratios were summarised by Rubin's rules (142).
[b]Fine and Gray, 1999 (143).
Source: EMPEROR-Preserved CSR, Figure 11.1.1.2: 1 (139).

B.2.6.2 Secondary outcomes

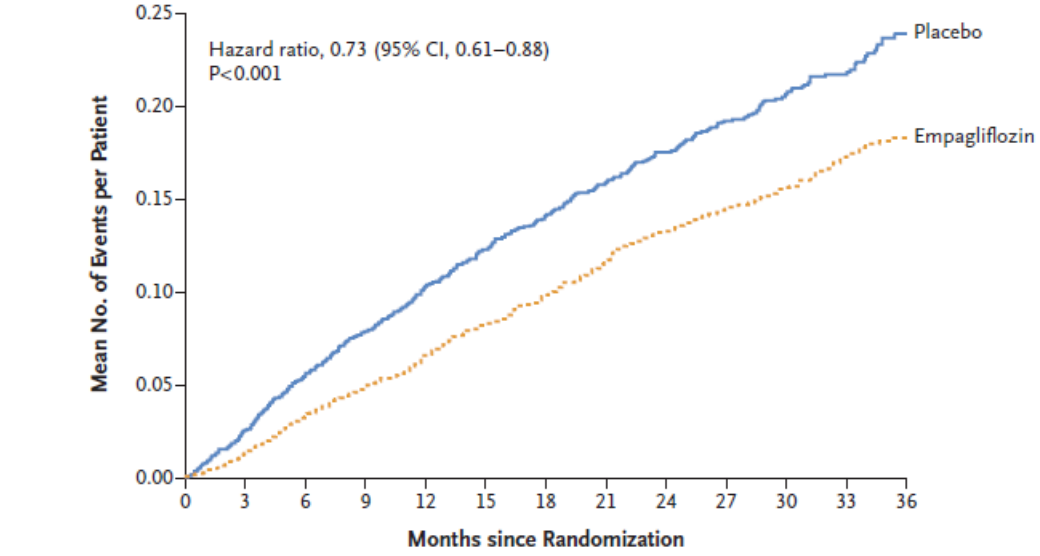
B.2.6.2.1 Total HHF (first and recurrent)

The total number of HHF event was lower in the empagliflozin group than in the placebo group with 407 events and 541 events, respectively. The mean cumulative incidence of HHF in the empagliflozin and placebo groups started to diverge shortly after randomisation and continued to segregate further over the course of the trial (

Source: Table S4, Anker et al 2021 (5).

). Primary analysis using joint frailty model with CV death as a competing risk demonstrated that the risk of recurrent HHF was significantly reduced with empagliflozin relative to placebo (HR, 0.73; 95% CI, 0.61 to 0.88, p=0.0009). The hazard of recurrent HHF was positively correlated to that of CV death, as indicated by a frailty exponent greater than zero (data not shown).

Figure 8. Key secondary outcome: Total number of HHF (first and recurrent)



Abbreviations: CI, confidence interval; HHF, hospitalisation for heart failure.
Source: Figure 3, Anker et al 2021 (5).

The results of the sensitivity analyses were consistent with the results of the primary analysis for the occurrence of adjudicated HHF (first and recurrent) (Table 20).

Table 20. Sensitivity analyses for the key secondary endpoint: total HHF

Sensitivity analyses	Hazard ratio (95% confidence interval)
Joint frailty model considering CV death as competing risk (primary analysis model)	0.73 (0.61 – 0.88)
Parametric joint gamma frailty model considering CV death as competing risk	0.73 (0.61-0.88)
Joint frailty model considering all-cause mortality as competing risk	0.75 (0.62-0.90)
Negative binomial model[a]	0.73 (0.60-0.89)
Negative binomial model without covariate adjustment[a]	0.74 (0.61-0.90)
Cox regression for time to first adjudicated HHF	0.71 (0.60-0.83)

Abbreviations: CV, cardiovascular; HHF, hospitalisation for heart failure; RS, randomised set; TS, treated set.
[a]Rate ratio is shown
Joint frailty model by Rogers et al. (2016) (144).
Source: Table S4, Anker et al 2021 (5).

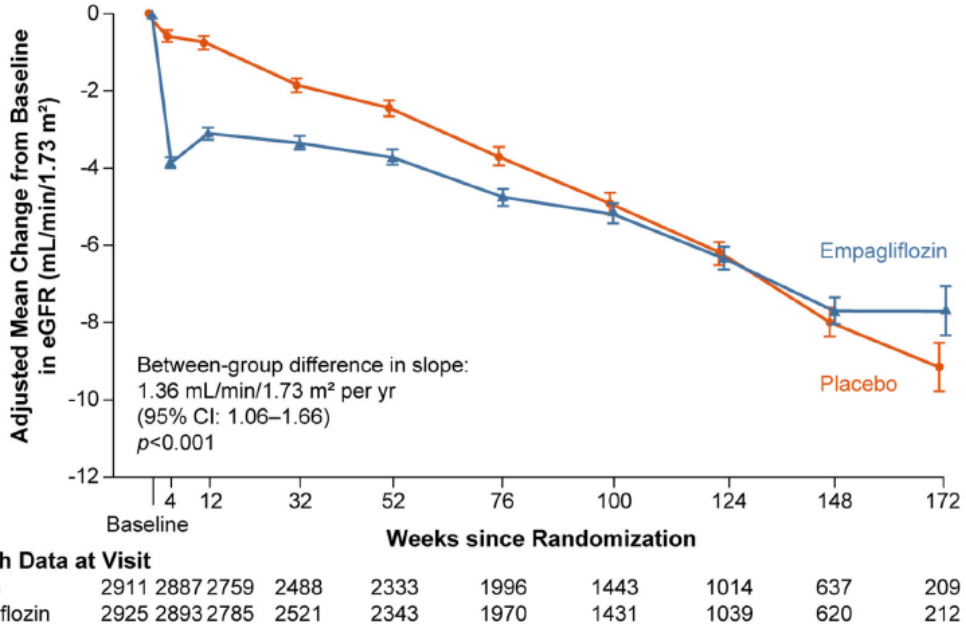
B.2.6.2.2 Deterioration of renal function

The other key secondary endpoint in the hierarchical testing procedure was mean slope of change in eGFR (mL/min/1.73 m^2) from baseline. Estimation of glomerular filtration rate (GFR) was based on (CKD-EPI)cr (145).

The primary analysis included only “on-treatment” data from the treated set (TS) and measurements up to one day after the last intake of study medication. In the empagliflozin group, the estimated slope was -1.25 ± 0.11 mL/min/1.73 m^2 per year. In the placebo group, eGFR declined more steeply over the duration of the treatment period, with an estimated slope of -2.62 ± 0.11 mL/min/1.73 m^2 per year. The estimated between-group difference in mean slope was 1.36 mL/min/1.73 m^2 per year (95% CI, 1.06 to 1.66; p<0.0001) (

Figure 9). In the RS, the adjusted mean eGFR change from baseline to follow-up was 2.4 (95% CI, 1.6 to 3.2) mL/min/1.73m^2 per year for empagliflozin *versus* placebo.

Figure 9. Changes in the estimated glomerular filtration rate, based on the TS and measurements up to one day after the last intake of study medication



Abbreviations: eGFR, estimated glomerular filtration rate; EF, left ventricular ejection fraction; MMRM, mixed model for repeated measures; TS, treated set.
Notes: Graph shows the adjusted mean changes from baseline in the eGFR as calculated using the Chronic Kidney Disease Epidemiology Collaboration equation. The bars indicate the standard error. The on-treatment data were analysed with MMRM. Age, baseline eGFR and EF were included as linear covariates, while sex, region, baseline diabetes status, last projected visit based on dates of randomisation and trial closure, baseline eGFR according to visit, and visit according to treatment interactions were included as fixed effects.
Source: Figure S4, Anker et al 2021 (5).

Thus, when measurements of renal function were compared at the start and after the discontinuation of empagliflozin and placebo, the eGFR declined significantly more in the placebo group than in the empagliflozin group, leading to increased risk of serious renal outcomes, as described in more detail in Section B.2.6.2.3. The initial dip in eGFR seen at the start of the treatment with empagliflozin represents a reversible

functional change in intrarenal haemodynamics commonly observed with SGLT2i and is not associated with an excess risk of investigator-reported acute kidney injury (146, 147).

B.2.6.2.3 Time to composite renal outcome

The composite renal endpoint occurred in 108 patients (3.6%) in the empagliflozin group and 112 patients (3.7%) in the placebo group, with the sustained reduction in eGFR from baseline of ≥40% being the first recorded renal event in most patients (Table 21). The risk of the composite renal endpoint was similar between the empagliflozin and the placebo treatment group (HR, 0.95; 95% CI, 0.73 to 1.24; nominal p=0.7243). The estimated cumulative incidence function for the time to the first event of the composite renal endpoint (considering all-cause mortality as a competing risk) is shown in Figure 10.

Table 21. Cox regression analysis of time to first renal event[a] in all randomised patients

Time to composite renal outcome[b]	Empagliflozin (N=2,997)	Placebo (N=2,991)
Patients with the composite renal endpoint, N (%)	108 (3.6)	[redacted]
Only sustained eGFR reduction ≥40% as the first event	95 (3.2)	[redacted]
Chronic dialysis as the first event	[redacted]	[redacted]
Sustained eGFR <15 mL/min/1.73 m^2 (baseline ≥30) or <10 mL/min/1.73 m^2 (baseline <30) as the first event	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
Hazard ratio *vs.* placebo (95% CI), composite renal outcome	[redacted]	
Nominal p-value	[redacted]	

Abbreviations: eGFR, estimated glomerular filtration rate; CI, confidence interval.
[a]Cox regression model included covariates age, baseline eGFR (CKD-EPI)cr, region, baseline diabetes status, sex, baseline EF, and treatment.
[b]The composite renal endpoint was comprised of chronic dialysis (with a frequency of twice per week or more for at least 90 days), renal transplant, sustained reduction in eGFR from baseline of ≥40%, sustained eGFR <15 mL/min/1.73 m^2 for patients with baseline eGFR ≥30 mL/min/1.73 m^2, or sustained eGFR <10 mL/min/1.73 m^2 for patients with baseline eGFR <30 mL/min/1.73 m^2. Sustained was determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (the first to last of the consecutive eGFR values).
Source: EMPEROR-Preserved CSR, Table 11.1.2.6: 1 (139).

Figure 10. Estimated cumulative incidence function for time to the first event of the composite renal endpoint in all randomised patients

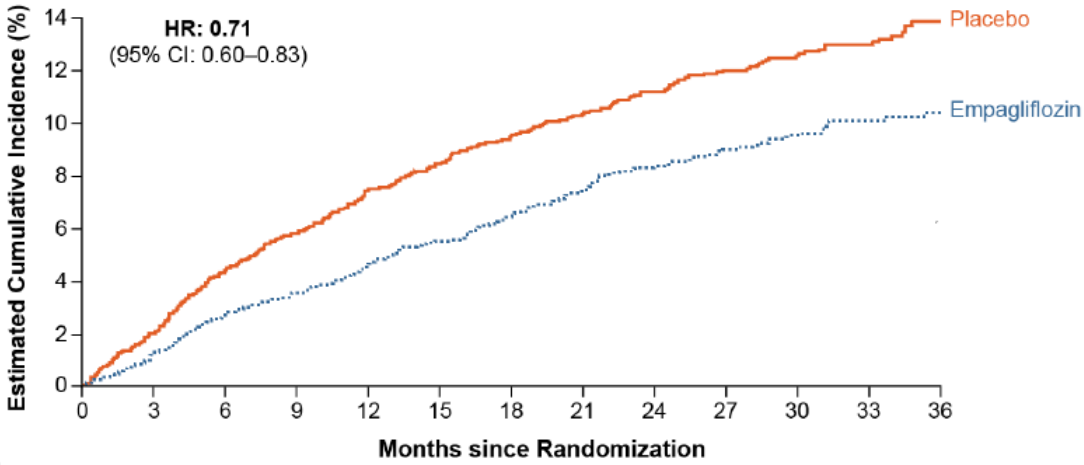
[redacted]

Source: EMPEROR-Preserved CSR, Figure 11.1.2.6.1: 1 (139).

B.2.6.2.4 Time to first adjudicated HHF

Over the duration of the trial, fewer patients experienced the event of first adjudicated HHF in the empagliflozin group (259 of 2,997, 8.6%) compared to placebo group (352 of 2,991, 11.8%). The estimated cumulative incidence of first adjudicated HHF, considering all-cause mortality as a competing risk, started to diverge between empagliflozin and placebo groups shortly after randomisation and continued to separate over the course of the trial (Figure 11). The risk of adjudicated HHF was significantly reduced with empagliflozin treatment *versus* placebo (HR, 0.71; 95% CI, 0.60 to 0.83; nominal p<0.0001), as determined by the Cox regression model adjusted for age, baseline eGFR (CKD-EPI)cr, region, baseline diabetes status, gender, baseline EF, and treatment.

Figure 11. Time to the first adjudicated HHF



No. at Risk													
Placebo	2991	2888	2786	2706	2627	2424	2066	1821	1534	1278	961	681	400
Empagliflozin	2997	2928	2843	2780	2708	2491	2134	1858	1578	1332	1005	709	402

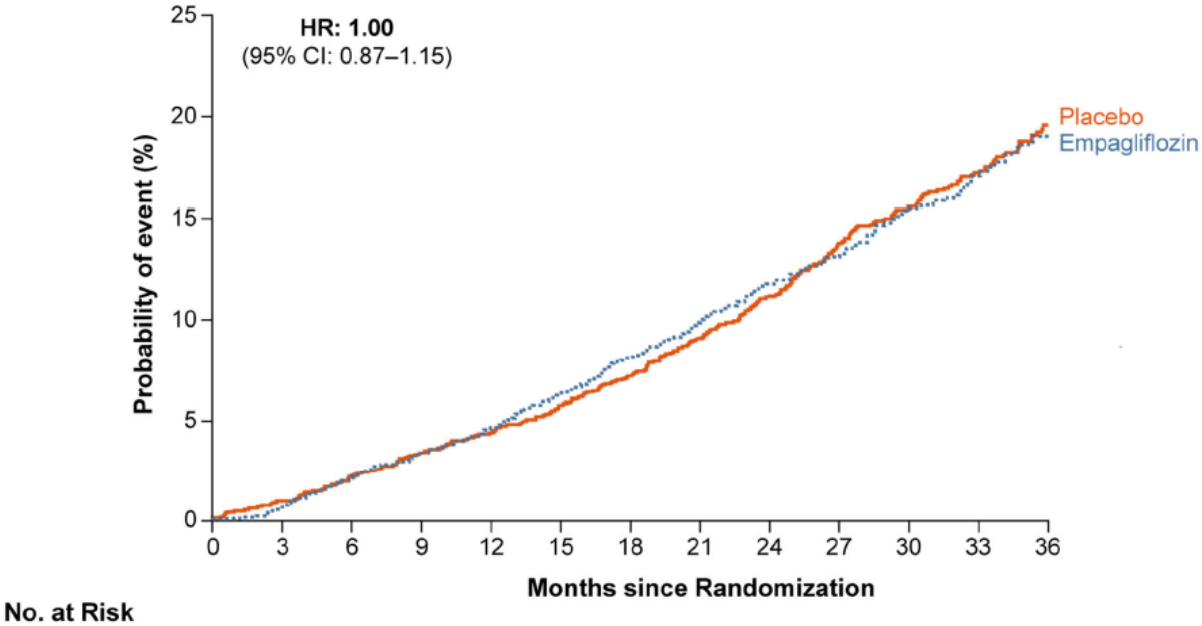
Source: Figure S3, Anker et al 2021 (5).

B.2.6.2.5 All-cause mortality

The Kaplan-Meier estimate of time to all-cause mortality in the RS is shown in

Figure 12. Death from any cause occurred in 422 patients (14.1%) in the empagliflozin group and 427 patients (14.3%) in the placebo group. Cox regression of time to all-cause mortality data for all randomised patients showed that the risk of death from any cause was of equivalence (HR, 1.00; 95% CI, 0.87 to 1.15) and the difference did not reach statistical significance (p=0.9893).

Figure 12. Kaplan-Meier estimate of time to all-cause mortality in all randomised patients

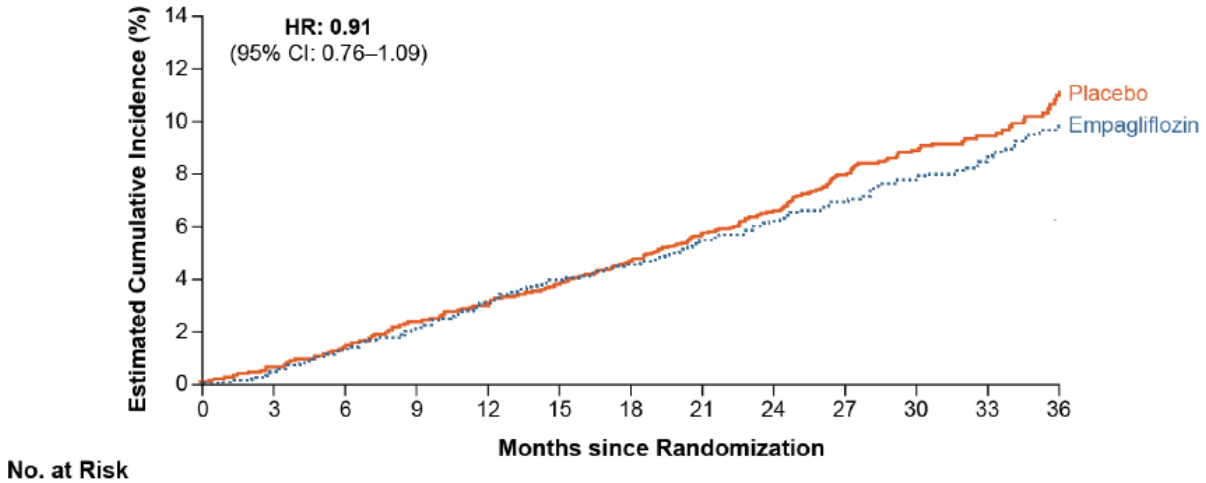


Source: Figure S5, Anker et al 2021 (5).

B.2.6.2.6 Cardiovascular mortality

Most of the deaths recorded during the study were due to CV causes, such as sudden cardiac death or HF. Adjudicated CV death occurred in 219 patients (7.3%) in the empagliflozin group and 244 patients (8.2%) in the placebo group. The risk of CV death was 0.9% lower with empagliflozin relative to placebo (HR, 0.91; 95% CI, 0.76 to 1.09), a difference that did not reach statistical significance (p=0.30). The cumulative incidence of adjudicated CV death in randomised patients, considering non-CV death as a competing risk, is shown in Figure 13.

Figure 13. Cardiovascular death



No. at Risk													
Placebo	2991	2960	2923	2883	2849	2656	2302	2046	1738	1459	1107	781	471
Empagliflozin	2997	2974	2930	2891	2847	2641	2287	2022	1725	1469	1118	794	462

Abbreviation: CI, confidence interval; HR, hazard ratio.
Source: Figure S3, Anker et al 2021 (5).

B.2.6.2.7 Time to onset of diabetes mellitus (DM) in patients with pre-DM

The onset of diabetes mellitus (DM) in patients with pre-DM occurred in 120 of 1,001 patients in the empagliflozin group (12.0%) and 137 of 979 patients (14.0%) in the placebo group. The observed reduction in risk of onset of DM with empagliflozin compared to placebo (HR, 0.84; 95% CI, 0.65 to 1.07) was not statistically significant (nominal p=0.15). The estimated cumulative incidence of time to onset of DM in patients with pre-DM, considering all-cause mortality as a competing risk, started to diverge after approximately 3 months, and was maintained over the remainder of the trial (Figure 14).

Figure 14. Estimated cumulative incidence function for time to onset of DM in patients with baseline pre-DM in the randomised set

[redacted]

Source: EMPEROR-Preserved CSR, Figure 11.1.2.8.1: 1 (139).

B.2.6.2.8 First and recurrent all-cause hospitalisation

All-cause hospitalisation occurred in 42.4% (1,271 of 2,997) of patients in the empagliflozin group and 44.8% (1,340 of 2,991) in the placebo group. The total number of hospitalisation events was lower in the empagliflozin group (2,566) than in the placebo group (2,769). Analysis of this endpoint using a joint frailty model that accounts for the dependence between recurrent all-cause hospitalisation and all-cause mortality demonstrated that the risk of recurrent all-cause hospitalisation was reduced with empagliflozin treatment compared to placebo (HR, 0.93; 95% CI, 0.85 to 1.01; nominal p=0.10). The mean cumulative incidence curves of all-cause hospitalisation in empagliflozin and placebo groups diverged at about 90 days after randomisation and maintained their separation throughout the study (Figure 15). Cox regression showed 2.33% reduction in risk of first all-cause hospitalisation with empagliflozin compared to placebo (HR, 0.92; 95% CI, 0.85 to 0.99; p=0.03).

Figure 15. Mean cumulative function for occurrence of all-cause hospitalisation (first and recurrent) in the randomised set

Source: EMPEROR-Preserved CSR, Figure 11.1.2.5.1: 1 (139).

B.2.6.2.9 Further secondary clinical endpoints

Results of further exploratory secondary endpoints from EMPEROR-Preserved trial, including measurement of health status by KCCQ, are presented in Table 22**.**

Table 22. Summary of further exploratory secondary endpoints from EMPEROR-Preserved study

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
Time to adjudicated MI (fatal or non-fatal), RS		
Patients with MI, N (%)	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
HR vs placebo (95% CI)	[redacted]	
Nominal p-value	[redacted]	
Time to adjudicated stroke (fatal or non-fatal), RS		
Patients with stroke, N (%)	[redacted]	[redacted]
Ischaemic	[redacted]	[redacted]
Haemorrhagic	[redacted]	[redacted]
Unclassified	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
HR vs placebo (95% CI)	[redacted]	
Nominal p-value	[redacted]	
Patients with fatal stroke	[redacted]	[redacted]
Time to new onset of Afib, as ECG finding or as AE, RS		
Patients without baseline or history of Afib[a], N (%)	[redacted]	[redacted]
Patients with new onset of Afib, N (%)	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
HR vs placebo (95% CI)	[redacted]	

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
Nominal p-value	■	
Blood pressure (mm Hg) changes from baseline to week 52 (mm Hg), RS		
Systolic blood pressure change	-1.8±0.3	-0.6±0.3
Adjusted mean difference (95% CI)	-1.2 (-2.1 to -0.3)	
p-value	0.01	
Diastolic blood pressure change	-0.9±0.2	-0.7±0.2
Adjusted mean difference (95% CI)	-0.2 (-0.7 to 0.3)	
p-value	0.46	
HbA1c (%) change from baseline to week 52, RS patients with diabetes		
Adjusted mean change from baseline	-0.16±0.02	-0.03±0.02
Adjusted mean difference (95% CI)	-0.19 (-0.25 to -0.14)	
p-value	<0.0001	
Body weight (kg) change from baseline to week 52, RS		
Adjusted mean change from baseline	-1.39±0.09	-0.11±0.09
Adjusted mean difference (95% CI)	-1.28 (-1.54 to -1.03)	
p-value	<0.0001	
Haematocrit (%) change from baseline to week 52, RS		
Adjusted mean change from baseline	1.94±0.07	-0.41±0.07
Adjusted mean difference (95% CI)	2.36 (2.17 to 2.54)	
p-value	<0.0001	
NT-proBNP (pg/mL) change from baseline to week 52, RS		
Adjusted median change from baseline (IQR)	-29 (-335 to 263)	-9 (-286 to 322)

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
Adjusted geometric mean ratio	0.95 (0.91 to 0.99)	
p-value	0.01	
Uric acid (mg/dL) change from baseline to week 52, RS		
Adjusted mean change from baseline	-0.90±0.03	-0.10±0.03
Adjusted mean difference (95% CI)	-0.80 (-0.88 to -0.72)	
p-value	<0.0001	
QoL measured by KCCQ at 52 weeks[b], TS		
Change in clinical summary score at 52 weeks	4.51±0.31	3.18±0.31
Adjusted mean change from baseline (95% CI)	1.32 (0.45 to 2.19)	
Nominal p-value	0.0028	
Change in overall summary score at 52 weeks	5.03±0.30	3.66±0.31
Adjusted mean change from baseline (95% CI)	1.37 (0.52 to 2.21)	
Nominal p-value	0.0015	
Change in total symptom score at 52 weeks	5.89±0.34	3.95±0.34
Adjusted mean change from baseline (95% CI)	1.94 (1.01 to 2.88)	
Nominal p-value	<0.0001	

Abbreviations: AE, adverse event; Afib, atrial fibrillation; CI, confidence interval; ECG, electrocardiogram; HbA1c, glycated haemoglobin; HR, hazard ratio; IQR, interquartile range; KCCQ, Kansas City Cardiomyopathy Questionnaire; MI, myocardial infarction; QoL, quality of life; RS, randomised set; SE, standard error; TS, treated set.
Note: Plus-minus values are means ± SE. Estimates of effect size for time to event endpoints (HR, 95% CI) were derived for the randomised set using Cox regression model which included covariates age, baseline eGFR (CKD-EPI)cr, region, baseline diabetes status, sex, baseline EF, and treatment. Continuous endpoints (blood pressure, KCCQ scores) were analysed using mixed model for repeated measures (MMRM).
[a]Based on investigator-reported medical history or baseline ECG.
[b]The clinical summary score on the Kansas City Cardiomyopathy Questionnaire ranges from 0 to 100, with higher scores indicating a better quality of life. Analysis of PRO data with a MMRM was based on the treated set and using on-treatment values only.
Source: Table 11.1.3.1:1, Table 15.2.4.2.1, Sections 11.1.2.7, 11.1.2.8.2 and 11.1.3.4, EMPEROR-Preserved CSR (139); Table S5, Anker et al 2021 (5).

Frequency of MI, stroke and atrial fibrillation were similar between the two treatment groups as based on Cox regression analysis and estimated cumulative incidence

analysis. There was no marked change in blood pressure in the empagliflozin group, with a placebo-corrected adjusted mean change at week 52 from baseline of -1.2 mmHg (95% CI, -2.1 to -0.3) for systolic and -0.2 mmHg (95% CI, -0.7 to 0.3) for diastolic blood pressure.

The change from baseline in health status was assessed by the KCCQ-CSS at week 52. The clinical summary score measures HF symptom frequency, symptom burden, and physical limitations. There was a greater improvement in the clinical summary score from baseline in the empagliflozin group than in the placebo group at Week 52. A similar improvement was also observed for the Kansas City Cardiomyopathy Questionnaire - total symptom score (KCCQ-TSS) and Kansas City Cardiomyopathy Questionnaire - overall summary score (KCCQ-OSS), as well as for the individual domains ‘quality of life’ and ‘social limitation’. There were no relevant differences between the treatment groups with regards to HRQoL as assessed by the EQ-5D-5L questionnaire.

B.2.7 Subgroup analyses

The pre-specified subgroup analyses for the efficacy endpoints of EMPEROR-Preserved were:

- Diabetes at baseline (diabetic, non-diabetic patients) (Appendix E)
- Renal function at baseline (eGFR ≥60 mL/min/1.73 m^2, <60 mL/min/1.73 m^2) (Appendix E)
- Gender
- Race (White, Black, Asian, other)
- Body mass index (BMI) (<30 kg/m^2 and ≥30 kg/m^2)
- Age (<70 years and ≥70 years) (Appendix E)
- SBP at baseline
- History of AF
- HHF in the last 12 months
- NYHA at baseline (II, III/IV)
- Uric acid, in thirds, at baseline
- HF physiology (reflected in baseline EF and level of NT-pro-BNP) (Appendix E)

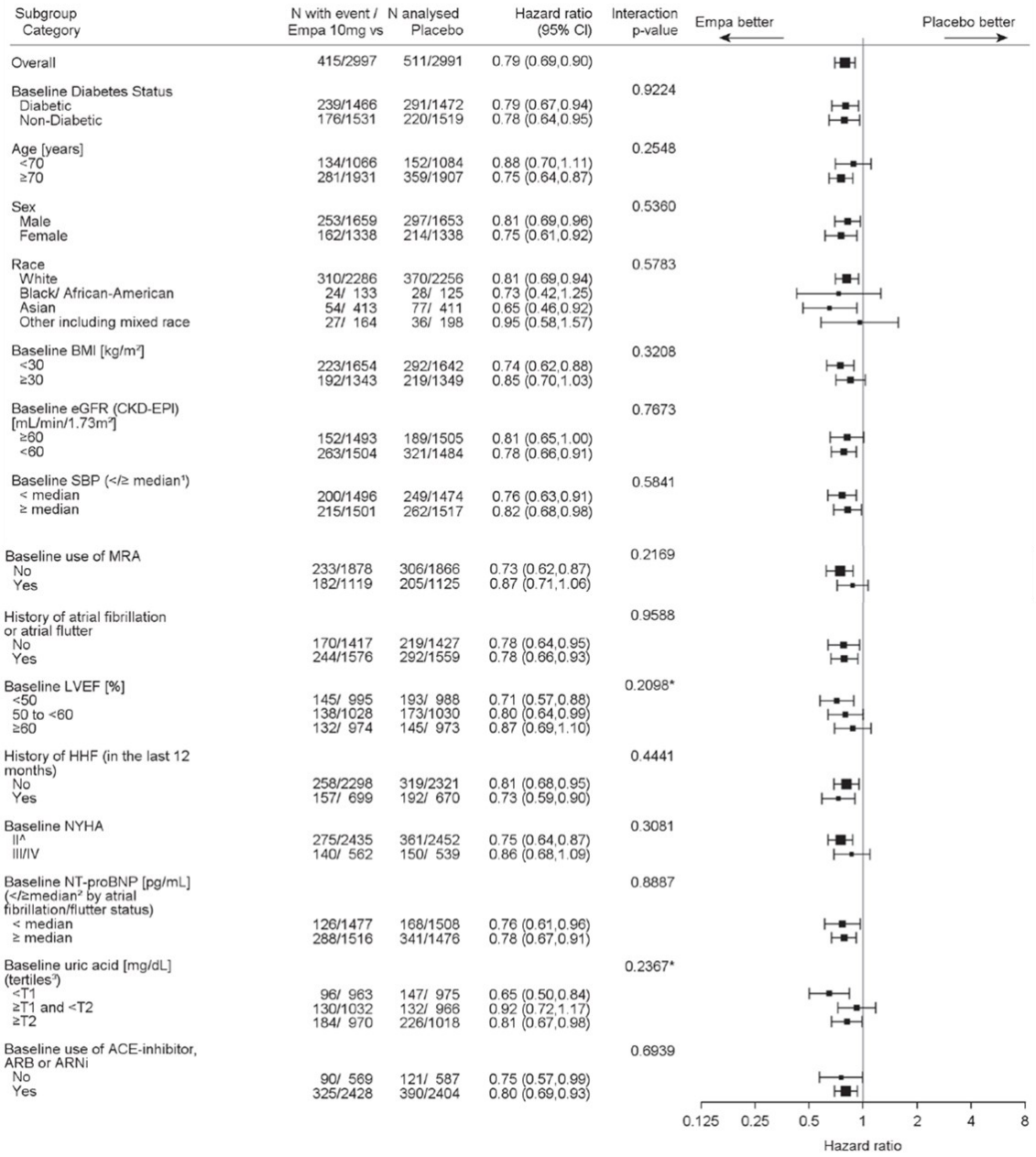
- Baseline use of mineralocorticoid receptor antagonist
- Baseline use of ACEI, ARB, or ARNI at baseline (Appendix E)
- Geographic region (Asia, Europe, Latin America, North America, and other)

A post hoc subgroup analysis according to KCCQ-CSS score at baseline (<62.5, 62.5-83.3, ≥83.3) was also conducted (148). It is to be noted that subgroup analyses were not adjusted for multiple testing. Hence, the subgroup findings were not specifically powered and were regarded as hypothesis generating (Appendix E).

The effect of empagliflozin on the combined risk of CV death or HHF was consistent across the subgroups, with the interaction p values above 0.05 for all analysed subgroups, including by diabetes status and EF (Figure 16). Age (interaction p-value=0.5162), eGFR (interaction p-value=0.6331), or EF (interaction p-value=0.4333) analysed as a continuous variable had no relevant impact on the results of the primary analysis (139). Of note, the consistent effect of empagliflozin in patients with an eGFR lower than 60 mL/min/1·73 m² (HR<1) provides evidence of an important reduction of CV death or HHF in this high-risk subgroup, including patients with eGFR as low as 20 mL/min/1.73 m^2 (5).

For the subgroup analyses by baseline NYHA class although the direction of the treatment effect remained consistent, magnitude of the benefit was smaller in the subgroup with NYHA class III-IV (more severe HF) versus NYHA class II (less severe HF) at baseline. Subgroup analyses by EF at baseline showed that the treatment effect deceased with increase in EF (5). No significant variation in treatment effect was seen across KCCQ scores (Appendix E) or NT-pro-BNP tertiles (5, 148). Furthermore, the point estimate HR remained less than one for each subgroup (Figure 16).

Figure 16. Primary outcome of EMPEROR-Preserved in pre-specified subgroups



Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; BMI, body mass index; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; HHF, hospitalisation for heart failure; eGFR, estimated glomerular filtration rate; EF, left ventricular ejection fraction; MRA, mineralocorticoid receptor antagonist; NT-pro-BNP, N-terminal prohormone of brain natriuretic peptide; NYHA, New York Heart Association; SBP, systolic blood pressure.

Note: The size of the squares for the hazard ratios is proportional to the size of the subgroup. Interaction p values are nominal; the subgroup analyses were not adjusted for multiple testing. The body mass index is the weight in kilograms divided by the square of the height in metres. Race was reported by the patients.

^Patients with NYHA class I are counted in subgroup NYHA class II.

[1]Baseline SBP median: 132 [mmHg].

[2]Baseline NT-proBNP median for patients with history of atrial fibrillation or atrial flutter: 1354 [pg/mL].

[2]Baseline NT-proBNP median for patients without history of atrial fibrillation or atrial flutter: 614 [pg/mL].

[3]Baseline uric acid tertile cut-offs for males are T1: 5.9 T2: 7.5 [mg/dL] [rounded].

[3]Baseline uric acid tertile cut-offs for females are T1: 5.4 T2: 6.9 [mg/dL] [rounded].

*= Trend test

Source: EMPEROR-Preserved CSR, Figure 11.1.1.3: 1 (139).

The use of the ITT population for the cost-effectiveness analysis (CEA) of empagliflozin in chronic HF (EF >40%) is also the most statistically robust approach since EMPEROR-Preserved was not powered to evaluate the treatment effect in any subgroups. With many subgroup analyses carried out without adjusting the overall significance level of the trial, it is unclear if the results represent spurious findings. Given that the results for the subgroups were generally consistent with the confirmatory analyses, only the ITT population was considered in the economic analysis. Results of the clinically relevant pre-specified subgroups can be found in Appendix E.

During TA733, the BMJ EAG explored the use of the Europe subgroup and whether this would be a better representation of UK clinical practice than the ITT population, and thus should be used for the preferred base case ICER. BI would like to proactively address this query as it is also relevant for this appraisal.

The Europe subgroup is unlikely to be a better representation of UK clinical practice than the ITT population. The baseline characteristics and key outcomes for the Europe subgroup were comparable to the ITT population, as reported in Appendix E. Thus, if the Europe subgroup was used as the base case to estimate the ICER, it is likely to be similar to the ITT population.

The use of data from Europe subgroup to assess generalisability is not appropriate and could contribute to existing ethnic inequalities in health (149), contrary to the NICE's Social Value Judgments and the Equality Act 2010 (race is one of the protected characteristics) (14). The Europe subgroup of EMPEROR-Preserved was 99.0% white and therefore not representative of the multi-ethnic UK population, which consists of 86% white, 3.3% black, 7.5% Asian and 3.2% other (139, 150). This difference is even wider in the metropolitan areas of the UK (44.9% white in London) (150). The ITT population of EMPEROR-Preserved, which was 75.9% white, 4.3% black, 13.8% Asian and 4.2% other (5, 139) is more generalisable to the ethnically diverse UK population and is, therefore, the population considered in the economic analysis.

B.2.8 Meta-analysis

Not applicable.

B.2.9 Indirect and mixed treatment comparisons

Not applicable.

B.2.10 Adverse events

Median exposure to study medication was approximately 23 months in both treatment groups, with 84% of patients treated for at least 1 year. Safety was assessed descriptively based on AE, adverse events of special interest (AESI), and specific AE.

A similar overall proportion of patients in the empagliflozin and placebo groups reported at least one AE, most of which were of mild or moderate intensity (Table 23). Proportions of patients experiencing severe AE and AE leading to premature discontinuation of study medication were also similar between the two groups (Table 23).

Table 23. Overall summary of AE in the TS

Category of AE	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Number of patients in the TS, N (%)	2,996 (100.0)	2,989 (100.0)
Patients with any AE	2,574 (85.9)	2,585 (86.5)
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]
Severe	[redacted]	[redacted]
Investigator-defined drug-related AE	[redacted]	[redacted]
AE leading to discontinuation of study medication	571 (19.1)	551 (18.4)
Serious AE	1,436 (47.9)	1,543 (51.6)
Serious AE		
Resulting in death	[redacted]	[redacted]
Life threatening	[redacted]	[redacted]
Persistent or significant disability/incapacity	[redacted]	[redacted]
Requires or prolongs hospitalisation	[redacted]	[redacted]

Category of AE	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Congenital anomaly or birth defect	[redacted]	[redacted]
Other medically important serious event[a]	[redacted]	[redacted]

Abbreviations: AE, adverse event; TS, treated set.
Note: Percentages calculated using total number of patients per treatment as the denominator. A patient may be counted in more than one seriousness criterion.
[a]Other medically important serious event was defined as any important medical event (when based upon appropriate medical judgement) which might jeopardise the patient and might require medical or surgical intervention to prevent one of the other serious outcomes included in the definition of serious adverse events shown in the table above. Examples of such events could be intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias, or convulsions that do not result in hospitalisation or development of dependency or abuse.
Source: EMPEROR-Preserved CSR, Table 15.3.1.1 (139).

The overall frequency of serious AE (SAE) was lower in the empagliflozin group than in the placebo group, consistent with the efficacy analyses of all-cause hospitalisations (Table 24). The most frequent SAE were cardiac failure, atrial fibrillation, pneumonia and acute kidney injury. All other SAE were reported in less than 3.0% of patients per treatment group.

Table 24. Serious AE with frequency >1% -exposure adjusted, in the TS

MedDRA SoC MedDRA PT	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Number of patients	2,996 (100.0)	2,989 (100.0)
Total with SAE	1,436 (47.9)	1,543 (51.6)
Cardiac disorders	[redacted]	[redacted]
Cardiac failure	[redacted]	[redacted]
Atrial fibrillation	[redacted]	[redacted]
Cardiac failure congestive	[redacted]	[redacted]
Acute myocardial infarction	[redacted]	[redacted]
Myocardial infarction	[redacted]	[redacted]
Coronary artery disease	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
COVID-19	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]
Renal and urinary disorders	[redacted]	[redacted]

MedDRA SoC MedDRA PT	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Acute kidney injury	[redacted]	[redacted]
Renal impairment	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]
Ischaemic stroke	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified	[redacted]	[redacted]
Basal cell carcinoma	[redacted]	[redacted]
Vascular disorders	[redacted]	[redacted]
Hypertensive crisis	[redacted]	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]	[redacted]
Chronic obstructive pulmonary disease	[redacted]	[redacted]
General disorders & administration site conditions	[redacted]	[redacted]
Death[a]	[redacted]	[redacted]
With investigator-defined drug-related SAE	[redacted]	[redacted]

Abbreviations: AE, adverse event; MedDRA, Medical dictionary for regulatory activities; SoC, system organ class; MedDRA PT, Medical dictionary for regulatory activities preferred term; SAE, serious adverse event; TS, treated set.
[a]Deaths not attributed to another PT by the investigator. The frequencies of patients with fatal AE were balanced between treatment groups.
Source: EMPEROR-Preserved CSR Table 12.2:1 (139).

Adverse events of special interest were pre-specified in the protocol as acute renal failure, hepatic injury, decreased renal function, ketoacidosis, and AE leading to lower limb amputation. Overall frequencies of AESI were comparable in the empagliflozin and placebo groups (Table 25).

Specific AE were defined as urinary and genital tract infections, volume depletion and hypotension, hypoglycaemic events, bone fractures and urinary tract malignancies. As known for the drug class, urinary tract infections were more common in the empagliflozin group. Uncomplicated genital tract infections also occurred more often with empagliflozin than with placebo, while complicated genital infections or those

leading to treatment discontinuation had similar frequency in both groups. There was a numerical but not clinically meaningful increase in volume depletion and hypotension with empagliflozin relative to placebo, including events that were reported as SAE or that led to treatment discontinuation. No increase in confirmed hypoglycaemic events was detected for patients with or without DM, and no severe hypoglycaemic events were reported in patients without DM. The frequencies of the remaining types of specific AE were similar between the groups (Table 25).

Table 25. Summary of AESI and specific AE, TS

Category of AESI and specific AE	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Number of patients	2,996 (100.0)	2,989 (100.0)
AESI		
Acute renal failure	363 (12.1)	384 (12.8)
Serious	[redacted]	[redacted]
Leading to discontinuation	[redacted]	[redacted]
Hepatic injury	115 (3.8)	155 (5.2)
Serious	[redacted]	[redacted]
Leading to discontinuation	[redacted]	[redacted]
Up to 30 days after treatment discontinuation	[redacted]	[redacted]
Ketoacidosis (broad[a])	[redacted]	[redacted]
Ketoacidosis (narrow[a])	4 (0.1)	5 (0.2)
AE leading to LLA up to trial completion (investigator-defined)	16 (0.5)	23 (0.8)
Specific AE		
Urinary tract infection	297 (9.9)	243 (8.1)
Complicated	57 (1.9)	45 (1.5)
Leading to discontinuation	[redacted]	[redacted]
Genital infection	67 (2.2)	22 (0.7)
Complicated	8 (0.3)	8 (0.3)

Category of AESI and specific AE	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Leading to discontinuation	[redacted]	[redacted]
Volume depletion	[redacted]	[redacted]
Hypotension (a subset of volume depletion)	311 (10.4)	257 (8.6)
Serious	[redacted]	[redacted]
Leading to discontinuation	[redacted]	[redacted]
Symptomatic hypotension (investigator-defined)	197 (6.6)	156 (5.2)
Confirmed hypoglycaemic events[b]	73 (2.4)	78 (2.6)
In patients with T1DM[c]	[redacted]	[redacted]
In patients with T2DM[c]	[redacted]	[redacted]
In patients with pre-diabetes[c]	[redacted]	[redacted]
In patients without diabetes or pre-diabetes[c]	[redacted]	[redacted]
Bone fracture	134 (4.5)	126 (4.2)
Serious	[redacted]	[redacted]
Leading to discontinuation	[redacted]	[redacted]
Up to trial completion	[redacted]	[redacted]
Urinary tract malignancy up to trial completion	[redacted]	[redacted]

Abbreviations: AE, adverse event; AESI, adverse event of special interest; LLA, lower limb amputation; T2DM, type 2 diabetes mellitus; TS, treated set.
[a]Ketoacidosis was investigated using both broad and narrow Boehringer Ingelheim customised MedDRA queries (BIcMQs)
[b]Hypoglycaemic AE with a plasma glucose value of ≤70 mg/dL or where assistance was required
[c]Patients with events/patients in subgroup (%)
Source: EMPEROR-Preserved CSR Table 12.1.3:1 (139).

B.2.11 Ongoing studies

There are no ongoing studies of empagliflozin relevant for this appraisal.

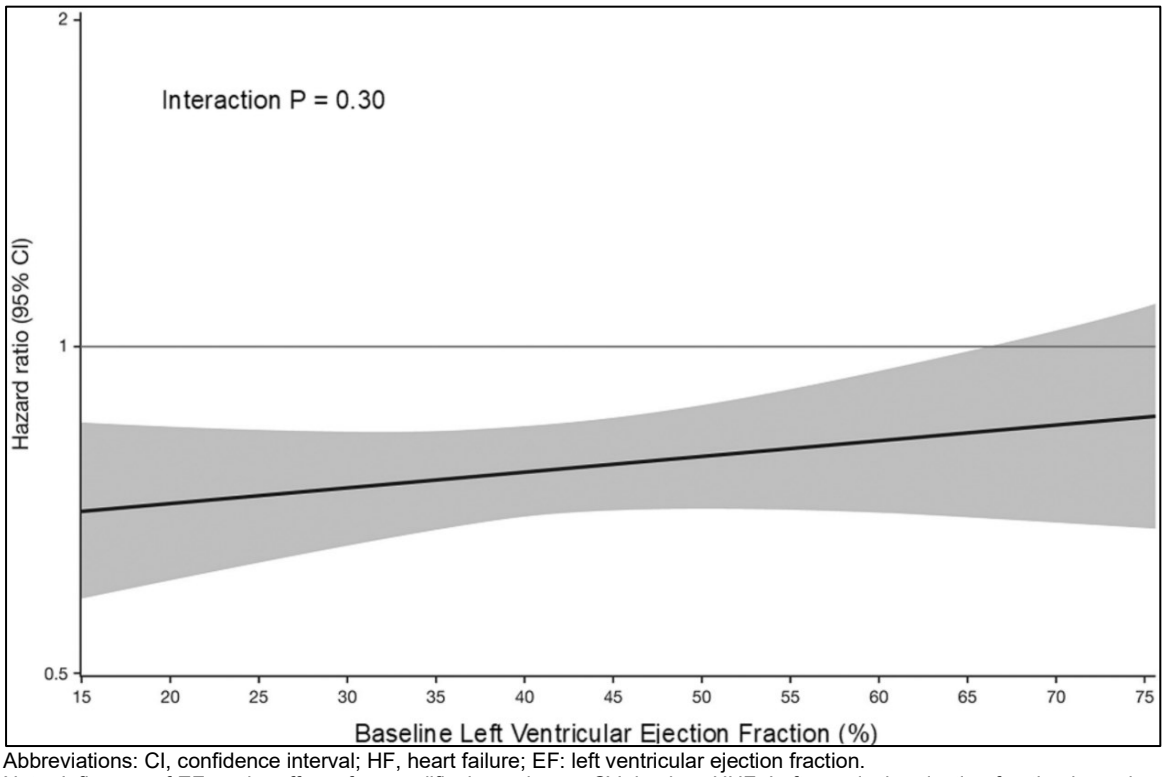
B.2.11.1 Effect of empagliflozin in patients with HF across spectrum of EF

Empagliflozin has demonstrated benefit across the EF spectrum. A combined HF analysis stratified by EF was performed on both the EMPEROR-Reduced and

EMPEROR-Preserved trials (9718 patients; 4860 empagliflozin and 4858 placebo) (120). Both trials had pre-specified subgroup analysis based on EF but a more granular stratification was done post hoc in this analysis. The patients were divided into six groups categorised by EF: <25%, 25–34%, 35–44%, 45–54%, 55–64%, and ≥65% in order to have a better understanding of the effect of empagliflozin across the full range of chronic HF patients (120).

The combined analysis demonstrated heterogeneity in the baseline characteristics across the EF groups. Patients with higher EF were mostly older females with greater comorbidity burden (120). In terms of efficacy, empagliflozin consistently reduced the primary composite endpoint of time to CV death or first HHF for HF patients across a broad spectrum of EF (**Error! Reference source not found.**). The benefit is mostly driven by the effect of empagliflozin on time to first HHF and total (first and recurrent) HHF in patients with EF ranging from <25% to <65%. The pattern of effects of empagliflozin across the EF intervals were similar for both genders. The combined analysis also showed that the improvement in patients' quality of life as measured by KCCQ-CSS had similar patterns as the HF outcomes. In conclusion, irrespective of differences in the demographic characteristics, the benefit of empagliflozin on HF outcomes and health status were consistently observed across the broad spectrum of EF.

Figure 17. Effect of empagliflozin on primary composite endpoint in patients with HF across spectrum of EF



Abbreviations: CI, confidence interval; HF, heart failure; EF: left ventricular ejection fraction.
Note: Influence of EF on the effect of empagliflozin on time to CV death or HHF. Left ventricular ejection fraction is analysed as a continuous variable, based on the assumption that the relationship is linear. Analysis of the influence of EF using cubic splines yielded a pattern similar to that observed in our six subgroups, showing a consistent risk reduction in patients with an ejection fraction <65% and an attenuated effect at the highest ejection fractions.
Shaded areas represent 95% confidence intervals.
Reference: Butler et. al. (2022) (120).

B.2.12 Innovation

As stated in **Error! Reference source not found.**, HF affects just under 1 million people in the UK, of which up to 50% are estimated to have chronic HF (EF >40%) (9, 151, 152). Currently NICE guideline does not recommend any specific therapy for the treatment of chronic HF (EF >40%) and the treatment focuses on the management of comorbidities (3).

Empagliflozin, either alone or in combination with all appropriate therapy for HF and comorbid conditions offers to be an important advancement in the treatment of chronic HF (EF >40%):

- It significantly reduces the risk of CV death or HHF while significantly improving renal outcomes and quality of life in a population with broad spectrum of

severity of chronic HF (EF >40%) regardless of age, gender, presence or absence of diabetes or EF at baseline (5).

- As a fixed dose, once-per-day, orally administered medication, empagliflozin is simple for physicians to initiate and for patients to adhere to, saving NHS professionals' time that would otherwise be spent on dose titration or on training patients to self-administer.
- Substantial reduction in HHF seen with empagliflozin combined with simplicity of initiation suggests that its adoption in primary care could support efficiency improvements in the allocation of NHS resources by releasing capacity in secondary care. A recently published report by NICE on implementation of NG106 noted that patients with HF often have comorbid diabetes and CKD that require visits to additional specialist clinics (153). SGLT2 inhibitors, like empagliflozin, offer an opportunity to promote a more holistic approach to treatment of adults with T2DM (153). Empagliflozin is already indicated in T2DM and is recommended in chronic HF (EF ≤40%); therefore, empagliflozin could further support the objective of a holistic treatment for adults with chronic HF (across the full spectrum of EF) and comorbid T2DM, as well as potentially reducing polypharmacy for comorbid patients.

B.2.13 Interpretation of clinical effectiveness and safety evidence

In the EMPEROR-Preserved trial, treatment with empagliflozin 10 mg once daily as an add-on to SoC in patients with chronic HF (EF >40%) demonstrated superiority compared to placebo for the primary endpoint, time to the first occurrence of adjudicated CV death or adjudicated HHF. The superiority over placebo was also demonstrated for key secondary endpoints, occurrence of adjudicated HHF (first and recurrent) and eGFR (CKD-EPI)cr slope of change from baseline.

Treatment with empagliflozin as an add-on to SoC led to a clinically and statistically significant reduction in risk of CV death or HHF by 21% compared with placebo added to SoC. During the trial period, the number of patients who needed to be treated with empagliflozin to prevent one primary event was 31 (95% CI, 20 to 69). The treatment effect of empagliflozin became apparent shortly after randomisation and was maintained throughout the trial. The results of all sensitivity analyses were consistent

with the results of the primary analysis (i.e., the HR was similar). The results were also consistent across the pre-defined subgroups stratified by demographics, baseline characteristics, and baseline medications as indicated by the point estimate HR for time to CV death or HHF for each subgroup being below the no-effect value of 1.

The risk of first occurrence of all-cause hospitalisation was reduced in the empagliflozin group compared with the placebo group. Furthermore, fewer patients receiving empagliflozin were reported with all-cause or CV mortality, although the treatment effect was not significant on a nominal level. Most of the CV deaths were classified as sudden cardiac death or HF death, as expected in this population. A recent analysis of data from EMPEROR-Preserved trial showed that empagliflozin significantly reduced worsening of HF event (CV death, HHF, or an emergency or urgent HF visit requiring intravenous treatment) 18 days after randomisation and maintained significance thereafter (124).

The decline in renal function, evaluated based on change in eGFR slope from baseline, was significantly slower in the empagliflozin group, with an estimated difference in slope of about 1.36 mL/min/1.73 m^2 per year vs. placebo. The risk of composite renal endpoint (chronic dialysis, renal transplant, or sustained reduction in eGFR) was similar between the empagliflozin and the placebo treatment group. The findings from the EMPEROR-Preserved study therefore have important clinical implications for the holistic treatment of indications comprising the “CRS”.

A higher proportion of patients in the empagliflozin group than in the placebo group showed a clinically meaningful improvement in KCCQ-CSS after 52 weeks of treatment of at least five points from baseline. Consistently, a lower proportion of patients in the empagliflozin group than in the placebo group showed deterioration. The favourable effect of empagliflozin was mainly driven by the domains of symptom frequency and symptom burden, and a positive trend in favour of empagliflozin was observed in domain of physical limitations. Supportive analyses of KCCQ-OSS and KCCQ-TSS were consistent with these findings.

Empagliflozin was well tolerated in chronic HF (EF >40%) patients with or without T2DM. AE reported in the trial were consistent with the known safety profile of

empagliflozin. As expected for the SGLT2 drug class, urinary tract infection and uncomplicated genital infections were more common in the empagliflozin group. The frequency of hypoglycaemia and bone fracture did not differ between the two groups, even though these AE have been associated with the use of SGLT2 inhibitors in trials with T2DM patients (154). The proportion of patients experiencing SAE was lower in the empagliflozin than in the placebo group, consistent with the efficacy analyses of all-cause hospitalisations. Safety concerns that have been seen with other drugs for HF (e.g., volume depletion, renal dysfunction, bradycardia, and hyperkalaemia) were not evident with empagliflozin in EMPEROR-Preserved.

The benefits of empagliflozin have been demonstrated across a broad range of EF in chronic HF patients (120). This combined HF analysis stratified by EF demonstrated that EF does not impact the treatment effect of empagliflozin on the primary endpoint (CV death or first HHF) and the key secondary endpoint (total HHF) (120).

In conclusion, the evidence submitted in the current company position combined with the evidence assessed in TA773 demonstrates empagliflozin to be associated with a clinically meaningful reduction in risk of CV death or HHF and a slower progressive decline of renal function in patients with chronic HF, regardless of EF. Empagliflozin is the first therapy demonstrating benefit across a broad range of EF, facilitating improvement in patient outcomes regardless of any delay to echocardiography to determine EF and appropriate treatment Thus, its broad recommendation in NG106 for all chronic HF patients will improve outcomes and benefit the full patient population treated by the NHS.

B.3 Cost-effectiveness

- A Markov model with health states defined by KCCQ-CSS quartiles was developed to estimate the lifetime costs and outcomes of patients with EF >40%. The model structure was identical to the model submitted for TA773 ≤40%; which the EAG and committee deemed appropriate.
- The results of the EF ≤40% model appraised in the TA773 NICE submission indicated that empagliflozin was cost-effective with an incremental cost-effectiveness ratio (ICER) of £4,717 per quality-adjusted life year (QALY) based on the company's revised base case post the technical engagement.
- In the EF >40% base case analysis, which evaluated the cost-effectiveness of empagliflozin as an add-on to SoC compared to SoC alone, empagliflozin was estimated to increase life years and QALYs by [redacted] and [redacted] per patient, respectively, and to reduce HHF by [redacted] events per 100 patient-years. The ICER of £14,429 per QALY indicated that empagliflozin is highly cost-effective as an addition to the SoC in symptomatic chronic HF and EF >40%.
- Deterministic sensitivity analyses for the EF >40% model demonstrated that the cost-effectiveness results were robust with respect to variation in individual model parameters. The treatment effect associated with HHF was identified as the most influential driver of model results. The ICER however remained below £22,000 per QALY across all parameter variations.
- Probabilistic sensitivity analyses indicated [redacted]% and [redacted]% probability of empagliflozin being cost-effective at the willingness-to-pay thresholds of £20,000 per QALY and £30,000 per QALY, respectively.
- Scenario analyses, including scenarios suggested by the EAG during technical engagement in TA773, show that there is certainty in the cost-effectiveness results for empagliflozin. All ICERs were below £30k/QALY except the scenario where transition probabilities between the KCCQ-CSS quartiles for the treatment arm [redacted] was set to the transition probabilities for the SoC arm after 8 months (ICER: £32,482/QALY). This means that the ICER is not significantly affected by assumptions including the choice of parametric distribution for mortality or treatment discontinuation, utility age adjustment, cost of non-CV death, increasing risk of HHF, or the mortality benefit of KCCQ-CSS membership.
- Empagliflozin therefore represents a highly cost-effective use of NHS resources in the treatment of symptomatic patients with chronic HF (EF >40%) and across the full phenotype spectrum of chronic HF.

B.3.1 Published cost-effectiveness studies

A SLR was conducted to identify existing economic evaluations for the treatment of chronic HF patients with EF >40%. The same SLR protocol was used for TA733, but only studies for patients with an EF<40% were included in the PRISMA diagram. Full details of the process and methods used are described in Appendix G.

In summary, following the abstract and full-text screening process, no relevant studies were identified for inclusion in the SLR and therefore no quality assessment was conducted.

B.3.2 Economic analysis

The SLR of cost-effectiveness studies, described in Appendix G, did not identify any suitable economic evaluations in the chronic HF with EF >40% population. Therefore, the cost-effectiveness model for economic evaluation of empagliflozin + SoC for chronic HF patients with EF >40% builds on the modelling approach previously accepted by the NICE committee for empagliflozin + SoC for patients with EF ≤40% (TA773) (1) and the economic model submitted for dapagliflozin in HFrEF to NICE (155), which were developed using Microsoft Excel® (Office 365, version 2008) with Visual Basic for Applications (VBA) functionality.

The selection of the model approach and structure considered the current treatment landscape for chronic HF with EF >40%, recent relevant health technology assessment (HTA) submissions to NICE for the treatment of chronic HF, the relevant questions in decision-making, and the EMPEROR-Preserved trial data, including adequately capturing the trends observed in the trial primary outcomes (HHF and CV death). Furthermore, the empagliflozin + SoC for patients with the HFrEF model submission (1) was appraised by NICE and agreed suitable for modelling chronic HF, hence that model has been used as the basis for the development of the cost-effectiveness model for EF >40%.

The economic model for dapagliflozin in HFrEF submitted to NICE as part of TA679 was found to adequately reflect the variation in risk with disease severity through the use of time-updated KCCQ-TSS covariates in risk equations for all-cause mortality, CV mortality and HHF (155). The NICE evaluation committee for TA679 concluded

that the KCCQ tool is a reasonable way to classify disease severity and is appropriate for decision-making (155).

A similar approach was therefore adopted for modelling cost-effectiveness of empagliflozin in HF patients with EF >40%, that is, with KCCQ-CSS rather than KCCQ-TSS-defined health states as explained in more detail in Section B.3.2.2**Error! Reference source not found.**.

B.3.2.1 Patient population

The patient population considered in the CEA is adults with symptomatic chronic HF with EF >40% in accordance with the anticipated MA of empagliflozin and the decision problem considered in this submission. Empagliflozin is indicated in adults for the treatment of symptomatic chronic HF with EF >40%. The population is also reflective of the ITT population of the EMPEROR-Preserved trial (124).

In accordance with the trial inclusion criteria, the modelled cohort comprised adults with chronic HF with EF >40%. The KCCQ-CSS distribution of patients in the ITT population at baseline across quartiles was used to inform the initial distribution of patients across alive health states at the start of the model and influenced the rates of all-cause death, CV death, and HHF (Table 26). The modelled cohorts in the empagliflozin + SoC and placebo + SoC arms were assigned the same baseline characteristics.

Table 26. Mean patient characteristics of the modelled cohort at model entry based on EMPEROR-Preserved trial, ITT population

Baseline characteristic	ITT population	SE
Demographics		
Age (years)	71.89	0.12
Age (≥65 years)	80%	0.01
Sex: Male	55%	0.01
Region		
Asia	11.5%	0.00
Europe	44.9%	0.01
Latin America	25.3%	0.01
North America	12.0%	0.00
Other	6.3%	0.00
KCCQ-CSS		

Baseline characteristic	ITT population	SE
KCCQ-CSS 0 to <55 (Quartile 1)	[redacted]%	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]%	[redacted]
KCCQ-CSS: 75 to <90 (Quartile 4)	[redacted]%	[redacted]
KCCQ-CSS: 90 to 100 (Quartile 4)	[redacted]%	[redacted]
NYHA class		
Baseline NYHA II	81.5%	0.01
Baseline NYHA III	18.1%	0.00
Baseline NYHA IV	0.3%	0.00
Treatment use at baseline		
ACEI	40.2%	0.01
ARB	38.7%	0.01
ARNI	2.2%	0.00
MRA	37.5%	0.01
BB	86.3%	0.00
Loop diuretics	67.7%	0.00
Medical history		
Ischaemic cause of HF	35.4%	0.01
Prior atrial fibrillation or flutter	52.4%	0.01

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; BB, beta-blocker; CSS, clinical summary score; HF, heart failure; ITT, intent to treat; KCCQ, Kansas City Cardiomyopathy Questionnaire; MRA, mineralocorticoid receptor antagonist; NYHA, New York Heart Association; SE, standard error.

B.3.2.2 Model structure

The model uses a Markov cohort state transition approach and describes the clinical course of HF with EF >40% using five discrete health states defined by quartiles of the baseline distribution of KCCQ-CSS in the combined empagliflozin and placebo treatment groups (KCCQ-CSS quartiles 1 to 4 corresponding to KCCQ-CSS scores of zero to <[redacted], [redacted] to <[redacted], [redacted] to <[redacted], and [redacted] to 100, respectively, with higher score corresponding to a better health status), and death, with health state-specific costs and utilities (Figure 18). The use of quarters vs tertiles was also explored for categorising KCCQ-TSS. Quartiles were found to provide a better fit to the observed data than tertiles while still retaining adequate patient numbers in each subgroup to permit statistically robust analysis and providing sufficient granularity in predicting

patient outcomes. Evenly spaced quarters were also rejected (i.e., 0-25, 26-50, 51-75, 76-100) as they did not contain adequate patient numbers in each group for a robust analysis. Similarly, health states defined by KCCQ-CSS tertiles of the baseline distribution was explored; however, the analysis of transition probabilities showed less differentiation between the treatment groups and over time, suggesting loss of sensitivity to differences.

The patient cohort entered the model according to the baseline distribution of KCCQ-CSS quartiles. From this state, patients could transition to a higher (i.e., regress/lower disease burden) or lower (i.e., progress/higher disease burden) KCCQ-CSS quartile, remain in the same state, or die. In each of the states, patients could experience an AE or HHF. Transitions between the health states occurred in 1-month cycles, and half-cycle correction was applied.

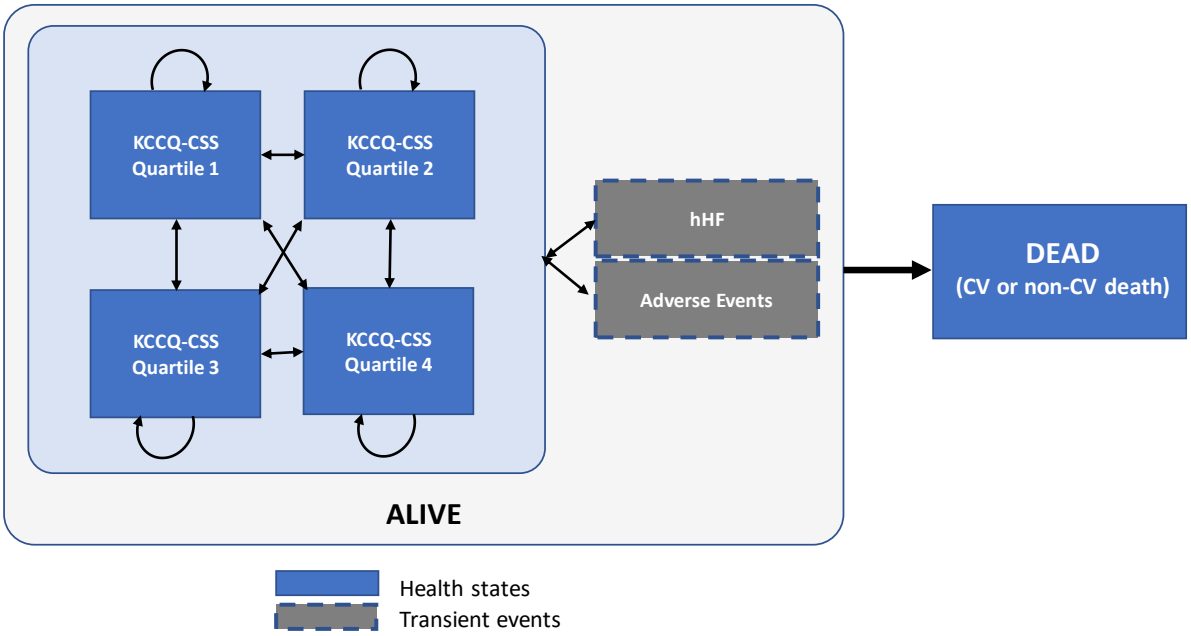
KCCQ score is an established disease-specific measure of health status derived from a 23-item self-administered questionnaire that quantifies a patient's perception of their health status (156-158). The KCCQ score has been shown to be valid, reliable, and sensitive to clinical changes, with low KCCQ score being an independent predictor of poor prognosis in HF (158-160). The test-retest reliability of the KCCQ with respect to HF was established in an outpatient cohort of 39 stable patients where a non-significant difference was observed in the overall summary scores (OSS) at baseline and a 3-month visit (66.2 *versus* 64.1; p=0.36) (156). The KCCQ has also been validated in patients with EF ≥45% and proven to be both reproducible and sensitive to important changes in a patient's health status (160, 161). The KCCQ domains quantify the patient's perception of their health status including HF symptoms (frequency and burden), physical and social limitation and quality of life (QoL). Different elements of the KCCQ are captured using various summary scores: (i) total symptom score (TSS) which includes the symptom domains (frequency and severity); (ii) clinical summary score (CSS) which includes the TSS and the physical limitation domain; and (iii) OSS which includes the CSS, QoL, and social limitation domains (162). Scores range from 0 to 100 with higher scores indicating better health status, lower symptom burden and better health-related quality of life (HRQoL) (162, 163).

KCCQ-CSS, unlike KCCQ-TSS, was an exploratory endpoint in EMPEROR-Preserved. Model health states based on KCCQ score are a more patient-centric assessment of disease burden in CHF with EF >40% rather than administrative measures such as hospitalisations. They also enable the impact of disease severity to be captured in the health state utility values and in the risk of events, and therefore allow the impact of disease severity to be more accurately modelled compared with the two-state Markov modelling approach previously used in NICE appraisals (TA267 and TA388 for HFrEF) (164, 165). The KCCQ score is also considered a better measure of disease severity compared with the commonly used NYHA functional classification (160). The NYHA classification is a physician's interpretation of patient's symptoms, and can often lead to biased assessments, whereas the KCCQ is a more robust, patient-centric questionnaire and is likely to be used more commonly in clinical trials given the sound KCCQ psychometric properties (31, 161-163).

The model captured the occurrence of first and subsequent HHF and treatment-related AE as transient events. Transition to the death state was modelled using parametric survival equations for CV mortality and all-cause mortality. All model equations were derived using the EMPEROR-Preserved trial data with KCCQ-CSS health states as time-varying predictors.

At the end of each cycle, patients transitioned from the alive health states to death based on the estimated all-cause death rate. The CV death equation was used to estimate the proportion of patients who die from CV causes. The difference between the all-cause death rate and the CV death rate represented the non-CV death rate. Patients could discontinue treatment with empagliflozin at any cycle. After discontinuation, patients received SoC treatment until death or till the end of the model time horizon. Patients who discontinued treatment with empagliflozin experienced thereafter the same event rates and health state transition probabilities as patients receiving placebo. The transition probability matrix for transitions between KCCQ-CSS quartiles was then applied to the remaining patients in the alive health states to calculate the health state distribution in the next cycle. Monthly transition probabilities were derived using longitudinal measurements of health status defined by KCCQ-CSS in the trial, as described in Section B.3.3.

Figure 18. Model schematic



Abbreviations: CSS, clinical summary score; CV, cardiovascular; HHF, hospitalisation for heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire.
Q1, Quartile 1: 0 to < [redacted]; Q2, Quartile 2: [redacted] to < [redacted]; Q3, Quartile 3: [redacted] to < [redacted]; Q4, Quartile 4: [redacted] to [redacted].

Costs included direct medical costs for treatment acquisition, clinical event management and disease management. Utilities were accrued based on time spent in each KCCQ-CSS quartile, adjusted for disutilities associated with HHF and AE. The model tabulates the cumulative number of clinical events experienced by the cohort, event rates per 100 person-years (PY), life years (LY), QALYs, costs, and ICER. A 3.5% annual discount rate was applied to costs and health outcomes

A Markov multi-state model structure based on disease severity was considered the most appropriate because it allowed explicit modelling of the relationship between disease progression and clinical outcomes through the specification of different rates of HHF and CV death depending on the KCCQ-CSS quartile, and enabled inclusion of short-term as well as long-term health benefits in the rate of health state transitions. This structure addresses concerns identified by NICE surrounding the two-state Markov models of chronic HF used in previous NICE submissions (TA388 and TA267) (164, 165) (since treatment effect can impact the rate of transitions between different health states) and follows the well-received approach in TA679 (155). Furthermore, the model allows for explicit modelling of the relationship between disease progression and clinical outcomes (e.g., different rates of HHF and CV death can be specified);

and patients' HRQoL and resource use over time can be tracked as they move between different health states. It was also considered simpler and more efficient than a patient-level simulation approach (used in NICE TA388) (165) requiring less computational time while still adequately capturing heterogeneity across patients with CHF with EF >40% through a tractable number of mutually exclusive and exhaustive health states whose occupancy is determined by patient characteristics related to individual's KCCQ-CSS state.

The chosen model structure is aligned with the clinical care pathway described in Section **Error! Reference source not found.** and reflects the anticipated positioning of empagliflozin as an add-on to usual therapy for CHF with EF >40%. Currently NICE guidelines do not specify specific therapy options for the treatment of chronic HF with EF >40% (3). However, even though the 2021 ESC guideline has recommended use of ACEI/ARB/ARNI, beta-blockers, MRAs and diuretics for management of patients with chronic HF and EF 40-50%, (20) the recommendations are low and are not well established by RCT evidence (3). Therefore, due to the lack of established efficacious treatments for chronic HF patients with EF >40%, the evidence supports the positioning and treatment pathway for empagliflozin. The main features of the cost-effectiveness model are summarised in Table 27.

Table 27. Features of the current economic analysis

Factor	Chosen values	Justification
Model structure	A Markov model with 5 states, including death, using KCCQ-CSS quartiles to capture disease severity and progression	A time-varying covariate by KCCQ-CSS quartile in risk/survival /utility equations allows modelling of the relationship between disease severity and outcomes (HHF, CV death, all-cause death)
Comparators	Standard care without empagliflozin. Standard care is defined as: • Usual therapies, comprised of the following: • ARNI (sacubitril / valsartan) • ACEI (captopril, enalapril, lisinopril, ramipril, trandolapril) • BB (bisoprolol, carvedilol, metoprolol, nebivolol) • ARB (candesartan, valsartan, losartan) • MRA (eplerenone)	See the decision problem **Error! Reference source not found.**.
Time horizon	Lifetime	HF is a chronic disease, with costs and effects of treatment accumulating over lifetime
Treatment waning effect?	No	No evidence of treatment waning in EMPEROR-Preserved
Source of utilities	EMEPEROR-Preserved trial	As per NICE manual for health technology evaluations (166)
Source of costs	NHS and PSS price sources, and literature for other cost inputs	As per NICE manual for health technology evaluations (166)

Factor	Chosen values	Justification
Perspective on health effects	Direct health effects	As per NICE manual for health technology evaluations (166)
Perspective on costs	NHS and PSS	As per NICE manual for health technology evaluations (166)
Discounting	3.5%	As per NICE manual for health technology evaluations (166)
Cycle length	One month, with half-cycle correction	The shortest practical cycle length, given the frequency of trial data collection and a lifetime horizon

Abbreviations: ACEI, angiotensin-converting enzyme inhibitors; ARB, angiotensin receptor blockers; ARNI, angiotensin receptor-neprilysin inhibitors; BB, beta-blockers; CV, cardiovascular; HF, heart failure; HHF, hospitalisation for heart failure; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire – clinical summary score; MRA, mineralocorticoid receptor antagonists; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PSS, Personal and Social Services.

B.3.2.3 Intervention technology and comparators

The base case analysis in this submission compares empagliflozin, as an add-on to background therapy for symptomatic relief and comorbidity management of HF in patients with an EF >40%.

This comparison is in line with the proposed positioning of empagliflozin in Section B.1.3.2.3 and current NICE guidance NG106 (3).

B.3.3 Clinical parameters and variables

The clinical inputs for the economic model of empagliflozin were derived from analyses of patient-level data for the ITT population of EMPEROR-Preserved and consisted of the following:

- Transition matrices describing the probability of moving to each of the KCCQ-CSS quartile health states over time with empagliflozin + SoC treatment and SoC alone, given the current $P_{x,y}$ (Figure 16)
- Projected survival distributions for all-cause and CV mortality as a function of current health state and treatment implemented as parametric survival equations with treatment and time-varying health state indicators
- Rate of HHF over time as a function of current health state and treatment implemented as a repeated measures Poisson regression with treatment and time-varying health state indicators
- Change in utility associated with HHF and AE derived from mixed-effects regression analyses relating the occurrence and timing of these events to changes in utilities
- Treatment discontinuation with empagliflozin implemented as a parametric survival equation

The choice of predictors in the risk equations was guided with the aim of preserving alignment between the observed and the predicted outcomes in the cohort Markov model. Following extensive baseline variable testing, the treatment and time-varying health state indicators were retained as predictors in the risk equations. Since the

treatment effect of empagliflozin was found to be consistent across subgroups defined by age, gender, body mass index, race, presence or absence of diabetes, baseline eGFR, and prior therapies (ARNI or MRA) to that observed in the ITT population (Section B.2.7 and Appendix E), only the ITT population was considered in the economic analysis.

B.3.3.1 Baseline characteristics

The risk equations for all-cause mortality, CV mortality, HHF, and treatment discontinuation underlying the model predict the expected outcomes in a population where individuals are homogeneous in terms of the set of predictors in the equations. A cohort Markov structure is inherently designed for modelling homogeneous populations where a profile defined by the mean of the patient characteristics closely resembles the true study population (167, 168). The baseline characteristics of the modelled cohort, representative of the ITT EMPEROR-Preserved population, are shown in Table 26. While KCCQ-CSS is included in risk equations directly as a time-varying predictor, the association of other baseline variables with outcomes is captured indirectly, through their correlation with KCCQ-CSS. Thus, the Markov cohort is fully defined with regards to the average characteristics of the corresponding ITT population in EMPEROR-Preserved and the observed outcomes in this population are directly comparable with the predicted results. This is achieved at the expense of loss of flexibility, however, since the model can only be run in the population for which equations have been derived.

B.3.3.2 Health state transition probabilities

Treatment effect was found to be not statistically significant with respect to change in KCCQ-CSS from baseline for baseline for CV mortality and AC mortality (detailed in Section B.3.3.1). Treatment effect was incorporated in the cost-effectiveness model (CEM) through KCCQ-CSS quartile transition probabilities for CV mortality, which is retained in the base case for CV mortality even though it was not statistically significant. The rationale for including treatment effect for CV mortality is that it is supported by clinical plausibility that empagliflozin could reduce CV mortality. However, for all-cause mortality, the treatment effect was removed from the base case

because it was not statistically significant and it is not clinically plausible to assume increased all-cause mortality for empagliflozin-treated patients.

As the CEM is a Markov model that tracks transitions between health states, it needs probability matrices as inputs, with rows representing current states, columns representing future states, and cells with the transition probabilities as elements. These matrices were constructed directly from the EMPEROR-Preserved trial data by stratifying patients by prior period health state (KCCQ-CSS defined) and calculating their observed proportions in the next period health state.

Treatment specific changes in health status were captured in the model through KCCQ-CSS quartile transition probabilities, with treatment specific transition matrices derived from the analysis of KCCQ-CSS data collected in EMPEROR-Preserved at baseline and at weeks 12, 32, and 52. Analysis of transition probabilities consisted of deriving the proportion of ITT population in each KCCQ-CSS health state at a given time stratified by the previous health state based on longitudinal measurements of KCCQ-CSS health status in the trial with imputation by last observation carried forward strategy for missing visits while patients were still alive and followed. Missing measurements due to early end of follow-up were not imputed since death status past the end of follow-up was unknown and the distribution of the last known KCCQ-CSS health states for patients with early end of follow-up was similar to the distribution among observed/imputed data. The observed transition probabilities were found to vary over the three time periods (baseline-week 12, week 12–32 and week 32–52), revealing inflection points at week 12 and week 32 (Table 28). Therefore, three sets of period-specific probabilities were used in the model. Each of the three derived matrices was then converted to monthly transition probabilities by finding the m-root of the observed transition matrix for a longer period (e.g., 12 weeks, 20 weeks). This yielded three sets of monthly transition probabilities representing progression in the three periods used in the analysis (Table 28) (121). The model uses the monthly transition matrices from the last period (month 9+) to predict progression after the first year, assuming the probabilities remain constant in the long-term.

Table 28. Monthly KCCQ-CSS transition matrix

KCCQ-CSS transitions [From, To]	Empagliflozin + SoC			Placebo + SoC		
	Months 1-3	Months 4-8	Months 9+	Months 1-3	Months 4-8	Months 9+
KCCQ [1,1]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [1,2]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [1,3]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [1,4]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [2,1]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [2,2]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [2,3]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [2,4]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [3,1]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [3,2]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [3,3]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [3,4]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [4,1]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [4,2]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [4,3]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ [4,4]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CSS, clinical summary score; KCCQ, Kansas City Cardiomyopathy Questionnaire; SoC, standard of care.

B.3.3.3 Mortality

B.3.3.3.1 CV mortality and all-cause mortality

A parametric survival analysis was conducted to allow for extrapolation of time to all-cause death and CV-related death as a function of treatment and time-varying KCCQ-CSS health states as measures of disease progression beyond the EMPEROR-Preserved trial duration. The analysis was conducted following recommendations of the NICE Decision Support Unit Technical Support Document 14 (169).

Selection of the best fitting parametric model consisted of fitting exponential, Weibull, Gompertz, log-logistic, log-normal, and the generalised gamma distributions to the observed data, including fitting of different models by treatment arm to explore and account for possible non-proportionality of effects. The goodness-of-fit was evaluated by graphical assessment of diagnostic plots, fit statistics (Akaike Information Criteria

with correction for a finite sample size [AICc] and Bayesian Information Criteria [BIC]) and clinical plausibility of extrapolations (Table 29). Time-varying indicators of the current health state were then introduced to a model based on the selected parametric distribution.

The joint arm Weibull model (i.e., the same parametric model fitted through both treatment arms) was selected as the base case distribution for all-cause and CV death. For all-cause mortality, the choice was based on Weibull being the best fitting distribution and yielded the most clinically plausible estimates of long-term survival (mean life expectancy of [redacted] months with placebo and [redacted] months with empagliflozin). To ensure the correct ordering of CV mortality and all-cause mortality predictions at all times, Weibull distribution was also chosen for modelling CV mortality. As CV mortality is cause-specific, it is expected for models to return higher survival estimates than their all-cause mortality counterparts. The log-normal and log-logistic models return implausibly high mean survival estimates across all formulations. The exponential model is also not plausible due to its constant hazard assumption. The Kaplan-Meier Weibull curves for all-cause and CV mortality are presented in Figure 18 and

Figure 19 respectively.

Table 29. Goodness-of-fit statistics (AICc/BIC) for alternative parametric distributions, all-cause mortality and CV mortality

Distribution	All-cause mortality		CV mortality	
	AICc	BIC	AICc	BIC
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AICc, corrected Akaike Information Criterion; BIC, Bayesian Information Criterion.

Figure 19. Kaplan-Meier curves for all-cause mortality from EMPEROR-Preserved and extrapolated survival curves (Weibull)

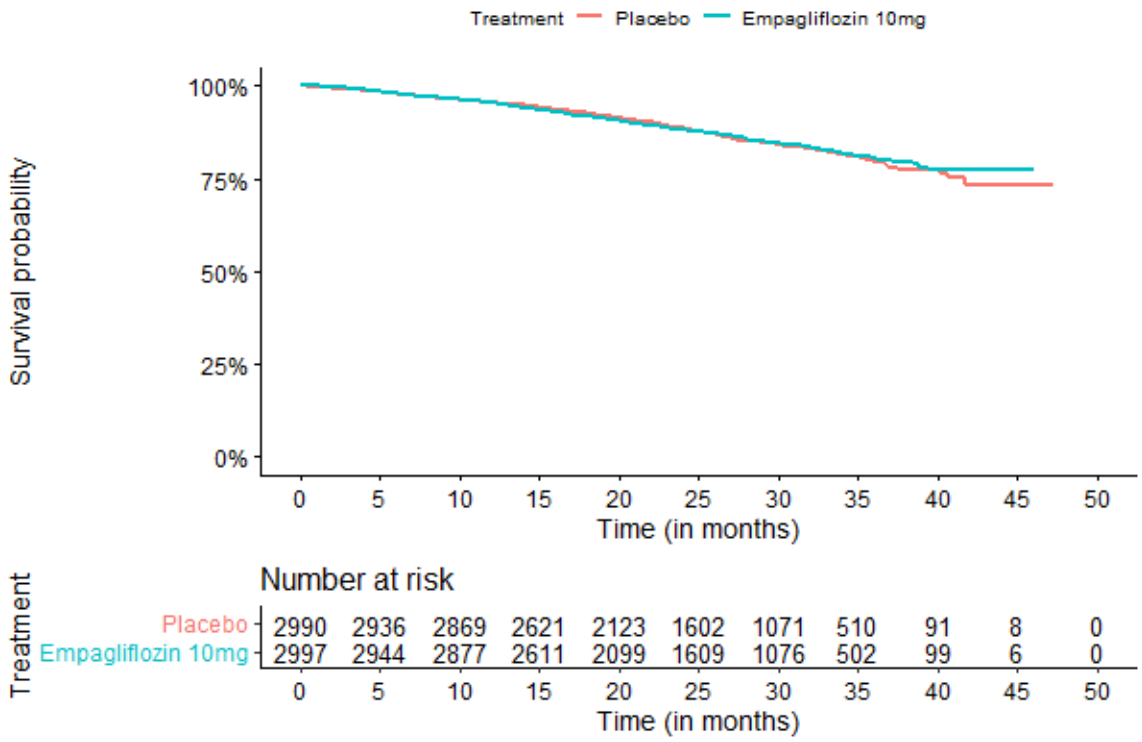
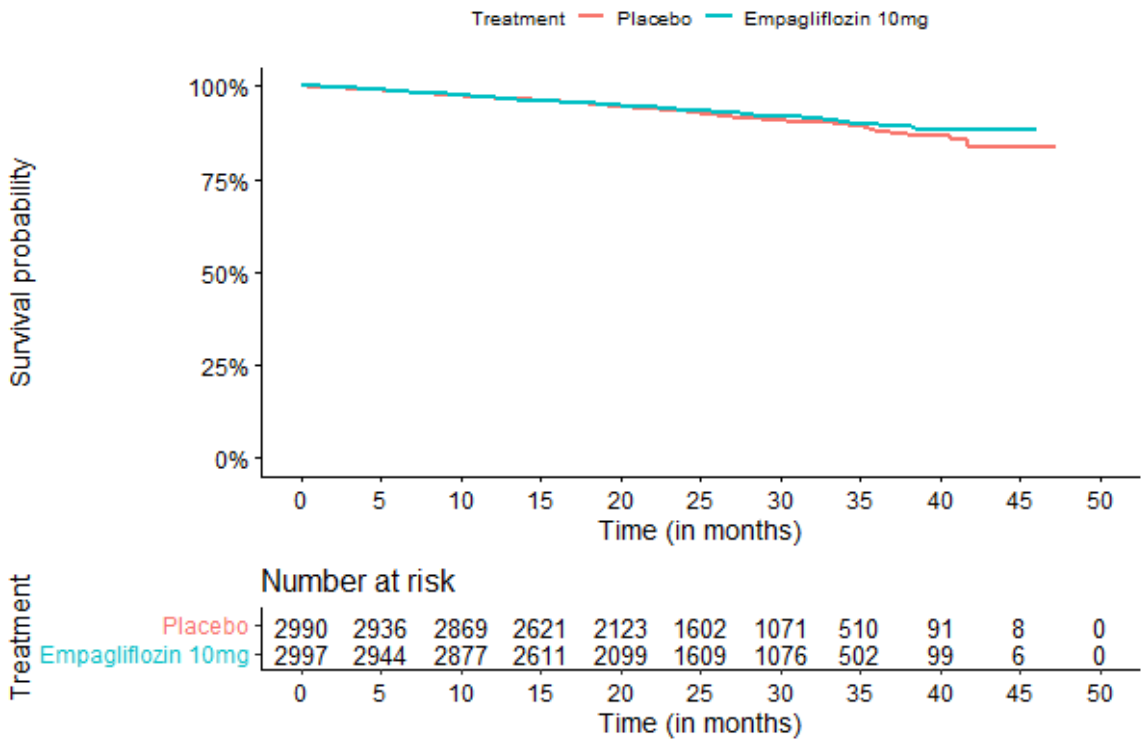


Figure 20. Kaplan-Meier curves for CV mortality from EMPEROR-Preserved and extrapolated survival curves (Weibull)



Coefficients of the Weibull risk equations for all-cause and CV mortality with time-updated KCCQ-CSS health states are shown in Table 30. In the CEM, the treatment effect was set to zero for AC mortality for the base case because the results from the respective statistical analysis were not statistically significant (i.e., no difference between the two comparators), and it is not clinically plausible to assume increased mortality for patients treated with empagliflozin. The treatment effect for CV mortality is also not statistically significant, but this was added in the model because a numerical difference between the two comparators is clinically plausible for this outcome.

Table 30. Parameterisation of survival equations for CV and all-cause death, ITT population of EMPEROR-Preserved trial, Weibull distribution (base case)

Parameter	All-cause death			CV death		
	Coefficients	SE	p-value	Coefficients	SE	p-value
Shape	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scale	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment effect Empagliflozin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Parameter	All-cause death			CV death		
	Coefficients	SE	p-value	Coefficients	SE	p-value
10 mg (Ref.: Placebo)						
KCCQ-CSS: [redacted] to < [redacted] (Quartile 2)*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: [redacted] to < [redacted] (Quartile 3)*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: [redacted] to [redacted] (Quartile 4)*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CV, cardiovascular; CSS, clinical summary score; ITT, intention to treat; KCCQ, Kansas City Cardiomyopathy Questionnaire; SE, standard error.

*vs KCCQ-CSS: 0 to 55.73 (Quartile 1)

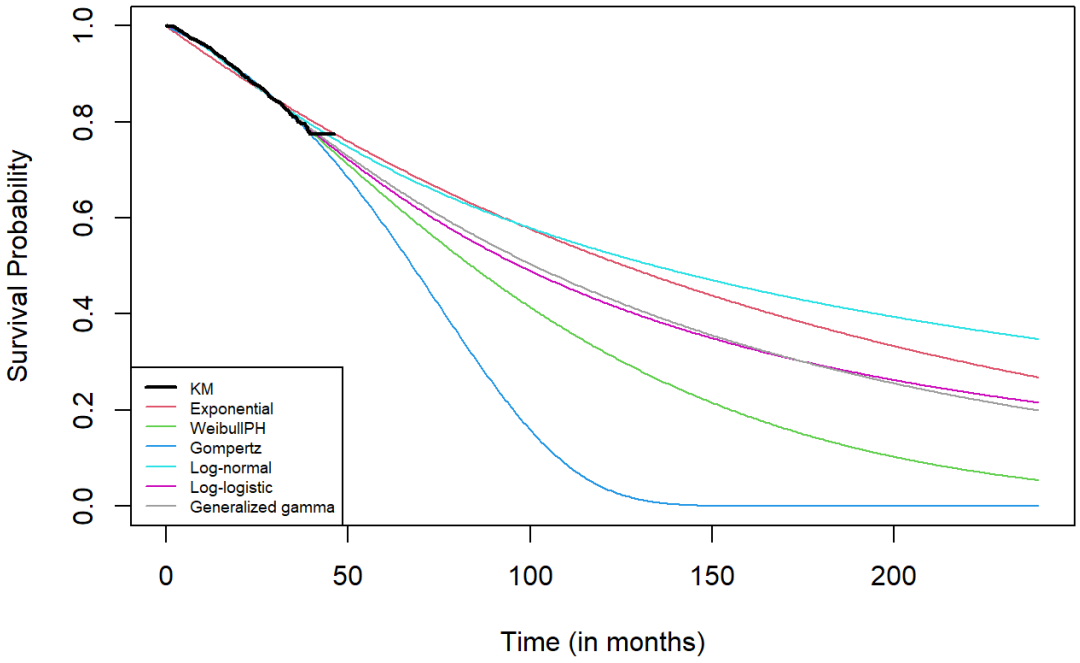
Other parametric distributions for all-cause mortality and CV mortality had the following disadvantages compared to Weibull:

- The constant hazard assumed by the exponential distribution was not realistic for a progressive disease, where the risk of CV events and death is likely to increase over time
- The log-normal and log-logistic distributions showed improbably long predicted mean survival times of around over 30 years in both arms in a population that is mostly over 65 years old, and improbably long predicted mean time to CV death
- The Gompertz fit yielded short survival predictions (less than five years in both arms) and a sharply increasing hazard of CV death, which likely overestimated risk in this population
- The joint fits with generalised gamma distribution were more plausible and comparable to those obtained with jointly fitted Weibull. Given the latter achieved a slightly better fit with fewer parameters, it was favoured as the optimal fit.

Long-term projections of all-cause and CV mortality using alternative parametric distributions are shown in Figure 20 and Figure 21 respectively. Risk equations derived with alternative distributions for sensitivity analyses are summarised in Table

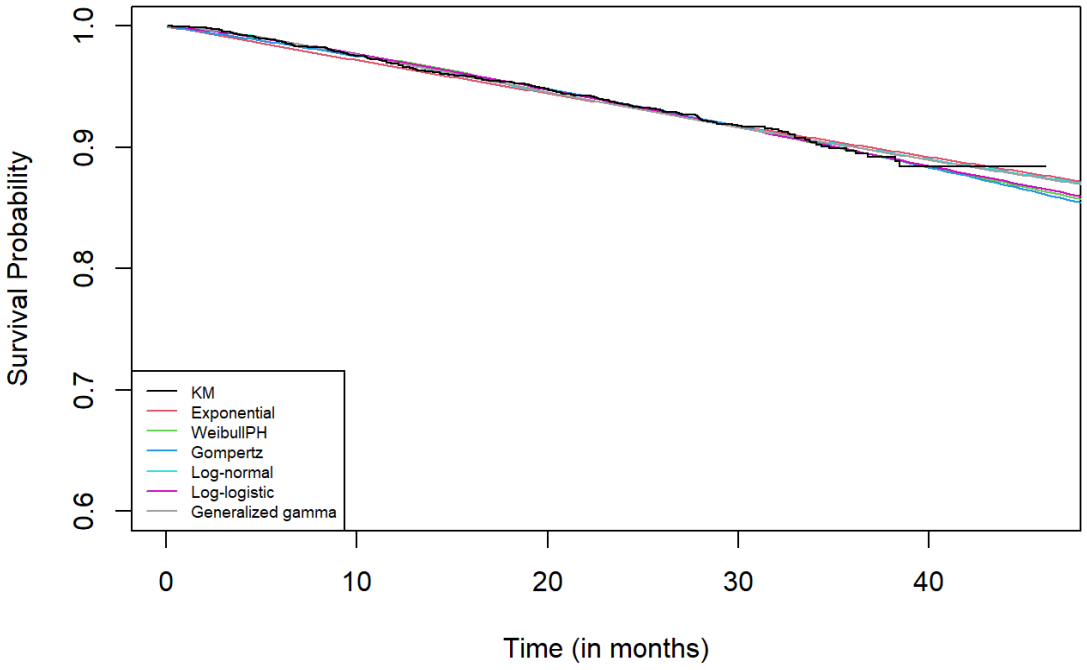
31. The proportional hazards assumption is assumed to hold and hence regression models jointly fit both arms using treatment as a predictor. Diagnostic plots that evaluated validity of the proportional hazards assumption for all the tested models are shown in Appendix N.

Figure 21. Alternative all-cause mortality survival curves (scenario analyses)



Abbreviations: KM, Kaplan-Meier curve.

Figure 22. Alternative CV mortality survival curves (scenario analyses)



Abbreviations: KM, Kaplan-Meier curve.

Table 31. Risk equations for alternative parametric distributions of CV mortality and all-cause mortality (KCCQ-based), ITT population of EMPEROR-Preserved (scenario analyses)

Coefficients	Exponential	Gompertz	Log-normal	Log-logistic	Generalised gamma
All-cause death					
P1		[redacted]	[redacted]	[redacted]	[redacted]
P2 (intercept)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment effect	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 90 to 100 (Quartile 4)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CV death					
P1		[redacted]	[redacted]	[redacted]	[redacted]
P2 (intercept)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P3					[redacted]

Coefficients	Exponential	Gompertz	Log-normal	Log-logistic	Generalised gamma
Treatment effect	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 90 to 100 (Quartile 4)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CSS, clinical summary score; CV, cardiovascular; ITT, intention to treat; KCCQ, Kansas City Cardiomyopathy Questionnaire.

B.3.3.3.2 Death from non-CV causes

Death attributable to non-CV causes was calculated from the all-cause death and CV death risk equations derived from the EMPEROR-Preserved trial (Section B.3.3.3.1). During each model cycle, the difference between the all-cause death rate and CV death rate was used to calculate non-CV death. If the probability of non-CV death was higher in a given cycle compared to the most recent age- and sex-specific life table probability for the general UK population, the latter was used to inform non-CV death in a given cycle. UK life tables were adjusted to exclude CV-related deaths to avoid double counting. UK life tables used to derive non-CV death for the model are reported in Appendix M.

A scenario analysis was carried out which used the CV death and all-cause death survival curves from EMPEROR-Preserved trial only, without applying the non-CV death rate from UK life tables (Section B.3.10.3**Error! Reference source not found.**).

B.3.3.4 Incidence of HHF

The monthly rate of first and recurrent HHF was modelled using a Poisson model fitted to patient-level data with generalised estimating equations (GEEs) which had an auto-regressive covariance structure to account for correlations between repeated measures as the data included a record for every month of follow-up for each patient. A negative binomial distribution was also considered but the fitting procedure failed and produced errors. The fit of the Poisson and negative binomial models were compared without the GEE correction and showed similar fit based on deviance

statistics, thus the negative binomial was not pursued further. The HHF rates appeared to be relatively constant over time and the analyses, therefore, assumed a constant rate in each treatment arm, but alternative scenarios where the benefit of treatment is turned off were considered.

The Poisson GEE model included treatment and time-varying KCCQ-CSS health states as predictors and was derived using the ITT population of the EMPEROR-Preserved. The parameters of the fitted GEE used to predict HHF in the ITT population in the model are reported in Table 32.

Table 32. Risk equation for hospitalisation for HF based on Poisson regression, ITT population from EMPEROR-Preserved trial

Covariate	Coefficient	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Treatment effect Empagliflozin 10 mg (Ref.: Placebo)	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55.73 to <73.96 (Quartile 2)*	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 73.96 to <88.02 (Quartile 3)*	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 88.02 to 100 (Quartile 4)*	[redacted]	[redacted]	[redacted]

Abbreviations: CSS, clinical summary score; ITT, intention to treat; KCCQ, Kansas City Cardiomyopathy Questionnaire; SE, standard error.

*vs KCCQ-CSS: 0 to <55 (Quartile 1)

B.3.3.5 Treatment discontinuation

A parametric survival analysis was applied to estimate the time to empagliflozin treatment discontinuation as observed in the EMPEROR-Preserved trial. Analyses considered treatment and time-varying KCCQ-CSS as predictors for discontinuation. Discontinuation due to death was not considered an event in these analyses as these were captured in mortality equations. Patients who died were censored for treatment discontinuation.

The Weibull, log-logistic, log-normal, Gompertz, exponential, and generalised gamma distributions were explored, and fits compared. Diagnostic plots and fit statistic are shown in Appendix N. Mean and median survival fit statistics suggested that the log-normal, log-logistic and Gompertz models return either an implausibly high or

diverging mean survival. Among the remaining models, the Weibull is the most optimistic, and the generalised gamma is the most pessimistic. While the generalised gamma is the most conservative model, it fits closer to the tail-end of the Kaplan-Meier than the other models, which tend to diverge upwards, resulting in higher survival estimates. Based on this consideration, as well as the plausibility of survival estimates, the generalised gamma model was chosen as the base case (Table 33). After discontinuation of empagliflozin + SoC, patients were assumed to receive SoC, and thus experience the same risk of clinical events, costs, and utility decrements as patients on SoC. The treatment discontinuation equations for alternative parametric distributions are provided in Table 34.

Table 33. Risk equation for treatment discontinuation from EMPEROR-Preserved trial, generalised gamma distribution

Covariate	Coefficient	SE	p-value
Mu	[redacted]	[redacted]	[redacted]
Sigma	[redacted]	[redacted]	[redacted]
Q	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)*	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 75 to <90 (Quartile 3)*	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 90 to 100 (Quartile 4)*	[redacted]	[redacted]	[redacted]

Abbreviations: CSS, clinical summary score; KCCQ, Kansas City Cardiomyopathy Questionnaire; SE, standard error.
*vs KCCQ-CSS: 0 to <55 (Quartile 1)

Table 34. Risk equations for alternative parametric distributions of treatment discontinuation, ITT population from EMPEROR-Preserved

Coefficients	Weibull	Exponential	Gompertz	Log-normal	Log-logistic
P1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P2 (intercept)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P3					
Treatment effect	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 90 to 100 (Quartile 4)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CSS, clinical summary score; ITT, intention to treat; KCCQ, Kansas City Cardiomyopathy Questionnaire.

B.3.3.6 Adverse event rates

The risk of experiencing AE from treatments was informed by the most common AE of special interest in the EMPEROR-Preserved trial by assuming a constant hazard. The rates of AE associated with empagliflozin + SoC and SoC were derived from the EMPEROR-Preserved trial. Patients who discontinued empagliflozin were subject to the risk of AE associated with the placebo arm of the EMPEROR-Preserved trial (Table 35).

Table 35. Rates of AE in the modelled cohort

	Rate per 100 patient-years in the EMPEROR-Preserved trial	
	Empagliflozin + SoC	**Placebo + SoC**
Urinary tract infection	[redacted]	[redacted]
Genital mycotic infection	[redacted]	[redacted]
Acute renal failure	[redacted]	[redacted]
Hepatic injury	[redacted]	[redacted]
Volume depletion	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]
Hypoglycaemic event*	[redacted]	[redacted]
Bone fracture	[redacted]	[redacted]
Ketoacidosis**	[redacted]	[redacted]

Abbreviation: AE, adverse event; SoC, standard of care.
*Defined as an event with a plasma glucose value of ≤70 mg/dL or where assistance was required.
**Not included in the base-case.

B.3.4 Measurement and valuation of health effects

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

The purpose of this section is to describe how HRQoL data was collected in the trial and comment on its consistency with NICE manual on health technology evaluations (166). The utility and disutility values associated with the model health states, AE and HHF were obtained from the pooled analysis of the patient-level data for the ITT population in the EMPEROR-Preserved trial (121). The values were derived from responses to the EQ-5D-5L questionnaire collected at baseline and at weeks 12, 32, 52, 100 and 148 following randomisation, at treatment discontinuation, and at a follow-up visit of 30 days following regular or premature completion of the treatment period.

Patients' responses to EQ-5D-5L questionnaire were then mapped to EQ-5D-3L descriptive system using the crosswalk mapping function developed by Hernández Alava et al. (170). The EQ-5D-3L responses were then converted to utility scores using the published UK utility values for EQ-5D health states, derived with the time trade-off method described by Dolan et. al. (171). Utility scores were analysed with a linear mixed-effects regression to account for the repeated measures on the same patients (172). The model included a random intercept for each patient to account for the within-patient correlations. To capture the short-and long-term effects of HHF events on utilities, the linear mixed model incorporated time-varying indicators reflecting whether a patient had a HHF in 0-1 month, 1-2 months, 2-4 months, and 4-12 months prior versus not hospitalised, as well as time-varying KCCQ-CSS quartiles. The reference group was no HHF events to date, and patients were classified back into the reference group once a year had passed from hospitalisation. This approach allowed estimation of the utilities based on patients' current severity level and HHF status. The assumption of disutility has been validated by patient organisations in the hearing to the empagliflozin EF ≤40% NICE submission (1). The choice of 12-month duration for the HHF disutility duration is based on precedent which was accepted by the NICE committee and EAG appraisal for dapagliflozin (155).

The AE effects were captured in the same way and assumed to be more acute. Indicators were created for each type of AE to flag whether it had occurred in the previous month and patients were returned to the reference group 1 month after the AE. The indicators for HHF and AE events were added to the baseline model, which was then trimmed down to remove predictors that became non-significant. Additionally, the model was adjusted for gender, age, region, and ischaemic cause.

The validity of the linear mixed model approach was verified by assessing the distribution of predicted values from the equations to ensure no ceiling effects were present. Specifically, the predicted values were within the expected ranges and less than 1% of the predicted values were above 1.00 (maximum at 1.004). The utility model used to inform health state utilities and utility decrements associated with clinical events is presented in Table 36.

Table 36. Health-related quality of life equation derived from EMPEROR-Preserved trial

Covariate	Coefficient	SE	p-value
Distribution/Type	Linear Mixed Model		
Intercept	[redacted]	[redacted]	[redacted]
Demographics			
Age ≥65	[redacted]	[redacted]	[redacted]
Male (ref: Female)	[redacted]	[redacted]	[redacted]
Region			
Latin America	[redacted]	[redacted]	[redacted]
North America	[redacted]	[redacted]	[redacted]
Asia	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Baseline KCCQ quartile (ref: [redacted])			
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 0 to 55 (Quartile 1)	[redacted]	[redacted]	[redacted]
Medical History			
Time Since HHF			
HHF: <1 month	[redacted]	[redacted]	[redacted]
HHF: 1 to <2 months	[redacted]	[redacted]	[redacted]
HHF: 2 to <4 months	[redacted]	[redacted]	[redacted]
HHF: 4 to <12 months	[redacted]	[redacted]	[redacted]
AE			
Urinary tract infection	[redacted]	[redacted]	[redacted]
Genital Mycotic Infection	[redacted]	[redacted]	[redacted]
Acute renal failure	[redacted]	[redacted]	[redacted]
Hepatic injury	[redacted]	[redacted]	[redacted]
Volume depletion	[redacted]	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]	[redacted]

Covariate	Coefficient	SE	p-value
Hypoglycaemic event	[redacted]	[redacted]	[redacted]
Bone fracture	[redacted]	[redacted]	[redacted]
Ketoacidosis*	[redacted]	[redacted]	[redacted]

Abbreviations: AE, adverse event; CSS, clinical summary score; EQ-5D, EuroQol five dimensions; HF, heart failure; HHF, hospitalisation for heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; SE, standard error.

*Not included in the base-case.

B.3.4.2 Mapping

The patient responses from the EQ-5D-5L questionnaire were mapped to the EQ-5D-3L questionnaire using the methodology developed by Hernández Alava et al. with UK value sets (170). This methodology is in line with the NICE manual on health technology evaluations (166).

B.3.4.3 Health-related quality of life studies

A SLR was conducted to identify relevant utilities evidence related to the treatment of chronic HF with EF >40%. The same SLR protocol was used for TA733, but only studies for patients with an EF<40% were included in the PRISMA diagram. Full details of the process and methods used are described in Appendix H.

Two randomised trials (SOCRATES-preserved trial and prehabilitation trial by Giovannini et al.) were included in the SLR and are summarised in Appendix H. The two randomized trials provided the utility values on treatment with vericiguat using the EQ-5D-3L (SOCRATES-Preserved trial), and with prehabilitation program using the EQ-5D (EQ-5D level not specified). The method of valuation was only reported in SOCRATES-Preserved trial. The responses were scored using the United States of America (USA) and United Kingdom (UK) societal value weights with the USA index score being reported in the study. Unlike SOCRATES, the analysis of the EMPEROR-Preserved trial allowed analysis of EQ-5D data by quartile, and thus was used for the economic analysis instead. Additionally, the population in EMPEROR-Preserved was considered the most relevant for this decision problem.

B.3.4.4 Adverse reactions

Utility decrements associated with adverse events of special interest (AESI) were derived either from the trial data as described in B.3.4.1 or from the literature and applied in the model over the month of incidence only Table 37. The disutility values for Hepatic injury, urinary tract infection, acute renal failure, volume depletion and bone fracture were derived from patient-level analysis of the EMPEROR-Preserved trial as the trial EQ-5D scores were deemed more reflective of the population of interest than the available estimates from the literature (172). The disutility associated with genital infection was sourced from Sullivan 2016 which provides a catalogue of disutility values for the UK (174). This study reported EQ-5D scores for diabetes-related chronic conditions, based on a nationally representative 12-item Short Form Health Survey response (n=20,705) from the US. These responses were mapped to EQ-5D-3L, and subsequently valued using UK-specific EQ-5D tariffs. The disutilities for hypoglycaemic events were identified as the most appropriate data based on the inclusion in the recently approved NICE appraisal for empagliflozin in the HFrEF population (1). The disutility value for hypotension was assumed equal to that of essential hypertension and taken from literature (175).

B.3.4.5 Health-related quality of life experienced in each health state

KCCQ-CSS health state-specific utility values were derived from the HRQoL equation shown in Table 35 based on EQ-5D data from EMPEROR-Preserved (172). The derived mean utility values were adjusted for gender, age (≥65 years), geographical region (Asia, Latin America, North America), baseline EQ-5D, and medical history (ischaemic cause of HF and history of HHF). The estimation of QALYs was determined by KCCQ-CSS state occupancy over time, and the incidence of discrete clinical events such as HHF and AE (urinary tract infection, genital mycotic infection, acute renal failure, hepatic injury, volume depletion, hypotension, hypoglycaemic event, bone fracture and ketoacidosis) which were captured as one-off utility decrements for the proportion of cohort experiencing the event in the month of incidence. No change in health state utility was considered based on age.

B.3.4.6 Health-related quality of life over the course of the disease

Changes in HRQoL over the course of the model were fully accounted for by changes in KCCQ-CSS health state occupancy, HHF events and occurrence of AE, which were included in the model as time-varying predictors (Section B.3.4.1). There were no statistically significant differences in EQ-5D scores between the two treatment groups in the EMPEROR-Preserved trial, hence the treatment was not a predictor in the utility equation.

B.3.4.7 Baseline HRQoL

The baseline utility values were contingent on the KCCQ-CSS quartile health states and were derived from the EMPEROR-Preserved data.

B.3.4.8 Adjusted health state utility values

As the trial-derived utility value for KCCQ-CSS quartile 4 ([redacted]) was higher than the utility of UK general population aged 60 to 69 years (0.723) reported by Sullivan et al. 2011, an age adjustment was applied (176). Under this adjustment, utility values for KCCQ-CSS quartile 1–3 were reduced by the relative difference between EMPEROR-Preserved observed utility for KCCQ-CSS quartile 4 and published utility of UK general population aged 70 to 79 years (176). The latter was assumed as the utility value for KCCQ-CSS quartile 4 utility in the model. The model provides flexibility to exclude the age-adjustment factor and instead use the EMPEROR-Preserved derived health state utility values, which was considered in a scenario analysis.

B.3.4.9 Summary of utility values

The health state utility values and disutilities associated with clinical events and AE included in the CEA are outlined in Table 37. Health state utility values and HHF disutility values were derived from the linear mixed-effects regression of EQ-5D data from the EMPEROR-Preserved trial. The AE disutility values were either derived from the trial data or identified from targeted literature searches.

Table 37. Summary of utility values for cost-effectiveness analysis

Parameter	Mean Utility	SE	Source and justification	Reference in the submission
Without Age adjustment				
KCCQ-CSS: 0 to <55 (Quartile 1)	[redacted]	[redacted]	Based on EMPEROR-Preserved trial data analyses (172)	B.3.4.1 and Table 36.
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]		B.3.4.1 and Table 36.
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]		B.3.4.1 and Table 36.
KCCQ-CSS: 90 to 100 (Quartile 4)	[redacted]	[redacted]		B.3.4.1 and Table 36.
With Age adjustment (base case)				
KCCQ-CSS: 0 to <55 (Quartile 1)*	[redacted]	[redacted]	Based on EMPEROR-Preserved trial data analyses (172)	B.3.4.1 and Table 36.
KCCQ-CSS: 55 to <75 (Quartile 2)*	[redacted]	[redacted]		B.3.4.1 and Table 36.
KCCQ-CSS: 75 to <90 (Quartile 3)*	[redacted]	[redacted]		B.3.4.1 and Table 36.
KCCQ-CSS: 90 to 100 (Quartile 4)^	[redacted]	[redacted]		B.3.4.1 and Table 36.
Clinical Event Disutility				
HHF	[redacted]	[redacted]	Based on EMPEROR-Preserved trial data analyses (172)	B.3.4.1 and Table 36.
AE Disutilities				
Hepatic injury	[redacted]	[redacted]	Based on EMPEROR-Preserved trial data analyses (172)	B.3.4.1, Table 36, and B.3.4.4.
Urinary tract infection	[redacted]	[redacted]		B.3.4.1, Table 36, and B.3.4.4.
Acute renal failure	[redacted]	[redacted]		B.3.4.1, Table 36, and B.3.4.4.
Volume depletion	[redacted]	[redacted]		B.3.4.1, Table 36, and B.3.4.4.
Bone fracture	[redacted]	[redacted]		B.3.4.1, Table 36, and B.3.4.4.
Hypotension^^	[redacted]	[redacted]	Sullivan 2006 (175)	B.3.4.1, Table 36, and B.3.4.4.

Parameter	Mean Utility	SE	Source and justification	Reference in the submission
Genital mycotic infection	[redacted]	[redacted]	Sullivan 2016 (174)	B.3.4.1, Table 36, and B.3.4.4.
Hypoglycaemic event	[redacted]	[redacted]	Pollard 2017, Peasgood 2016 (177, 178)	B.3.4.1, Table 36, and B.3.4.4.
Ketoacidosis	[redacted]	[redacted]	Pollard 2017, Peasgood 2016 (177, 178)	B.3.4.1, Table 36, and B.3.4.4.

Abbreviations: AE, adverse event; CV, cardiovascular; CSS, clinical summary score; HHF, hospitalisation due to heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; SE, standard error.

*Relative differences from EMPEROR-Preserved study applied to general population utility for people aged 60 to 69 years reported by Sullivan et al. 2011 (176)

^ Set equal to UK general population utility for people aged 60 to 69 years reported by Sullivan et al. 2011 (176)

^^ Disutility for hypotension in the US population reported by Sullivan 2006 due to lack of UK population values (175)

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

A SLR was conducted to identify cost and healthcare resource use evidence related to the treatment of chronic HF with EF >40. The same SLR protocol was used for TA733, but only studies for patients with an EF<40% were included in the PRISMA diagram. Full details of the process and methods are provided in Appendix I. In summary, no relevant studies were identified for inclusion in this SLR.

B.3.5.1 Intervention and comparators' costs and resource use

Intervention and comparator costs in the model consist of drug acquisition costs and monitoring costs. In the base case analysis, empagliflozin + SoC is compared to SoC alone (i.e., usual therapies treating CV comorbidities). The drug costs for empagliflozin and the SoC therapies were extracted from the NHS Electronic Drug Tariff (179).

Patients were assumed to receive appropriately titrated doses of SoC therapies, (i.e., the stable maintenance dosage for each SoC treatment was applied and the titration process was not modelled). The background therapy used within the trial was as per national or international guideline recommendations (3, 20). Costs of devices were not included as patients were assumed to have undergone procedures for these treatments before entering the model.

A summary of the pack cost, pack size, strength, dosage, daily and monthly cost are provided in Table 38.

Table 38: Technology and comparator unit costs

Drug class	Treatment	Pack cost (MIMS)	Pack size	Strength (mg)	Daily dosage	Daily cost	Monthly cost	Source
SGLT2i	**Empagliflozin**	£36.59	28 pills	10 mg	10 mg	£1.31	£39.78	NHS electronic drug tariff database March 2022 (180)
ARNI	**Sacubitril valsartan**	£91.56	56 pills	200 mg	400 mg	£3.27	£99.53	
Loop diuretics	**Furosemide**	£0.70	28 pills	40 mg	80 mg	£0.05	£1.52	
ACEI	**Captopril**	£1.66	56 pills	50 mg	100 mg	£0.06	£1.80	
	Enalapril	£4.95	28 pills	20 mg	20 mg	£0.18	£5.38	
	Lisinopril	£0.91	28 pills	20 mg	20 mg	£0.03	£0.99	
	Ramipril	£1.06	28 pills	10 mg	10 mg	£0.04	£1.15	
	Trandolapril	£1.68	14 pills	0.5 mg	1.5 mg	£0.36	£10.96	
BB	**Bisoprolol**	£0.92	28 pills	10 mg	10 mg	£0.03	£1.00	
	Carvedilol	£1.24	28 pills	25 mg	50 mg	£0.09	£2.70	
	Metoprolol	£1.44	28 pills	100 mg	100 mg	£0.05	£1.57	
	Nebivolol	£27.39	28 pills	10 mg	10 mg	£0.98	£29.77	
ARBs	**Candesartan**	£1.41	28 pills	32 mg	32 mg	£0.05	£1.53	
	Valsartan	£14.69	28 pills	160 mg	320 mg	£1.05	£31.94	
	Losartan	£1.20	28 pills	100 mg	150 mg	£0.06	£1.96	
MRA	**Eplerenone**	£9.43	28 pills	50 mg	50 mg	£0.34	£10.25	

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; BB, beta-blockers; MRA, mineralocorticoid receptor antagonist; SGLT2i, sodium-glucose co-transporter-2 inhibitor.

B.3.5.2 Health state unit costs and resource use

The health state resource use and unit costs associated with chronic HF with EF >40% are outlined in Table 39 and Table 40. The costs of managing clinical events were obtained from UK national databases or published literature.

The acute cost of HHF was based on NHS reference costs for non-elective long inpatient stay, computed as the weighted average of reference costs for healthcare resource group (HRG) codes EB03A to EB03E and the number of finished consultant episodes (FCEs) (181).

The cost of CV death was estimated from a study by Alva et al., who estimated inpatient costs for T2D-related complications during the UK Prospective Diabetes Study post-trial monitoring period from 1997 to 2007 using hospitalisation records for patients in England (n=2,791) (182). Their analysis produced an equation with coefficients interpreted as linear effects on expected inpatient costs for complications, which was used in the CEA to estimate costs of a fatal MI, fatal IHD, and fatal stroke for a male aged <65 years and ≥65 years and a female aged <65 years and ≥65 years. Characteristics of EMPEROR-Preserved participants (e.g., percentage of male or female and percentage of aged <65 or ≥65 years) were applied to derive weighted average costs for each event, which were themselves averaged to derive the cost of CV death for the model. Non-CV deaths were assumed to incur no cost in the base case and were assumed to equal the cost of CV death in a scenario analysis.

The HF-related disease management costs associated with GP, cardiologist visits and A&E referral were computed based on the frequency of use and unit cost for each type of care. Resource use was based on data from the CPRD, as reported by McMurray and colleagues (2018), which was converted from annual to monthly frequency (183). Unit costs were retrieved from national sources. In particular, the cost of GP and cardiologist visits were based upon per patient contact lasting 9.22 minutes (code 10.3b) and a consultant-led non-admitted face to face follow-up appointment in cardiology (code 320), respectively, while A&E referral cost was a weighted mean derived from national average unit costs and number of FCEs for non-admitted

emergency medicine (codes VB01Z to VB11Z, and VB99Z) (184, 185). All disease management costs were inflated to 2021 by applying the consumer price health inflation factor from Eurostat (179).

Table 39. CV events management unit costs

CV events	Unit cost per event (inflated to 2021)	Cost source
HHF	£3,258.58	NHS 2019-2020; Weighted average of non-elective long stay HRG codes; EB03A: EB03E (181)
CV death	£4,295.01	Alva 2015 (182)
Non-CV death*	£0.00	Assumption

Abbreviations: CV, cardiovascular; HHF, hospitalisation for heart failure; HRG, healthcare resource group
Notes: Non-CV death is used in model background calculations to correctly compute the number of patients remaining alive (and on-treatment) from year to year.
**Source: EMPEROR-Preserved Clinical trials report (172).

Table 40. Disease management resource use and unit costs

Disease management costs	Unit cost per event	Monthly frequency for all KCCQ-CSS quartiles*	Monthly cost per patient for all KCCQ-CSS quartiles	Cost source
GP visit	£39.23	0.999	£39.23	PSSRU 2021, Code 10.3b (9.22 minutes per patient contact) (184)
Cardiologist visit	£143.12	0.0022	£0.31	PSSRU 2021, Cardiology non-admitted face to face and follow-up visit (184)
A&E referral	£188.63	0.0004	£0.08	PSSRU 2021, weighted mean (HRG codes: VB01Z-VB11Z, VB99Z) (184)
Total cost			**£39.62**	

Abbreviations: A&E, accident and emergency; GP, general practitioner; HRG, healthcare resource group; KCCQ,-CSS Kansas City Cardiomyopathy Questionnaire – clinical summary score; PSSRU, personal social services research unit.
*Monthly frequency: McMurray et al (2018) (183)

B.3.5.3 Adverse reaction unit costs and resource use

The adverse reaction unit costs and resource use are provided in Table 41**Error! Reference source not found.** below. The acute cost of an outpatient visit was based on costs for GPs, assuming per patient contact lasting 9.22 minutes (code 10.3b), taken from unit costs of health and social care by personal social services research unit (PSSRU) 2021 (184). In addition, NHS reference costs for non-elective long and short stays for HRG codes related with each AE served as the basis for the cost of inpatient episodes (Table 41), while self-treated patients were assumed to receive over-the-counter treatment, thus incurring no costs to the health care payer. The distribution of visit types for management of AE was based on assumption, as UK-specific data was not available and the event rate was derived from the EMPEROR-Preserved trial (172, 186). All AE costs were inflated to 2021 by applying the consumer price health inflation factor from Eurostat (179).

Table 41. Adverse event management unit costs, event rate and frequencies of distribution

Adverse event management costs	Weighted average cost (2021)	Outpatient*	Inpatient	Event rate for overall CHF with EF >40% (per 1,000 patient-years)**		Inpatient cost source
		Unit cost (inflated to 2021) (%)	Unit cost (inflated to 2021) (%)	SoC	Empagliflozin + SoC	
Urinary tract infection	£39.23	£39.23 (100%)	£1,799.08 (0%)	4.53	5.56	NHS 2019-20; weighted average of HRG codes: LA04H, LA04J-N, LA04P-S, kidney or urinary tract infections, non-elective long or short stay
Genital mycotic infection	£39.23	£39.23 (100%)	£1,300.67 (0%)	0.39	1.2	NHS 2019-20 weighted average of HRG codes: WJ03A-G, standard infection diseases, non-elective long or short stay
Acute renal failure	£2,067.24	£39.23 (0%)	£2,067.24 (100%)	7.26	6.87	NHS 2019-20; weighted average of HRG codes: LA07H, LA07J-N, LA07P, acute kidney injury, non-elective long or short stay
Hepatic injury	£1,301.74	£39.23 (50%)	£2,564.25 (50%)	2.84	2.08	NHS 2019-20; weighted average of HRG codes: GC01C-F, liver failure disorders, non-elective long or short stay
Volume depletion	£39.23	£39.23 (100%)	£1,507.16 (0%)	5.38	6.78	NHS 2019-20; weighted average of HRG codes: KC05G-H, KC05J-N, fluid or electrolyte disorders, non-elective long or short stay

Adverse event management costs	Weighted average cost (2021)	Outpatient*	Inpatient	Event rate for overall CHF with EF >40% (per 1,000 patient-years)**		Inpatient cost source
		Unit cost (inflated to 2021) (%)	Unit cost (inflated to 2021) (%)	SoC	Empagliflozin + SoC	
Hypotension	£39.23	£39.23 (100%)	£1,976.92 (0%)	4.8	5.88	NHS 2019-20; weighted average of HRG codes: EB14A-E, other acquired cardiac conditions, non-elective long or short stay
Hypoglycaemic event**	£699.99	£39.23 (50%)	£1,360.75 (50%)	1.41	1.31	NHS 2019-20; weighted average of HRG codes: KA08A-C, other endocrine disorders, non-elective long or short stay
Bone fracture	£3,098.62	£39.23 (0%)	£3,098.62 (100%)	2.30	2.43	NHS 2019-20; weighted average of HRG codes: HD39D-H, pathological fractures, non-elective long or short stay
Cough	£39.23	£39.23 (100%)	N/A	0.00	0.00	N/A
Dizziness	£39.23	£39.23 (100%)	N/A	0.00	0.00	N/A
Ketoacidosis	£632.55	£39.23 (50%)	£1,225.87 (50%)	0.78	0.90	NHS 2019-20; (HRG codes: KB02G-K)

Abbreviations: CHF, Chronic heart failure; EF, Left ventricle ejection fraction; HRG, healthcare resource group; N/A, not applicable; NHS, National Health Service; SoC, standard of care; UTI, Urinary tract infection.
*Outpatient source: PSSRU 2021, Code 10.3b (9.22 minutes per patient contact) and inflated to 2021 (184)
**Event rate source: EMPEROR-Preserved Clinical trials report (172)
NHS 2019-20 reference: National Health Service (NHS). National Schedule of NHS Costs - Year 2019-20 (181)

B.3.5.4 Miscellaneous unit costs, indirect costs and resource use

There are no miscellaneous unit costs, indirect costs and resource use are not captured.

B.3.6 Severity

Not applicable.

B.3.7 Uncertainty

Not applicable.

B.3.8 Summary of base case analysis inputs and assumptions

B.3.8.1 Summary of base case analysis inputs

Table 42. Summary of variables applied in the economic model

Variable	Value	SE	Distribution		Reference
	Baseline characteristics				
Age (years)	71.9	0.12	NA		Table 28
Male	55.3%	0.01	NA		
KCCQ-CSS Q1: 0-<55.73	[redacted]	[redacted]	NA		
KCCQ-CSS Q2: 55.73-<73.96	[redacted]	[redacted]	NA		
KCCQ-CSS Q3: 73.96-<88.02	[redacted]	[redacted]	NA		
KCCQ-CSS Q4: 88.02-100	[redacted]	[redacted]	NA		
Ischaemic HF	35.4%	0.01	NA		
Treatment use at baseline			NA		
ACEI	40.2%	0.01	NA		
ARB	38.7%	0.01	NA		
ARNI	2.2%	0.00	NA		
MRA	37.5%	0.01	NA		
BB	86.3%	0.00	NA		
Loop Diuretics	67.7%	0.00	NA		
	Monthly KCCQ-CSS transition matrix – months 0–3, Empagliflozin + SoC				
KCCQ [1,1]	[redacted]	[redacted]	Dirichlet		Table 28
KCCQ [1,2]	[redacted]	[redacted]	Dirichlet		

Variable	Value	SE	Distribution		Reference
KCCQ [1,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,4]	[redacted]	[redacted]	Dirichlet		
	Monthly KCCQ-CSS transition matrix – months 4-8, Empagliflozin + SoC				
KCCQ [1,1]	[redacted]	[redacted]	Dirichlet		Table 28
KCCQ [1,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,4]	[redacted]	[redacted]	Dirichlet		
	Monthly KCCQ-CSS transition matrix – months 9+, Empagliflozin + SoC				
KCCQ [1,1]	[redacted]	[redacted]	Dirichlet		**Error! Reference source not found.**
KCCQ [1,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,1]	[redacted]	[redacted]	Dirichlet		

Variable	Value	SE	Distribution		Reference
KCCQ [4,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,4]	[redacted]	[redacted]	Dirichlet		
	Monthly KCCQ-CSS transition matrix – months 0–3, Placebo + SoC				
KCCQ [1,1]	[redacted]	[redacted]	Dirichlet		Table 28
KCCQ [1,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,4]	[redacted]	[redacted]	Dirichlet		
	Monthly KCCQ-CSS transition matrix – months 4-8, Empagliflozin + SoC				
KCCQ [1,1]	[redacted]	[redacted]	Dirichlet		Table 28
KCCQ [1,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,4]	[redacted]	[redacted]	Dirichlet		
	Monthly KCCQ-CSS transition matrix – months 9+, Empagliflozin + SoC				
KCCQ [1,1]	[redacted]	[redacted]	Dirichlet		Table 28
KCCQ [1,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [1,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [2,3]	[redacted]	[redacted]	Dirichlet		

Variable	Value	SE	Distribution		Reference
KCCQ [2,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [3,4]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,1]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,2]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,3]	[redacted]	[redacted]	Dirichlet		
KCCQ [4,4]	[redacted]	[redacted]	Dirichlet		
	Adjusted CV mortality survival equation* (Weibull)				
Shape	[redacted]	[redacted]	Multivariate normal		Table 30
Scale	[redacted]	[redacted]	Multivariate normal		
Treatment effect	[redacted]	[redacted]	Multivariate normal		
KCCQ Q2	[redacted]	[redacted]	Multivariate normal		
KCCQ Q3	[redacted]	[redacted]	Multivariate normal		
KCCQ Q4	[redacted]	[redacted]	Multivariate normal		
	Adjusted all-cause mortality survival equation* (Weibull)				
Shape	[redacted]	[redacted]	Multivariate normal		Table 30
Scale	[redacted]	[redacted]	Multivariate normal		
Treatment effect	[redacted]	[redacted]	Multivariate normal		
KCCQ Q2	[redacted]	[redacted]	Multivariate normal		
KCCQ Q3	[redacted]	[redacted]	Multivariate normal		
KCCQ Q4	[redacted]	[redacted]	Multivariate normal		
	Adjusted generalised estimating equations for HHF events* (Poisson)				
Intercept	[redacted]	[redacted]	Multivariate normal		Table 32
Treatment effect	[redacted]	[redacted]	Multivariate normal		
KCCQ Q2	[redacted]	[redacted]	Multivariate normal		
KCCQ Q3	[redacted]	[redacted]	Multivariate normal		
KCCQ Q4	[redacted]	[redacted]	Multivariate normal		
	Treatment discontinuation equations* (Generalised Gamma)				
Mu	[redacted]	[redacted]	Multivariate normal		Table 30
Sigma	[redacted]	[redacted]	Multivariate normal		

Variable	Value	SE	Distribution		Reference
Q	[redacted]	[redacted]	Multivariate normal		
KCCQ Q2	[redacted]	[redacted]	Multivariate normal		
KCCQ Q3	[redacted]	[redacted]	Multivariate normal		
KCCQ Q4	[redacted]	[redacted]	Multivariate normal		
	Adverse events rates per 100 patient-years – Empagliflozin + SoC				
Urinary tract infection	[redacted]	[redacted]	Gamma		Table 35
Genital Mycotic Infection	[redacted]	[redacted]	Gamma		
Acute renal failure	[redacted]	[redacted]	Gamma		
Hepatic injury	[redacted]	[redacted]	Gamma		
Volume depletion	[redacted]	[redacted]	Gamma		
Hypotension	[redacted]	[redacted]	Gamma		
Hypoglycaemic event	[redacted]	[redacted]	Gamma		
Bone fracture	[redacted]	[redacted]	Gamma		
Ketoacidosis	[redacted]	[redacted]	Gamma		
	Adverse events rates per 100 patient-years – Placebo + SoC				
Urinary tract infection	[redacted]	[redacted]	Gamma		Table 35
Genital Mycotic Infection	[redacted]	[redacted]	Gamma		
Acute renal failure	[redacted]	[redacted]	Gamma		
Hepatic injury	[redacted]	[redacted]	Gamma		
Volume depletion	[redacted]	[redacted]	Gamma		
Hypotension	[redacted]	[redacted]	Gamma		
Hypoglycaemic event	[redacted]	[redacted]	Gamma		
Bone fracture	[redacted]	[redacted]	Gamma		
	Utility values – health states and events (with age adjustment)				
KCCQ Q1	[redacted]	[redacted]	Beta		Table 36
KCCQ Q2	[redacted]	[redacted]	Beta		
KCCQ Q3	[redacted]	[redacted]	Beta		
KCCQ Q4	[redacted]	[redacted]	Beta		
HHF (decrement)	[redacted]	[redacted]	Beta		
	Disutility values – adverse events				
Urinary tract infection	[redacted]	[redacted]	Beta		Table 36

Variable	Value	SE	Distribution		Reference
Genital Mycotic Infection	[redacted]	[redacted]	Beta		
Acute renal failure	[redacted]	[redacted]	Beta		
Hepatic injury	[redacted]	[redacted]	Beta		
Volume depletion	[redacted]	[redacted]	Beta		
Hypotension	[redacted]	[redacted]	Beta		
Hypoglycaemic event	[redacted]	[redacted]	Beta		
Bone fracture	[redacted]	[redacted]	Beta		
Ketoacidosis	[redacted]	[redacted]	Beta		
	Treatment acquisition costs per cycle				
Empagliflozin + SoC	[redacted]	[redacted]	N/A		Table 38**Error! Reference source not found.**
SoC	[redacted]	[redacted]	N/A		
	Health state and event costs				
HHF	[redacted]	[redacted]	Gamma		Table 40
CV death	[redacted]	[redacted]	Gamma		
Non-CV death	0.000	N/A	N/A		
	Adverse event unit costs				
Urinary tract infection	[redacted]	[redacted]	Gamma		Table 41
Genital Mycotic Infection	[redacted]	[redacted]	Gamma		
Acute renal failure	[redacted]	[redacted]	Gamma		
Hepatic injury	[redacted]	[redacted]	Gamma		
Volume depletion	[redacted]	[redacted]	Gamma		
Hypotension	[redacted]	[redacted]	Gamma		
Hypoglycaemic event	[redacted]	[redacted]	Gamma		
Bone fracture	[redacted]	[redacted]	Gamma		
Ketoacidosis	[redacted]	[redacted]	Gamma		
	Resource use costs: Monitoring costs per cycle				
GP visit	[redacted]	[redacted]	Gamma		Table 40
Cardiologist visit	[redacted]	[redacted]	Gamma		
A&E referral	[redacted]	[redacted]	Gamma		

Abbreviations: A&E, Accident & Emergency; ACEI, Angiotensin-converting enzyme inhibitors; ARB, Angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; BB, beta-blocker; CSS, clinical summary score; CV, Cardiovascular; GP, General Practitioner; HHF, hospitalisation for heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; MRAs, Mineralocorticoid receptor antagonists; SE, standard error; SoC, Standard of Care, .
*The standard errors reported here are different from the standard errors obtained directly from regression models due to Cholesky decomposition.

B.3.8.2 Assumptions

The CEM uses the best available evidence to inform the decision problem, in line with the NICE reference case and guidance on methods of appraisal (166). Table 43 outlines the simplifying assumptions which were needed either in the absence of data or to ensure internal validity (i.e. preserve alignment between the modelled outcomes and those observed in the EMPEROR-Preserved trial). All assumptions have bused in the EF ≤40% TA773 submission and have been included in this EF >40% model.

Table 43. Summary of key assumptions of the economic analysis

#	Assumption	Justification	Likely bias	Was this assumption included in initial Company Submission in TA733?
Model structure				
1	Clinical event rates observed in clinical practice mirror those observed in the EMPEROR-Preserved trial	Generalisability of trial outcomes to clinical practice is a common assumption in economic modelling. In this case, the plausibility of the assumption is strengthened by the trial protocol requirement for consistency with local guidelines in standards of HF care. Thus, event rates observed in the trial have direct relevance to clinical practice. To assess this assumption, the event rates in EMPEROR-Preserved vs PULSE are compared in *Document B, B.3.13.* Exploring uncertainty	None	Yes
2	There may be unmodelled comorbidities that could have influenced the shapes of the statistical extrapolations for HHF and death (CV or all-	This choice was made to ensure that the model retains internal validity i.e., the model-predicted outcomes match those observed in the EMPEROR-Preserved trials	None	Yes

#	Assumption	Justification	Likely bias	Was this assumption included in initial Company Submission in TA733?
	cause). The risk equations incorporated only treatment and time-varying KCCQ-CSS states as predictors			
3	The rate of clinical events beyond trial duration is based on extrapolation of the observed trial outcomes	This is a limitation inherent to most cost-effectiveness models. No reliable external source was available for estimating the rate of clinical events beyond the duration of the trial. An external validation exercise was undertaken to assess the validity of the long-term extrapolations (see B.3.13.1.2).	None	Yes
4	The HHF risk equations/rates only account for non-fatal HHF events	Fatal HHF events were captured by the CV death risk equations.	None	Yes
5	Patients can experience non-CV death in any KCCQ-CSS health state. Non-CV death was computed separately	This is a commonly utilised approach in cost-effectiveness analysis to model death from other causes.	None	Yes
6	Temporal changes in serum concentration of NT-pro-BNP, a prognostic biomarker of HF morbidity and mortality, were not modelled	This is unlikely to affect cost-effectiveness results since risk equations account for the time-updated KCCQ-CSS health states and therefore capture the impact of disease severity on all-cause mortality, CV mortality, and the risk of HHF.	None	Yes
7	The model assumed that all patients receive appropriately titrated doses of HF medications (e.g., ACEI/ARB). This does not imply that patients receive the maximum effective dose but a dosage that achieves the best trade-off between effectiveness and tolerability	This is a simplifying assumption not expected to lead to bias in favour of any treatment.	None	Yes
8	The modelled rate of treatment discontinuation is derived from the	Patients were on SoC before receiving add-on empagliflozin and it is reasonable to assume that they will continue on SoC	None	Yes

#	Assumption	Justification	Likely bias	Was this assumption included in initial Company Submission in TA733?
	EMPEROR-Preserved trial, with a rate of discontinuation applied to all patients receiving empagliflozin + SoC in each modelled cycle (based on the selected distribution). Following discontinuation of empagliflozin, patients are assumed to have the same event risks and costs as patients in the control arm	after discontinuing empagliflozin to the same extent as they discontinued SoC in the comparator arm.		
HRQoL				
9	KCCQ-CSS health state-specific utility values and disutilities associated with AE and HHF were derived from pooled analysis of the EMPEROR-Preserved ITT population after mapping EQ-5D-5L data to EQ-5D-3L and applying the UK value sets	This is common practice when EQ-5D data is available from the clinical trial and in line with the NICE reference case (166)	None	Yes
10	The model assumed no decline in HRQoL with increasing age	This is a simplifying assumption . The model includes functionality to adjust utilities to reflect those of age-matched UK general population and in that way already partially reflects the expected utility for the corresponding age groups.	None	Yes
Costs and resource use				
11	The cost of non-CV death was assumed to be zero under all interventions	The rate of non-CV death is expected to be the same across all interventions, hence this assumption is unlikely to	None	Yes

#	Assumption	Justification	Likely bias	Was this assumption included in initial Company Submission in TA733?
		have an impact on incremental cost-effectiveness. Nevertheless, a non-zero cost was provided to allow testing of the alternative costing scenario.		
12	The model does not include the cost of medical devices or their implantation. It is assumed that patients with ICD/CRT had the device implanted prior to entering the model	This is a simplifying assumption and is not expected to bias any treatment.	None	Yes

Abbreviations: AE, adverse events; ACEI, Angiotensin-converting enzyme inhibitors; ARB, Angiotensin receptor blocker; NT-pro-BNP, N-terminal pro hormone B-type natriuretic peptide; CHF, Chronic heart failure; CV, cardiovascular; CRT, cardiac resynchronisation therapy; EQ-5D-5L, EuroQol-5 dimensions-5 levels; EQ-5D-3L, EuroQol-5 dimensions-3 levels; HHF, hospitalisation for heart failure; HRQoL, health-related quality of life; ICD, implantable cardioverter-defibrillator; ITT, intention to treat; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire Clinical Summary Score; EF, Left ventricular ejection fraction; NICE, National Institute for Health and Care Excellence; NT-pro-BNP, N-terminal pro; SGLT2i, sodium-glucose co-transporter 2 inhibitor; UK, United Kingdom.

B.3.9 Base case results

B.3.9.1 Base case incremental cost-effectiveness analysis results

Table 44 shows the discounted results of the EF >40% base case comparison of empagliflozin as an add-on to standard care (ACEI/ARB + BB ± MRA) against SoC alone over a lifetime horizon. SoC is associated with [redacted] LYs, [redacted] QALYs, and [redacted] per patient. Treatment with empagliflozin as an add-on to SoC resulted in an increase in LYs (+[redacted] per person) and QALYs (+[redacted]) per person at an additional cost of £1,682 per person. Empagliflozin as an add-on to SoC was cost-effective against SoC alone at usual threshold values with an ICER of £14,429 per QALY gained. During Technical Engagement for TA773, the BMJ EAG noted that the LYs gained in the model was lower than the QALYs gained, which appears counterintuitive. BI would like to proactively address this query as it is also relevant for this appraisal. The overall sum of QALYs across all health states ([redacted]) is slightly higher than the overall sum of LYs ([redacted]). This is because each health state is associated with a difference utility

value and the final QALY value is the utility weighted sum of the life years across all KCCQ-CSS health states.

The clinical outcomes of the model and disaggregated results of the base case analysis are presented in Appendix J.

Table 44. Base case analysis: deterministic results for empagliflozin as an add-on to standard care

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
SoC	[redacted]	[redacted]	[redacted]	-	-	-	-
Empagliflozin + SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14,429

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; SoC, standard of care; QALYs, quality-adjusted life year.

Currently, empagliflozin is indicated for patients with EF ≤40% with a base case ICER from the EF ≤40% technical engagement model of £4,999 per QALY demonstrating that empagliflozin is also cost-effective in this patient population as an add-on to SoC.

Empagliflozin can be considered cost-effective across the whole chronic HF population irrespective of LVEF status. This is based on the present analysis for the LVEF >40% population with an ICER of £14,429 and the cost-effectiveness results for the LVEF ≤40% population, which when combined, suggests empagliflozin is cost-effective across the full phenotype spectrum of chronic HF.

B.3.10 Exploring uncertainty

This appraisal represents a low risk to the committee in making a recommendation for routine commissioning for HF patients with an EF ≥40%. To provide the committee with confidence that empagliflozin represents value for money, and thus is suitable for routine commissioning, structural and parameter uncertainty has been assessed in the probabilistic and deterministic sensitivity analyses (B.3.10.1 to B.3.11). An internal and external validation of the cost-effective outcomes are provided in B.3.13.1.

Further, the BMJ EAG provided a thorough critique of the evidence in TA733, based on the EMPEROR-Reduced and PULSE (CPRD) study. The same cost utility and budget impact models have been re-used in this appraisal, however utilizing the pivotal EMPEROR-Preserved trial instead. PULSE has also been used in this appraisal to validate model outcomes. To expedite the assessment of this appraisal, the company has run scenario analyses for the issues that the BMJ identified during technical engagement (B.3.14). A rationale is provided where an issue hasn't been explored in this appraisal.

B.3.10.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was undertaken to translate uncertainty in the model parameters to decision uncertainty through simultaneous sampling of critical parameters from their respective distributions. The PSA encompassed parameters which inform the calculated rates of HHF, all-cause and CV mortality, as well as those informing the estimates of health state utilities, unit costs, AE and HHF utility decrements, and transition probabilities. The observed standard error was used to determine the probabilistic distribution of all parameters except costs and transition probabilities, where the standard error was calculated as a proportion of the mean value. For the costs, the standard error was assumed equal to 20% of the mean. However, for KCCQ-CSS transition probabilities, the standard error was assumed equal to 10% of the mean value to avoid iterations where the observed trend in the transition probabilities with empagliflozin + SoC versus SoC was inconsistent with the deterministic analyses (i.e., to ensure that relative probability of a given transition remains the same as in deterministic analysis across all probabilistic iterations).

Covariance matrices for parameters informing the rate of CV death, all-cause death, HHF and baseline utility estimates were included in the model. Using these, a Cholesky decomposition was performed and the resulting lower-triangular matrix was then used to jointly draw samples of these parameters from a normal distribution. All cost parameters were assigned a gamma distribution, while disutilities associated with AE and HHF were assigned the beta distribution. Details on the parameters, SEs, and assumptions are provided in Sections B.3.8.1 and B.3.8.2. One thousand PSA

iterations were run to ensure that stable estimates of the required model outputs were obtained.

Results of the PSA are summarised in the cost-effectiveness scatterplot (Figure 23). Each point on the chart represents a single probabilistic iteration of the model. Of the 1000 iterations, [redacted]% were below the WTP threshold of £25,000, indicating a high likelihood that empagliflozin + SoC would be costlier and more effective than SoC alone. The cost-effectiveness acceptability curve in Figure 23 illustrates the probability of empagliflozin + SoC being cost-effective at different willingness-to-pay thresholds. At willingness-to-pay thresholds of £20,000 and £30,000 per QALY, the probability of empagliflozin + SoC being cost-effective is always above [redacted]% ([redacted]% and [redacted]%, respectively) meaning that in the majority of the iterations of the PSA the ICER was below the £20,000 threshold. Those probabilities reflect the finding that empagliflozin + SoC is more costly compared to SoC, which is aligned with the findings of the deterministic analysis. Drug acquisition costs are the main driver of the observed higher healthcare costs for empagliflozin + SoC, with some cost savings occurring due to the lower rates of HHF for empagliflozin + SoC. The higher incremental costs for empagliflozin + SoC result in most PSA iterations being located in the top (north) quadrants, and specifically in the top right quadrant, demonstrating the higher incremental costs, but also the higher incremental QALYs for empagliflozin + SoC compared to SoC alone.

The ICER from the PSA converged at £14,564 per QALY (Table 45) which was comparable to the deterministic ICER of £14,429/QALY (Table 44). Overall, the results from the PSA are in accordance with the results of the deterministic analysis, indicating the robustness of the findings, which shows that the ICER is below the £20,000 threshold.

Table 45. Base case analysis: probabilistic results for empagliflozin as an add-on to standard care

Technology	Total costs (£), mean	Total LYG, mean	Total QALYs, mean	Incremental costs (£), mean	Incremental LYG, mean	Incremental QALYs - mean	ICER incremental (£/QALY)
SoC	[redacted]	[redacted]	[redacted]	-	-	-	-
Empagliflozin + SoC				[redacted]	[redacted]	[redacted]	£14, 564

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; SoC, standard of care.

Figure 23. Base case analysis: cost-effectiveness scatterplot

[redacted]

Abbreviations: QALY, Quality-adjusted Life Years; WTP, willingness to pay

Key: WTP threshold line drown at £25,000.

Figure 24. Base case analysis: cost-effectiveness acceptability curve

[redacted]

Abbreviations: QALY, quality-adjusted life-year; SoC, standard of care; WTP, willingness to pay

B.3.10.2 Deterministic sensitivity analysis

A deterministic (or one-way) sensitivity analysis was performed to assess how changes in one specified parameter at a time impact the predicted costs and outcomes of empagliflozin + SoC compared to SoC alone, and to identify the main drivers of cost-effectiveness in the model. The most influential parameter was the treatment effect of empagliflozin + SoC associated with HHF. When this parameter was set to zero, the ICER increased by 48% from the deterministic base case value of £14,429/QALY to a high of £21,339/QALY (Table 46). Other highly impactful parameters included the disutility for HHF, the treatment effect associated with empagliflozin + SoC for CV mortality, inclusion of treatment discontinuation for empagliflozin + SoC, and cost of treatment for HHF.

Table 46. Deterministic sensitivity analyses inputs and results

Scenario	Base Case Input	Alternative Input	Description	ICER per QALY
Base case	-	-	-	**£14,429**
Clinical Inputs				
CV & All-cause death: Distribution	Weibull	Exponential	Alternative distribution	£14,802
CV death: Treatment effect	-[redacted]	0	-0.Removing treatment effect	£15,297
All-cause death: Adjust with UK lifetable?	Yes	No	Lifetable adjustment	£14,685
HHF: Treatment effect	-[redacted]	0	Removing treatment effect	£21,339
Discontinuation: Distribution	Generalised Gamma	Weibull	Alternative distribution	£14,610
Include discontinuation?	Yes	No	Alternative scenario for discontinuation	£15,126
Costs and Resource Use				
Cost of HHF	[redacted]	[redacted]	Decrease by 20%	£14,884
		[redacted]	Increase by 20%	£13,974
Cost of CV death	[redacted]	[redacted]	Decrease by 20%	£14,557
		[redacted]	Increase by 20%	£14,301

Scenario	Base Case Input	Alternative Input	Description	ICER per QALY
Unit Costs of Disease monitoring	Multiple Values	Multiple Values	Decrease by 20%	£14,357
			Increase by 20%	£14,501
Monthly Cost of Disease Monitoring: KCCQ-CSS 1st Quartile	[redacted]	[redacted]	Decrease by 20%	£14,503
		[redacted]	Increase by 20%	£14,355
Monthly Cost of Disease Monitoring: KCCQ-CSS 2nd Quartile	[redacted]	[redacted]	Decrease by 20%	£14,492
		[redacted]	Increase by 20%	£14,366
Monthly Cost of Disease Monitoring: KCCQ-CSS 3rd Quartile	[redacted]	[redacted]	Decrease by 20%	£14,377
		[redacted]	Increase by 20%	£14,481
Monthly Cost of Disease Monitoring: KCCQ-CSS 4th Quartile	[redacted]	[redacted]	Decrease by 20%	£14,272
		[redacted]	Increase by 20%	£14,512
Cost of AE management	Multiple Values	Multiple Values	Decrease by 20%	£14,480
			Increase by 20%	£14,378
Utilities				
Utility: KCCQ-CSS 1st Quartile	[redacted]	[redacted]	Lower 95% CI	£14,512
		[redacted]	Upper 95% CI	£14,512
Utility: KCCQ-CSS 2nd Quartile	[redacted]	[redacted]	Lower 95% CI	£14,359
		[redacted]	Upper 95% CI	£14,499
Utility: KCCQ-CSS 3rd Quartile	[redacted]	[redacted]	Lower 95% CI	£14,491
		[redacted]	Upper 95% CI	£14,368
Utility: KCCQ-CSS 4th Quartile	[redacted]	[redacted]	Lower 95% CI	£14,633
		[redacted]	Upper 95% CI	£14,233
Disutility: HHF	[redacted]	[redacted]	Lower 95% CI	£16,102
		[redacted]	Upper 95% CI	£12,947
Disutility: AE	Multiple Values	Multiple Values	Lower 95% CI	£14,395
			Upper 95% CI	£14,447
Settings				
Time horizon	Lifetime	10 years	Lower range	£16,890
		20 years	Upper range	£13,860
Discount rate: cost	3.5%	[redacted]	Lower range	£16,033
		[redacted]	Upper range	£14,273
Discount rate: health	3.5%	[redacted]	Lower range	£11,758

Scenario	Base Case Input	Alternative Input	Description	ICER per QALY
		[redacted]	Upper range	£15,591

Abbreviations: AE, adverse events; CV, cardiovascular; HR, hazard ratio; HHF, hospitalisation for heart failure; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire Clinical Summary Score; UK, United Kingdom.

A tornado diagram showing the impact of model parameters on the ICER is provided in Figure 25. In conclusion, the ICER for empagliflozin + SoC relative to SoC alone remained below £20,000/QALY across all parameter variations except for HHF treatment effect (£21,339/QALY), with the majority of analyses resulting in ICERs below £18,000/QALY.

Figure 25.Tornado diagram

Abbreviations: AE, adverse events; CV, cardiovascular; hHF, hospitalisation for heart failure; ICER, incremental cost-effectiveness ratio; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire-Clinical Summary Score

B.3.10.3 Scenario analysis

Scenario analyses were run to test the validity of the structural assumptions, including the choice of parametric models to inform mortality and time-to-treatment discontinuation extrapolations, and the impact of utility age adjustment. **Error! Reference source not found.** provides a description of each scenario along with its resulting ICER.

Table 47.Scenario Analyses: ICERs for empagliflozin as add-on to standard care compared to standard care alone

Scenario	Description	ICER (Cost in £ / QALY)	% Change Relative to Base Case ICER
Base Case		£14,429	N/A
One Inflection Point	Use the KCCQ quartile transition matrix used for months 4 to 8 in the model base case from month 4 to the end of the time horizon.	£22,000	52.47
Mortality: Log-normal	Extrapolate CV and all-cause mortality outcomes using a log-normal distribution.	£15,752	9.17
Mortality: Log-logistic	Extrapolate CV and all-cause mortality outcomes using a log-logistic distribution.	£15,030	4.17
Mortality: Exponential	Extrapolate CV and all-cause mortality outcomes using an exponential distribution.	£14,802	2.59
Mortality: Generalised Gamma	Extrapolate CV and all-cause mortality outcomes using a generalised gamma distribution.	£14,473	0.30
Mortality: Gompertz	Extrapolate CV and all-cause mortality outcomes using a Gompertz distribution.	£17,553	21.65
Discontinuation: Weibull	Extrapolate time to discontinuation for empagliflozin using a Weibull distribution.	£14,610	1.25
Discontinuation: Log-normal	Extrapolate time to discontinuation for empagliflozin using a log-normal distribution.	£14, 808	2.63
Discontinuation: Log-logistic	Extrapolate time to discontinuation for empagliflozin using a log-logistic distribution.	£14,735	2.12
Discontinuation: Generalised Gamma	Extrapolate time to discontinuation for empagliflozin using an exponential distribution.	£14,565	0.94
Discontinuation: Gompertz	Extrapolate time to discontinuation for empagliflozin using a Gompertz distribution.	£14,592	1.13

Scenario	Description	ICER (Cost in £ / QALY)	% Change Relative to Base Case ICER
Utility: Age Adjustment Off	Use utility data as collected in the trial (KCCQ 4: [redacted]; KCCQ 3: [redacted]; KCCQ 2: [redacted]; KCCQ 1: [redacted]), without adjusting KCCQ 4 to be equal to UK general population utility.	£12,964	-10.15
Non-CV Death Costs	Assuming that non-CV deaths incur the same costs as CV deaths.	£14,958	3.67

Abbreviations: CV, cardiovascular; ERG, evidence review group; HHF, hospitalisation with heart failure; ICER, incremental cost-effectiveness ratio; ITT, intention to treat population; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire-Clinical Summary Score; KM, Kaplan-Meier; LVEF, left ventricular ejection fraction; OS, overall survival; SoC, standard of care; SGLT2i, sodium-glucose co-transporter-2 inhibitors; TP, transition probabilities; UK, United Kingdom.

B.3.10.4 Summary of the assessment of uncertainty

The overall agreement between deterministic and probabilistic results suggests that the economic model is adequately linear. Furthermore, the results of the deterministic and probabilistic sensitivity analyses, and scenario analyses, demonstrate that ICER is robust with respect to changes in model inputs (Section B.3.10.2 and B.3.10.1, B.3.10.3, respectively).

In the deterministic sensitivity analyses, the most influential parameter was the treatment effect of empagliflozin + SoC on the risk of HHF. The assumption of no beneficial effect of empagliflozin + SoC on the reduction of risk of HHF relative to SoC increases the ICER by 48% from the base case value of £14,429/QALY to a high of £21,339/QALY. Other highly impactful parameters included the disutility for HHF, the treatment effect associated with empagliflozin + SoC for CV mortality, inclusion of treatment discontinuation for empagliflozin + SoC and cost of treatment for HHF.

B.3.11 Subgroup analysis

As mentioned in Section B.2.7, the use of the ITT population for the CEA of empagliflozin in chronic HF (EF >40%) was the most statistically robust approach since EMPEROR-Preserved was not powered to evaluate the treatment effect in subgroups. In the EMPEROR-Preserved trial, a reduction in HHF or adjudicated CV death (composite primary outcome) was shown across multiple subgroups, including age (<70 />70 years), sex (male/female), race (White, Black, Asian, other), body mass index and prior therapies (ARNI/no ARNI). Given that the results for the subgroups

were generally consistent with the confirmatory analyses, only the ITT population was considered in the economic analysis. Results of the clinically relevant pre-specified subgroups can be found in Appendix E.

B.3.12 Benefits not captured in the QALY calculation

The impact on the carer has not been included in the QALY calculation.

B.3.13 Validation

B.3.13.1 Validation of cost-effectiveness analysis

B.3.13.1.1 Internal validity

Internal validation was undertaken to assess the model's ability to accurately predict the observed outcomes from the EMPEROR-Preserved trial. The rates of HHF, CV death and non-CV death observed during the average trial follow-up of 26 months were compared with the economic model predictions over an 18-month time horizon. Figure 26 shows the observed and predicted CV death, all-cause death and HHF rates per 100 patient-years, respectively, obtained with the simple risk equations containing only treatment and time-varying KCCQ-CSS health states as predictors as well as those including the full set of predictors in the ITT population. The reduced set of equation resulted in estimates that were closer to the estimates from the observed data, meaning that the model's predictions were more validated when using the reduced list of predictors compared to using the extended list of predictors. Thus, the reduced list of predictors was used as the base case.

Figure 26. Comparison of observed and model-predicted event rates in the ITT population: CV death, non-CV death, and HHF rates per 100 patient-years

[redacted]

B.3.13.1.2 External validity

As previously mentioned in Section B.2.1.2, the PULSE study was the best available evidence to support this decision problem and considered appropriate for the validity of the event rates projected by the model (48). The external validity of the economic model predictions were checked against the observed rates of the PULSE study (48) and the long-term predictions of non-CV mortality were closely aligned with the observed rates of non-CV death in the PULSE study.

HHF and CV death rates predicted by the empagliflozin model were notably higher than those observed in the PULSE study. the difference in HHF and CV-mortality rates observed in EMPEROR-R and PULSE is likely due inaccurate recording of events in PULSE rather than a clinically meaningful difference in the patient characteristics. Although the differences in outcomes could be due to EMPEROR-R being enriched with a 'sicker' population, this conclusion is very uncertain. The significant amount of

missing data observed in the baseline characteristics in PULSE makes it impossible to characterise a typical UK HF patient. Since PULSE relied on Office for National Statistics (ONS) mortality data for primary cause of death, the rate of CV death is likely to be under-recorded compared to EMPEROR-Preserved, where CV death was an adjudicated endpoint. The summary of the comparison between the predicted model outcomes and the observed rates from PULSE are presented in Table 48.

Table 48: Model-predicted vs observed rates per 100 patient-years in PULSE

Characteristics	PULSE SoC (Events per 100 patient-years)	Model simulation Pulse - Placebo (60 months) (Events per 100 patient-years)
Non-CV mortality	[redacted]	[redacted]
CV mortality	[redacted]	[redacted]
All-cause mortality	[redacted]	[redacted]
HF hospitalisation	[redacted]	[redacted]

Abbreviations: CV, cardiovascular; HF, heart failure; SoC, standard of care.

B.3.14 Uncertainty explored during Technical Engagement for TA733

During Technical Engagement for TA773, the BMJ EAG identified 12 issues or areas of uncertainty in the clinical and economic evidence case. To leverage the work already completed by the BMJ EAG for TA733, the same issues have been explored in this appraisal for HF patients with an EF>40%.*Table **49*** describes each key issue and how it has been addressed in this appraisal. For any issue that has not been addressed in this appraisal, a rationale is provided. **Table *49***Table 50 shows that, like in TA733, the ICERs are robust. This exploration of uncertainty provides the committee with further confidence that empagliflozin in those with an EF >40% represents value for money.

Table 49. Description of issues identified by the BMJ EAG during Technical Engagement for TA733

Issue	Description	Included in this appraisal	Rationale
Issue 1 Generalisability older subgroup	Uncertainty around the generalisability of the results from EMPEROR-Reduced to the older heart failure with reduced ejection fraction population expected in clinical practice. The ERG considered the >65 years and <65 years subgroup.	No	The PULSE study reported a mean age of 73 years in the HFpEF population, which was comparable to 72 years in the ITT population of EMPEROR-Preserved. Therefore, there is limited evidence that using a subgroup by age as the preferred base case in the economic model would offer a better representation of UK clinical practice than the ITT population.
Issue 2 Generalisability Europe subgroup	Uncertainty around the difference in efficacy of empagliflozin compared with standard of care in the Europe subgroup of EMPEROR-Reduced	No	As stated in B.2.7, the Europe subgroup is unlikely to be a better representation of UK clinical practice than the ITT population. This is because the baseline characteristics and key outcomes for the Europe subgroup were comparable to the ITT population, as reported in Appendix E. Thus, if the Europe subgroup was used in the base case to estimate the ICER, it is likely to be similar to the ITT population. Further, the use of data from Europe subgroup to assess generalisability is not appropriate and could contribute to existing ethnic inequalities in health (149), contrary to the NICE's Social Value Judgments and the Equality Act 2010 (race is one of the protected characteristics) (14). The Europe subgroup of EMPEROR-Preserved was 99.0% white and therefore not representative of the multi-ethnic UK population, which consists of 86% white, 3.3% black, 7.5% Asian and 3.2% other (139, 150)
Issue 3 Comparator	Uncertainty around the efficacy of empagliflozin compared with dapagliflozin	No	Dapagliflozin is not a relevant comparator for this appraisal
Issue 4 Waning of treatment effect	The EAG asked that the company conducted scenario analyses where the effect of empagliflozin on KCCQ-CSS (sustained by the combination of the proportion of	Yes	The Company base case for this appraisal assumed that the benefit of empagliflozin on KCCQ is maintained from the end of the trial period (Week 52) for the remaining duration of a patient's life. Although treatment waning is explored in Table 50, it is not appropriate to adopt this as the preferred base case for decision making. This is because there is limited evidence to suggest

Issue	Description	Included in this appraisal	Rationale
	patients in the better KCCQ-CSS states in the empagliflozin arm at month 8 and the low probability of disease progression for both SoC and empagliflozin arms in month 9+) was waned over time in the model		that there is a waning effect. The EMPA-REG study – a CVOT trial of T2DM patients with multiple CV risk factors –showed that all-cause mortality, CV-mortality and heart failure hospitalisation was sustained for 3.1 years (Shown in Figure 1 of reference (130)). Although KCCQ was not collected as part of the EMPA-REG study, it demonstrated a sustained consistent effect for empagliflozin over time. This assumption is consistent with prior NICE appraisals for products in the same drug class. In the dapagliflozin appraisal in HFrEF (TA679), the Committee concluded that "there was no evidence for or against treatment waning in the long-term. Clinical experts and stakeholders confirmed that treatment with dapagliflozin would likely be lifelong".
Issue 5 Assuming a constant rate of HHF over time	A constant rate of hospitalisation was assumed in the TA773 submission; however, the EAG noted that this was misaligned with increasing hospitalisation rates over time as observed for patients with HF in clinical practice.	Yes	A scenario analysis was conducted to fit different models to show that the hazards are either similar to the Poisson model or they are clinically implausible. A constant increase in the rate of HHF was added to the HHF rate for each model cycle (1 month) at a constant increase over time.
Issue 6 Overestimation of HHF in the UK population	The EAG was concerned that the HHF was overestimated in the UK population when using the trial data compared to PULSE. .	Yes	A scenario was implemented in the model to adjust the Poisson model used to estimate HHF by fitting a joint regression model of the individual patient-level data from PULSE and the EMPEROR-Preserved trial. This was done with a negative binomial model for the total number of HHF events observed for patients with follow-up duration as an offset, yielding a 0.44 rate-ratio adjustment factor when comparing PULSE EF >40% patients and the ITT Placebo arm of the EMPEROR-Preserved trial
Issue 7 Modelling of mortality	The EAG noted that there is a risk for double counting the treatment effect on mortality because mortality is different for each KCCQ state, but the treatment effect also impacts	Yes	A scenario was implemented in the model where it was assumed that empagliflozin has no direct or indirect survival benefit over SoC, including the residency in KCCQ-CSS health states. For this scenario, a Weibull function was fitted to the OS KM data for the placebo arm of the EMPEROR-Preserved trial with no treatment effect (i.e., CV or non-CV mortality was assumed to be the same) or KCCQ-CSS predictors for the ITT population.

Issue	Description	Included in this appraisal	Rationale
	mortality through the treatment coefficients in the risk equations.		
Issue 8 Overestimation of mortality in the UK population	In TA733, the EAG noted differences in the number of CV-deaths and all-cause deaths observed in PULSE vs the Company's model for the ≥65 years subgroup. When an adjustment to CV-mortality and all-cause mortality to reflect the PULSE HFrEF (prevalent and incident patients at index) was applied to the >65 years subgroup in the CE model, the length of patient's life was extended life-years and resulted in higher reduction in number of HHFs. This in turn, resulted in a decrease in the ICER from £4,717 to £4,325/QALY gained.	No	In this appraisal, the ITT population was used to estimate cost effectiveness in the CE model. This is because the mean age in the trial was consistent with the mean age of patients in PULSE for patients with an EF >40%. (73 years vs 72 years, respectively).
Issue 9 HHF disutility	An additional technical engagement scenario was considered (Issue 9) which assumed the HHF disutility changed from 1 year (base case) to 3 months (scenario).	No	This scenario was not implemented based on feedback from the patient experts at the TA773 committee meeting (1) who unanimously agreed that the disutility could last for 1 year.
Issue 10 Quality of life	During clarification questions, the EAG requested an	No	Please see response to issue 1.

Issue	Description	Included in this appraisal	Rationale
regressions for the UK population	ICER for >65 subgroup. During Technical Engagement, the ERG requested the Company to re-estimate the QoL regression analysis for the >65 years subgroup, as the ITT population had been previously used with a binary predictor for <65 and >65 years.		
Issue 11 Sex distribution underlying utility estimates	In TA733, the utility value, 0.7740. applied to KCCCQ-CSS health state 4 was taken from Sullivan et al. This value was based on a UK dataset where 52% were female. This differed from the gender distribution in EMPEROR-R and PULSE (24% females and 35% females, respectively).	No	Compared with EMPEROR-Reduced, the sex distribution in EMPEROR-Preserved more closely matched the distribution reported in Sullivan et al (44.65% vs 52%, respectively). Thus, exploration of this scenario is not relevant for this appraisal
Issue 12 Quality of life gains in EMPEROR-Reduced	Scenario combining issue 4,6,8,9,10, and 11	No	N/A as not all scenarios have been explored in this appraisal.

Table 50. Scenario Analyses: ICERs for each issue explored in this appraisal, based on the scenarios run by the BMJ EAG for TA733

Scenario	Description		ICER (Cost in £ / QALY)	% Change Relative to Base Case ICER
Issue 4 Waning of Treatment Effect	Set the proportion of patients in the KCCQ-CSS quartiles under the empagliflozin arm equal to those proportions in the SoC arm at 5, 3, 2 and 1 years	5 years	£16,139	11.85
		3 years	£17,187	19.11
		2 years	£17,457	20.99
		1 year	£16,985	17.71
	The TP between KCCQ-CSS quartiles for treatment arm are set to the TPs for the SoC arm after 8 months		£32,482	125.12
Issue 5 Increased Risk of HHF Over Time	A constant increase to the HHF rate for each treatment cycle. The increase varies by KCCQ-CSS quartile: KCCQ-CSS 1st quartile: 0.4% KCCQ-CSS 2nd quartile: 0.3% KCCQ-CSS 3rd quartile : 0.2% KCCQ-CSS 4th quartile : 0.1%		£13,861	-3.94
Issue 6 Overestimation of hospitalisation for heart failure in the UK population	Rate-ratio adjustment factor (0.44) when comparing PULSE EF > 40% patients and the ITT Placebo arm of the EMPEROR-Preserved trial		£18,288	26.74
Issue 7 Remove the direct/indirect benefit in mortality that is realised through KCCQ state membership	A fitted Weibull function to the CV-mortality or AC-mortality data for the placebo arm of the EMPEROR-Preserved trial with no treatment effect or KCCQ-CSS predictors for the ITT population		£26,424	83.13

Abbreviations: CV, cardiovascular; EAG, evidence assessment group; HHF, hospitalisation with heart failure; ICER, incremental cost-effectiveness ratio; ITT, intention to treat population; KM, Kaplan-Meier; KCCQ, Kansas City Cardiomyopathy Questionnaire; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire – clinical summary score; OS, overall survival; SoC, standard of care; SGLT2i, sodium-glucose co-transporter-2 inhibitors; TP, transition probabilities; UK, United Kingdom.

The value in the parentheses for the SGLT2i scenario corresponds to the probabilistic ICER. All the remaining ICER estimates are deterministic

B.3.15 Interpretation and conclusions of economic evidence

The CEM for the economic evaluation of empagliflozin + SoC for chronic HF patients with EF >40% builds on the modelling approach previously accepted by the NICE committee for empagliflozin + SoC for patients with HFrEF (1). Model inputs were primarily derived from the EMPEROR-Preserved trial, including inputs for baseline characteristics, health state transition probabilities, health state utility values, disutilities associated with clinical events, survival equations, risk equations, AE incidence rates, and treatment discontinuation rates. Additional model inputs for disutilities of AE, unit costs and resource use were identified from the published literature or from NHS National Reference Costs. The model was able to reproduce the EMPEROR-Preserved trial results over the mean trial follow-up period of 26 months and was used to extrapolate those results to a lifetime horizon.

In the base case analysis, over the lifetime horizon, patients treated with empagliflozin + SoC experienced a lower rate of HHF ([redacted] per 100 PY vs [redacted] per 100 PY on SoC) and CV death ([redacted] per 100 PY vs [redacted] per 100 PY on SoC) compared to those treated with SoC alone. The difference in the rate of non-CV death between empagliflozin + SoC and SoC arms was minimal ([redacted] per 100 PY vs. [redacted] per 100 PY). Reduction in clinical event rates with empagliflozin + SoC compared to SoC was the key driver of the incremental benefits, while incremental costs were largely attributable to empagliflozin + SoC drug costs which were in part due to the longer survival and treatment duration of patients receiving empagliflozin as an add-on to SoC (average time of receiving empagliflozin + SoC was [redacted] years). The base case analysis estimated a deterministic ICER of £14,429 and a probabilistic ICER of £14,564 per QALY gained, with both < £20,000, suggesting that empagliflozin + SoC offers a good use of NHS resources and should be preferred over SoC alone based on usual threshold values.

Sensitivity and scenario analyses showed that the CEM was robust to variation in model parameters. The probabilistic and deterministic base case results were closely aligned, with [redacted] % of the iterations below the WTP threshold of £25,000, suggesting

a high probability of empagliflozin + SoC being costlier and more effective than SoC alone.

The most influential parameter in the deterministic sensitivity analysis was the treatment effect of empagliflozin + SoC associated with HHF. When this parameter was set to 0, the ICER increased 48% from the base case value of £14,429/QALY to a high of £21,339/QALY (Table 46). However, this is an unrealistic assumption as it assumes no treatment benefit with empagliflozin, which is not supported by the clinical trial data primary endpoint. The estimate was also statistically significant and formed the basis of the positive opinion and MA for empagliflozin in this population. Other highly impactful parameters included the disutility for HHF, the treatment effect associated with empagliflozin + SoC for CV mortality, inclusion of treatment discontinuation for empagliflozin + SoC and cost of treatment for HHF.

Results of the scenario analyses indicate that the ICER is not significantly affected by structural assumptions including the choice of parametric distribution for mortality or treatment discontinuation, utility age adjustment or cost of non-CV death, although it is highly sensitive to assumptions regarding the treatment effect through mortality, with an ICER of £26,424 per QALY. However, the latter scenario reflects an unrealistic assumption as it assumes no treatment benefit with empagliflozin which is not supported by the clinical trial data. The treatment effect for CV mortality is in the model because a numerical difference between the two comparators is clinically plausible for this outcome and the clinical data from the EMPEROR-Preserved trial shows a trend in favour of empagliflozin for CV mortality benefit. Finally, changing the model to align transition probabilities after month 8 for both the empagliflozin treatment and SoC arms (i.e., treatment waning scenario) had the largest impact on the ICER with an ICER of £34,482 per QALY.

The results of the analysis should be interpreted considering its limitations. First, the EMPEROR-Preserved trial offered only short-term data (median follow-up time was 26 months in the EMPEROR-Preserved trial) and therefore, long-term outcomes had to be extrapolated at the expense of uncertainty; however, the sensitivity analyses

indicated that the choice of the parametric model did not have a significant impact on the estimated ICER. Next, no reliable external resource is available for estimating the rate of clinical events beyond the trial time horizon as these outcomes are closely tied to underlying SoC. However, sensitivity analyses varying the distribution of the risk equations did not affect conclusions about the cost-effectiveness of empagliflozin + SoC compared to SoC, even though there was some variation in the incremental health benefits. Furthermore, the economic analysis assumes that the HF event rates observed in UK clinical practice mirror those observed in the EMPEROR-Preserved trial. The relevance of the trial to the UK clinical practice is strengthened by the protocol requirement for patients to receive stable doses of guideline-recommended HF therapies at baseline. Last, the analysis only captured direct medical costs. Costs associated with non-direct medical resources and indirect costs (i.e., productivity loss) were not considered given the perspective of this analysis.

In conclusion, the CEA presented here demonstrates that empagliflozin represents a cost-effective use of NHS resources as an add-on to standard care therapy for the treatment of chronic HF patients with EF >40%. In addition, when considering the favourable ICER from the TA773 submission and the combined weighted HF ICER, empagliflozin demonstrates value for money across the full phenotype spectrum of chronic HF, irrespective of EF.

B.4 References

1. National Institute for Health and Care Excellence. Empagliflozin for treating chronic heart failure with reduced ejection fraction [ID3826] 2021 [Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10719].
2. European Medicines Agency. Jardiance : EPAR - Summary for the public 2022 [Available from: https://www.ema.europa.eu/en/medicines/human/summaries-opinion/jardiance-0].
3. National Institute for Health and Care Excellence. Chronic heart failure in adults: diagnosis and management NICE guideline [NG106] 2018 [Available from: https://www.nice.org.uk/guidance/ng106].
4. Packer M, Butler J, Filippatos G, Zannad F, Ferreira JP, Zeller C, et al. Design of a prospective patient-level pooled analysis of two parallel trials of empagliflozin in patients with established heart failure. Eur J Heart Fail. 2020;22(12):2393-8.
5. Anker SD, Butler J, Filippatos G, Ferreira JP, Bocchi E, Böhm M, et al. Empagliflozin in heart failure with a preserved ejection fraction. New England Journal of Medicine. 2021;385(16):1451-61.
6. Packer M, Anker SD, Butler J, Filippatos G, Pocock SJ, Carson P, et al. Cardiovascular and renal outcomes with empagliflozin in heart failure. New England Journal of Medicine. 2020;383(15):1413-24.
7. National Institute for Health and Care Excellence. Empagliflozin in combination therapy for treating type 2 diabetes Technology appraisal guidance [TA336] 2015 [Available from: https://www.nice.org.uk/guidance/ta336].
8. National Institute for Health and Care Excellence. Canagliflozin, dapagliflozin and empagliflozin as monotherapies for treating type 2 diabetes Technology appraisal guidance [TA390] 2016 [Available from: https://www.nice.org.uk/guidance/ta390].
9. Conrad N, Judge A, Tran J, Mohseni H, Hedgecott D, Crespillo AP, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. The Lancet. 2018;391(10120):572-80.
10. Lawson Claire A, Zaccardi F, Squire I, Okhai H, Davies M, Huang W, et al. Risk Factors for Heart Failure. Circulation: Heart Failure. 2020;13(2):e006472.
11. Moscelli G, Siciliani L, Gutacker N, Cookson R. Socioeconomic inequality of access to healthcare: Does choice explain the gradient? J Health Econ. 2018;57:290-314.
12. McCartney G, Hart C, Watt G. How can socioeconomic inequalities in hospital admissions be explained? A cohort study. BMJ Open. 2013;3(8):e002433.
13. National Institute for Health and Care Excellence. The NICE strategy 2021 to 2026 2021 [Available from: https://www.nice.org.uk/about/who-we-are/corporate-publications/the-nice-strategy-2021-to-2026].
14. National Institute for Health and Care Excellence. Our Priniciples: Social Value Judgements 2021 [Available from: https://www.nice.org.uk/about/who-we-are/our-principles].
15. The Health Foundation. Build Back Fairer: The COVID-19 Marmot Review. 2020.
16. Ndefo UA, Anidiobi NO, Basheer E, Eaton AT. Empagliflozin (Jardiance): A Novel SGLT2 Inhibitor for the Treatment of Type-2 Diabetes. P T. 2015;40(6):364-8.

17. Lopaschuk GD, Verma S. Mechanisms of Cardiovascular Benefits of Sodium Glucose Co-Transporter 2 (SGLT2) Inhibitors: A State-of-the-Art Review. JACC Basic Transl Sci. 2020;5(6):632-44.
18. Cowie MR, Fisher M. SGLT2 inhibitors: mechanisms of cardiovascular benefit beyond glycaemic control. Nature Reviews Cardiology. 2020;17(12):761-72.
19. Boehringer Ingelheim Ltd. Summary of Product Characteristics: Jardiance (empagliflozin) 10 mg film-coated tablets (GB). 2022. Available from: https://www.medicines.org.uk/emc/product/5441/smpc.
20. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. European heart journal. 2021;42(36):3599-726.
21. Murphy SP, Ibrahim NE, Januzzi JL, Jr. Heart Failure With Reduced Ejection Fraction: A Review. JAMA. 2020;324(5):488-504.
22. Caraballo C, Desai NR, Mulder H, Alhanti B, Wilson FP, Fiuzat M, et al. Clinical Implications of the New York Heart Association Classification. J Am Heart Assoc. 2019;8(23):e014240.
23. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., Drazner MH, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013;62(16):e147-239.
24. Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. Nat Rev Cardiol. 2016;13(6):368-78.
25. Kenchaiah S, Evans JC, Levy D, Wilson PWF, Benjamin EJ, Larson MG, et al. Obesity and the Risk of Heart Failure. New England Journal of Medicine. 2002;347(5):305-13.
26. Matthews JC, Dardas TF, Dorsch MP, Aaronson KD. Right-sided heart failure: diagnosis and treatment strategies. Curr Treat Options Cardiovasc Med. 2008;10(4):329-41.
27. Mangini S, Pires PV, Braga FG, Bacal F. Decompensated heart failure. Einstein (Sao Paulo). 2013;11(3):383-91.
28. Nhfa Csanz Heart Failure Guidelines Working Group, Atherton JJ, Sindone A, De Pasquale CG, Driscoll A, MacDonald PS, et al. National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand: Guidelines for the Prevention, Detection, and Management of Heart Failure in Australia 2018. Heart Lung Circ. 2018;27(10):1123-208.
29. Voelkel NF, Quaife RA, Leinwand LA, Barst RJ, McGoon MD, Meldrum DR, et al. Right Ventricular Function and Failure. Circulation. 2006;114(17):1883-91.
30. American Heart Association. Ejection Fraction Heart Failure Measurement 2021 [Available from: https://www.heart.org/en/health-topics/heart-failure/diagnosing-heart-failure/ejection-fraction-heart-failure-measurement].
31. Raphael C, Briscoe C, Davies J, Ian Whinnett Z, Manisty C, Sutton R, et al. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. Heart. 2007;93(4):476-82.

32. Bhatt DL, Szarek M, Steg PG, Cannon CP, Leiter LA, McGuire DK, et al. Sotagliflozin in Patients with Diabetes and Recent Worsening Heart Failure. New England Journal of Medicine. 2020;384(2):117-28.
33. ClinicalTrials.gov. Dapagliflozin Evaluation to Improve the LIVEs of Patients With PReserved Ejection Fraction Heart Failure (DELIVER) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT03619213].
34. ClinicalTrials.gov. Spironolactone Initiation Registry Randomized Interventional Trial in Heart Failure With Preserved Ejection Fraction (SPIRRIT) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT02901184].
35. Massie BM, Carson PE, McMurray JJ, Komajda M, McKelvie R, Zile MR, et al. Irbesartan in patients with heart failure and preserved ejection fraction. N Engl J Med. 2008;359(23):2456-67.
36. Pitt B, Pfeffer MA, Assmann SF, Boineau R, Anand IS, Claggett B, et al. Spironolactone for heart failure with preserved ejection fraction. N Engl J Med. 2014;370(15):1383-92.
37. Solomon SD, McMurray JJV, Anand IS, Ge J, Lam CSP, Maggioni AP, et al. Angiotensin-Neprilysin Inhibition in Heart Failure with Preserved Ejection Fraction. N Engl J Med. 2019;381(17):1609-20.
38. Yusuf S, Pfeffer MA, Swedberg K, Granger CB, Held P, McMurray JJ, et al. Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: the CHARM-Preserved Trial. Lancet. 2003;362(9386):777-81.
39. Boehringer Ingelheim International GmbH. UK EMPEROR-Preserved ad board report [data on file]. 2021.
40. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail. 2016;18(8):891-975.
41. British Heart Foundation. Heart statistics: UK factsheet 2021 [Available from: https://www.bhf.org.uk/-/media/files/research/heart-statistics/bhf-cvd-statistics-uk-factsheet.pdf?la=en&rev=b9ac86ab1ef349c2ae73177bb176174d&hash=B3B23C403C3B08EC42AC5D3B928C1CEE54C67DCA].
42. Boehringer Ingelheim International GmbH. Understanding the heart failure patient journey [data on file]. 2020.
43. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail. 2020;22(8):1342-56.
44. British Lung Foundation. Chronic obstructive pulmonary disease (COPD) statistics 2020 [Available from: https://statistics.blf.org.uk/copd].
45. Mamas MA, Sperrin M, Watson MC, Coutts A, Wilde K, Burton C, et al. Do patients have worse outcomes in heart failure than in cancer? A primary care-based cohort study with 10-year follow-up in Scotland. Eur J Heart Fail. 2017;19(9):1095-104.
46. Morton G, Masters J, Cowburn PJ. Multidisciplinary team approach to heart failure management. Heart. 2018;104(16):1376-82.

47. Masters H, Freeman J, Dixon S. Variable structure and provision of guideline-based care in specialist heart failure centres in the UK: a survey of 100 health professionals. British Journal of Cardiac Nursing. 2020;15(5).
48. Boehringer Ingelheim International GmbH. Incidence prevalence and resource utilisation of heart failure in England: a linked descriptive analysis of English primary and secondary care – PULSE study. [Data on File]. 2021.
49. Cancer Research UK. Cancer incidence statistics 2017 [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence#heading-One].
50. Rangaswami J, Bhalla V, Blair JEA, Chang TI, Costa S, Lentine KL, et al. Cardiorenal Syndrome: Classification, Pathophysiology, Diagnosis, and Treatment Strategies: A Scientific Statement From the American Heart Association. Circulation. 2019;139(16):e840-e78.
51. Dhingra R, Vasan RS. Diabetes and the risk of heart failure. Heart failure clinics. 2012;8(1):125-33.
52. Cherney DZI, Repetto E, Wheeler DC, Arnold SV, MacLachlan S, Hunt PR, et al. Impact of Cardio-Renal-Metabolic Comorbidities on Cardiovascular Outcomes and Mortality in Type 2 Diabetes Mellitus. Am J Nephrol. 2020;51(1):74-82.
53. Savarese G, Stolfo D, Sinagra G, Lund LH. Heart failure with mid-range or mildly reduced ejection fraction. Nat Rev Cardiol. 2022;19(2):100-16.
54. Lofman I, Szummer K, Hagerman I, Dahlstrom U, Lund LH, Jernberg T. Prevalence and prognostic impact of kidney disease on heart failure patients. Open Heart. 2016;3(1):e000324.
55. House AA, Wanner C, Sarnak MJ, Piña IL, McIntyre CW, Komenda P, et al. Heart failure in chronic kidney disease: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int. 2019;95(6):1304-17.
56. Lawson CA, Jones PW, Teece L, Dunbar SB, Seferovic PM, Khunti K, et al. Association Between Type 2 Diabetes and All-Cause Hospitalization and Mortality in the UK General Heart Failure Population: Stratification by Diabetic Glycemic Control and Medication Intensification. JACC Heart Fail. 2018;6(1):18-26.
57. Kenny HC, Abel ED. Heart Failure in Type 2 Diabetes Mellitus. Circ Res. 2019;124(1):121-41.
58. Yang H, Negishi K, Otahal P, Marwick TH. Clinical prediction of incident heart failure risk: a systematic review and meta-analysis. Open Heart. 2015;2(1):e000222.
59. Hamo C, Kwak L, Florido R, Echouffo-Tcheugui J, Blumenthal R, Loehr LR, et al. Abstract 14768: Modifiable Risk Factor Control and Risk for Heart Failure With Preserved Ejection Fraction (HFpEF) versus Reduced Ejection Fraction (HFrEF): The Atherosclerosis Risk in Communities Study (ARIC). Circulation. 2019;140(1).
60. Kubicki DM, Xu M, Akwo EA, Dixon D, Munoz D, Blot WJ, et al. Race and Sex Differences in Modifiable Risk Factors and Incident Heart Failure. JACC Heart Fail. 2020;8(2):122-30.
61. Zhang J, Hobkirk J, Carroll S, Pellicori P, Clark AL, Cleland JG. Exploring quality of life in patients with and without heart failure. Int J Cardiol. 2016;202:676-84.
62. Aguilar D. Heart Failure, Diabetes Mellitus, and Chronic Kidney Disease: A Clinical Conundrum. Circulation: Heart Failure. 2016;9(7).

63. Taylor CJ, Ordóñez-Mena JM, Roalfe AK, Lay-Flurrie S, Jones NR, Marshall T, et al. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000-2017: population based cohort study. BMJ. 2019;364:l223.
64. Taylor CJ, Ryan R, Nichols L, Gale N, Hobbs FR, Marshall T. Survival following a diagnosis of heart failure in primary care. Family Practice. 2017;34(2):161-8.
65. van Riet EE, Hoes AW, Wagenaar KP, Limburg A, Landman MA, Rutten FH. Epidemiology of heart failure: the prevalence of heart failure and ventricular dysfunction in older adults over time. A systematic review. Eur J Heart Fail. 2016;18(3):242-52.
66. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. Circulation. 2019;139(10):e56-e528.
67. Boonman-de Winter LJ, Rutten FH, Cramer MJ, Landman MJ, Liem AH, Rutten GE, et al. High prevalence of previously unknown heart failure and left ventricular dysfunction in patients with type 2 diabetes. Diabetologia. 2012;55(8):2154-62.
68. Lin AH, Chin JC, Sicignano NM, Evans AM. Repeat Hospitalizations Predict Mortality in Patients With Heart Failure. Mil Med. 2017;182(9):e1932-e7.
69. Cowie MR. The heart failure epidemic: a UK perspective. Echo Res Pract. 2017;4(1):R15-R20.
70. Hancock HC, Close H, Fuat A, Murphy JJ, Hungin AP, Mason JM. Barriers to accurate diagnosis and effective management of heart failure have not changed in the past 10 years: a qualitative study and national survey. BMJ Open. 2014;4(3):e003866.
71. Dulai R, Sheikh AS, Qureshi A, Katechia S, Peysakhova Y, Johns M, et al. Prevalence, clinical characteristics and outcomes of HF with preserved versus reduced ejection fraction. The British Journal of Cardiology. 2016;23(1).
72. Chan MM, Lam CS. How do patients with heart failure with preserved ejection fraction die? Eur J Heart Fail. 2013;15(6):604-13.
73. Vaduganathan M, Patel RB, Michel A, Shah SJ, Senni M, Gheorghiade M, et al. Mode of Death in Heart Failure With Preserved Ejection Fraction. J Am Coll Cardiol. 2017;69(5):556-69.
74. Crespo-Leiro MG, Anker SD, Maggioni AP, Coats AJ, Filippatos G, Ruschitzka F, et al. European Society of Cardiology Heart Failure Long-Term Registry (ESC-HF-LT): 1-year follow-up outcomes and differences across regions. Eur J Heart Fail. 2016;18(6):613-25.
75. Waldum B, Westheim AS, Sandvik L, Flonaes B, Grundtvig M, Gullestad L, et al. Renal function in outpatients with chronic heart failure. J Card Fail. 2010;16(5):374-80.
76. Tromp J, Bamadhaj S, Cleland JGF, Angermann CE, Dahlstrom U, Ouwerkerk W, et al. Post-discharge prognosis of patients admitted to hospital for heart failure by world region, and national level of income and income disparity (REPORT-HF): a cohort study. Lancet Glob Health. 2020;8(3):e411-e22.
77. Dauriz M, Targher G, Temporelli PL, Lucci D, Gonzini L, Nicolosi GL, et al. Prognostic Impact of Diabetes and Prediabetes on Survival Outcomes in Patients With Chronic Heart Failure: A Post-Hoc Analysis of the GISSI-HF (Gruppo Italiano per lo Studio della Sopravvivenza nella Insufficienza Cardiaca-Heart Failure) Trial. J Am Heart Assoc. 2017;6(7).

78. Kristensen SL, Mogensen UM, Jhund PS, Petrie MC, Preiss D, Win S, et al. Clinical and Echocardiographic Characteristics and Cardiovascular Outcomes According to Diabetes Status in Patients With Heart Failure and Preserved Ejection Fraction: A Report From the I-Preserve Trial (Irbesartan in Heart Failure With Preserved Ejection Fraction). Circulation. 2017;135(8):724-35.
79. Lam CS, Donal E, Kraigher-Krainer E, Vasan RS. Epidemiology and clinical course of heart failure with preserved ejection fraction. Eur J Heart Fail. 2011;13(1):18-28.
80. Lawson CA, Testani JM, Mamas M, Damman K, Jones PW, Teece L, et al. Chronic kidney disease, worsening renal function and outcomes in a heart failure community setting: A UK national study. International journal of cardiology. 2018;267:120-7.
81. Patel HC, Hayward C, di Mario C, Cowie MR, Lyon AR, Rosen SD. Heart failure with preserved ejection fraction: the impact of stricter definitions. Eur J Heart Fail. 2014;16(7):767-71.
82. Bottle A, Kim D, Aylin P, Cowie MR, Majeed A, Hayhoe B. Routes to diagnosis of heart failure: observational study using linked data in England. Heart. 2018;104(7):600.
83. Murcia AM, Hennekens CH, Lamas GA, Jimenez-Navarro M, Rouleau JL, Flaker GC, et al. Impact of diabetes on mortality in patients with myocardial infarction and left ventricular dysfunction. Arch Intern Med. 2004;164(20):2273-9.
84. MacDonald MR, Petrie MC, Varyani F, Ostergren J, Michelson EL, Young JB, et al. Impact of diabetes on outcomes in patients with low and preserved ejection fraction heart failure: an analysis of the Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM) programme. Eur Heart J. 2008;29(11):1377-85.
85. Shah AM, Uno H, Kober L, Velazquez EJ, Maggioni AP, MacDonald MR, et al. The inter-relationship of diabetes and left ventricular systolic function on outcome after high-risk myocardial infarction. Eur J Heart Fail. 2010;12(11):1229-37.
86. National Institute for Health and Care Excellence. Resource impact report: Chronic heart failure in adults: diagnosis and management (NG106). 2018.
87. Du H, Farmer R, Ustyugova A. Non-elective in-patient burden attributable to Cardio-Renal-Metabolic (CRM) conditions in England. [Data on File]. 2021.
88. Olchanski N, Vest AR, Cohen JT, DeNofrio D. Comparing inpatient costs of heart failure admissions for patients with reduced and preserved ejection fraction with or without type 2 diabetes. Cardiovasc Endocrinol Metab. 2020;9(1):17-23.
89. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. Int J Cardiol. 2014;171(3):368-76.
90. OECD. Exchange rates (indicator) 2020 [Available from: https://data.oecd.org/conversion/exchange-rates.htm].
91. National institute for cardiovascular outcomes research. About Heart Failure 2021 [Available from: https://www.nicor.org.uk/national-cardiac-audit-programme/about-heart-failure/].
92. Hollingworth W, Biswas M, Maishman RL, Dayer MJ, McDonagh T, Purdy S, et al. The healthcare costs of heart failure during the last five years of life: A retrospective cohort study. International Journal of Cardiology. 2016;224:132-8.

93. Lesyuk W, Kriza C, Kolominsky-Rabas P. Cost-of-illness studies in heart failure: a systematic review 2004-2016. BMC Cardiovasc Disord. 2018;18(1):74.
94. Shafie AA, Tan YP, Ng CH. Systematic review of economic burden of heart failure. Heart Fail Rev. 2018;23(1):131-45.
95. Calado F, Viriato D, Proenca CC, Lahoz R, Agrawal R, Cotton S, et al. Burden of chronic heart failure on caregivers: a multinational cross-sectional survey in real world setting Value in Health. 2017;20(A623).
96. O'Donnell J, Smith-Byrne K, Velardo C, Conrad N, Salimi-Khorshidi G, Doherty A, et al. Self-reported and objectively measured physical activity in people with and without chronic heart failure: UK Biobank analysis. Open Heart. 2020;7(1):e001099.
97. Hossain MZ, Chew-Graham CA, Sowden E, Blakeman T, Wellwood I, Tierney S, et al. Challenges in the management of people with heart failure with preserved ejection fraction (HFpEF) in primary care: A qualitative study of general practitioner perspectives. Chronic Illness. 2021:1742395320983871.
98. Dzhioeva O, Belyavskiy E. Diagnosis and Management of Patients with Heart Failure with Preserved Ejection Fraction (HFpEF): Current Perspectives and Recommendations. Ther Clin Risk Manag. 2020;16:769-85.
99. Mullens W, Damman K, Harjola VP, Mebazaa A, Brunner-La Rocca HP, Martens P, et al. The use of diuretics in heart failure with congestion - a position statement from the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail. 2019;21(2):137-55.
100. Forsyth F, Brimicombe J, Cheriyan J, Edwards D, Hobbs R, Jalaludeen N, et al. Characteristics of patients with heart failure with preserved ejection fraction in primary care: Cross-sectional analysis. BJGP Open. 2021.
101. Maggioni AP, Anker SD, Dahlstrom U, Filippatos G, Ponikowski P, Zannad F, et al. Are hospitalized or ambulatory patients with heart failure treated in accordance with European Society of Cardiology guidelines? Evidence from 12,440 patients of the ESC Heart Failure Long-Term Registry. Eur J Heart Fail. 2013;15(10):1173-84.
102. Steinmann E, Brunner-La Rocca HP, Maeder MT, Kaufmann BA, Pfisterer M, Rickenbacher P. Is the clinical presentation of chronic heart failure different in elderly versus younger patients and those with preserved versus reduced ejection fraction? Eur J Intern Med. 2018;57:61-9.
103. Graves A, Hartshorne-Evans N. Heart failure specialist nurse care: more questions than answers. Br J Cardiol. 2019;26:86-7.
104. Zheng A, Cowan E, Mach L, Adam RD, Guha K, Cowburn PJ, et al. Characteristics and outcomes of patients with suspected heart failure referred in line with National Institute for Health and Care Excellence guidance. Heart. 2020;106(20):1579-85.
105. Fersia O, Bryant S, Nicholson R, McMeeken K, Brown C, Donaldson B, et al. The impact of the COVID-19 pandemic on cardiology services. Open Heart. 2020;7(2).
106. Heart Failure Policy Network. Heart failure and COVID-19. 2020.
107. Farmakis D, Mehra MR, Parissis J, Filippatos G. Heart failure in the course of a pandemic. Eur J Heart Fail. 2020;22(10):1755-8.
108. Sankaranarayanan R, Hartshorne-Evans N, Redmond-Lyon S, Wilson J, Essa H, Gray A, et al. The impact of COVID-19 on the management of heart failure: a United Kingdom patient questionnaire study. ESC Heart Fail. 2021;8(2):1324-32.

109. Zhang YB, Chen C, Pan XF, Guo J, Li Y, Franco OH, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. BMJ. 2021;373:n604.
110. Sannigrahi S, Pilla F, Basu B, Basu AS, Molter A. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. Sustain Cities Soc. 2020;62:102418.
111. Williamson E. Factors associated with COVID-19-related death using OpenSAFELY. Nature. 2020;584:430-6.
112. Au AG, McAlister FA, Bakal JA, Ezekowitz J, Kaul P, van Walraven C. Predicting the risk of unplanned readmission or death within 30 days of discharge after a heart failure hospitalization. American heart journal. 2012;164(3):365-72.
113. Farré N, Vela E, Clèries M, Bustins M, Cainzos-Achirica M, Enjuanes C, et al. Real world heart failure epidemiology and outcome: A population-based analysis of 88,195 patients. PloS one. 2017;12(2):e0172745.
114. Lawson CA, Zaccardi F, Squire I, Ling S, Davies MJ, Lam CSP, et al. 20-year trends in cause-specific heart failure outcomes by sex, socioeconomic status, and place of diagnosis: a population-based study. Lancet Public Health. 2019;4(8):e406-e20.
115. British Heart Foundation. Heart failure: a blueprint for change. 2021.
116. Cleland JG, Tendera M, Adamus J, Freemantle N, Polonski L, Taylor J. The perindopril in elderly people with chronic heart failure (PEP-CHF) study. Eur Heart J. 2006;27(19):2338-45.
117. Flather MD, Shibata MC, Coats AJ, Van Veldhuisen DJ, Parkhomenko A, Borbola J, et al. Randomized trial to determine the effect of nebivolol on mortality and cardiovascular hospital admission in elderly patients with heart failure (SENIORS). Eur Heart J. 2005;26(3):215-25.
118. van Veldhuisen DJ, Cohen-Solal A, Böhm M, Anker SD, Babalis D, Roughton M, et al. Beta-blockade with nebivolol in elderly heart failure patients with impaired and preserved left ventricular ejection fraction: Data From SENIORS (Study of Effects of Nebivolol Intervention on Outcomes and Rehospitalization in Seniors With Heart Failure). J Am Coll Cardiol. 2009;53(23):2150-8.
119. Bavishi C, Chatterjee S, Ather S, Patel D, Messerli FH. Beta-blockers in heart failure with preserved ejection fraction: a meta-analysis. Heart Fail Rev. 2015;20(2):193-201.
120. Butler J, Packer M, Filippatos G, Ferreira JP, Zeller C, Schnee J, et al. Effect of empagliflozin in patients with heart failure across the spectrum of left ventricular ejection fraction. European heart journal. 2022;43(5):416-26.
121. Anker SD, Butler J, Filippatos GS, Jamal W, Salsali A, Schnee J, et al. Evaluation of the effects of sodium-glucose co-transporter 2 inhibition with empagliflozin on morbidity and mortality in patients with chronic heart failure and a preserved ejection fraction: rationale for and design of the EMPEROR-Preserved Trial. Eur J Heart Fail. 2019;21(10):1279-87.
122. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD: The Task Force for diabetes, pre-diabetes, and cardiovascular diseases of the European Society of Cardiology (ESC) and the

European Association for the Study of Diabetes (EASD). European Heart Journal. 2020;41(2):255-323.
123. Cookson R, Propper C, Asaria M, Raine R. Socio-Economic Inequalities in Health Care in England. Fiscal Studies. 2016;37(3-4):371-403.
124. Packer M, Butler J, Zannad F, Filippatos G, Ferreira JP, Pocock SJ, et al. Effect of Empagliflozin on Worsening Heart Failure Events in Patients With Heart Failure and Preserved Ejection Fraction: EMPEROR-Preserved Trial. Circulation. 2021;144(16):1284-94.
125. Propper C, Stoye G, Zaranko B. The Wider Impacts of the Coronavirus Pandemic on the NHS*. Fiscal Studies. 2020;41(2):345-56.
126. ClinicalTrials.gov. EMPagliflozin outcomE tRial in Patients With chrOnic heaRt Failure With Preserved Ejection Fraction (EMPEROR-Preserved) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT03057951].
127. ClinicalTrials.gov. EMPagliflozin outcomE tRial in Patients With chrOnic heaRt Failure With Reduced Ejection Fraction (EMPEROR-Reduced) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT03057977].
128. ClinicalTrials.gov. This Study Tests Empagliflozin in Patients With Chronic Heart Failure With Preserved Ejection Fraction (HFpEF). The Study Looks at How Far Patients Can Walk in 6 Minutes and at Their Heart Failure Symptoms (EMPERIAL - Preserved) 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT03448406].
129. ClinicalTrials.gov. This Study Tests Empagliflozin in Patients With Chronic Heart Failure With Reduced Ejection Fraction (HFrEF). The Study Looks at How Far Patients Can Walk in 6 Minutes and at Their Heart Failure Symptoms (EMPERIAL-reduced) 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT03448419].
130. ClinicalTrials.gov. BI 10773 (Empagliflozin) Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME) 2016 [Available from: https://clinicaltrials.gov/ct2/show/NCT01131676].
131. ClinicalTrials.gov. A Study to Test the Effect of Empagliflozin in Patients Who Are in Hospital for Acute Heart Failure (EMPULSE) 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT04157751].
132. ClinicalTrials.gov. EMPA-KIDNEY (The Study of Heart and Kidney Protection With Empagliflozin) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT03594110].
133. ClinicalTrials.gov. A Study That Looks at the Function of the Heart in Patients With Heart Failure Who Take Empagliflozin (EMPA-VISION) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT03332212].
134. ClinicalTrials.gov. EMPACT-MI: A Study to Test Whether Empagliflozin Can Lower the Risk of Heart Failure and Death in People Who Had a Heart Attack (Myocardial Infarction) 2021 [Available from: https://clinicaltrials.gov/ct2/show/NCT04509674].
135. ClinicalTrials.gov. Comparative Effectiveness of Empagliflozin in the US 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT03363464].
136. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. EUPAS20677 2017 [Available from: http://www.encepp.eu/encepp/viewResource.htm?id=21657].

137. Oo M, Lim R, Lang C, Mordi I. Differences in clinical characteristics and outcomes in real-world HFpEF patients according to clinical definitions. European Heart Journal. 2021;42(Supplement_1):ehab724. 0745.
138. Uijl A, Savarese G, Vaartjes I, Dahlström U, Brugts JJ, Linssen GCM, et al. Identification of distinct phenotypic clusters in heart failure with preserved ejection fraction. Eur J Heart Fail. 2021;23(6):973-82.
139. Boehringer Ingelheim International GmbH. A phase III randomised, double-blind trial to evaluate efficacy and safety of once daily empagliflozin 10 mg compared to placebo, in patients with chronic Heart Failure with preserved Ejection Fraction (HFpEF). 2021.
140. CONSORT. CONSORT 2010 2010 [Available from: http://www.consort-statement.org/checklists/view/32--consort-2010/66-title].
141. Hicks Karen A, Tcheng James E, Bozkurt B, Chaitman Bernard R, Cutlip Donald E, Farb A, et al. 2014 ACC/AHA Key Data Elements and Definitions for Cardiovascular Endpoint Events in Clinical Trials. Journal of the American College of Cardiology. 2015;66(4):403-69.
142. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York1987 [updated November 2020. Available from: https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/9780470316696.fmatter].
143. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999;94(446):496-509.
144. Rogers JK, Yaroshinsky A, Pocock SJ, Stokar D, Pogoda J. Analysis of recurrent events with an associated informative dropout time: Application of the joint frailty model. Statistics in Medicine. 2016;35(13):2195-205.
145. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Annals of internal medicine. 2009;150(9):604-12.
146. Heerspink HJ, Cherney DZ. Clinical Implications of an Acute Dip in eGFR after SGLT2 Inhibitor Initiation. Clinical Journal of the American Society of Nephrology. 2021.
147. Rosano GMC, Moura B, Metra M, Bohm M, Bauersachs J, Ben Gal T, et al. Patient profiling in heart failure for tailoring medical therapy. A consensus document of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail. 2021;23(6):872-81.
148. Butler J, Filippatos G, Siddiqi TJ, Brueckmann M, Böhm M, Chopra V, et al. Empagliflozin, Health Status, and Quality of Life in Patients with Heart Failure and Preserved Ejection Fraction: The EMPEROR-Preserved Trial. Circulation. 2021.
149. Public Health England. Local action on health inequalities Understanding and reducing ethnic inequalities in health. . 2018.
150. Gov.uk. Population of England and Wales 2020 [Available from: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest].
151. Chioncel O, Lainscak M, Seferovic PM, Anker SD, Crespo-Leiro MG, Harjola V-P, et al. Epidemiology and one-year outcomes in patients with chronic heart failure

and preserved, mid-range and reduced ejection fraction: an analysis of the ESC Heart Failure Long-Term Registry. European Journal of Heart Failure. 2017;19(12):1574-85.
152. Koh AS, Tay WT, Teng THK, Vedin O, Benson L, Dahlstrom U, et al. A comprehensive population-based characterization of heart failure with mid-range ejection fraction. Eur J Heart Fail. 2017;19(12):1624-34.
153. National Institute for Health and Care Excellence. NICE impact cardiovascular disease management 2021 [Available from: https://www.nice.org.uk/about/what-we-do/into-practice/measuring-the-use-of-nice-guidance/impact-of-our-guidance/nice-impact-cardiovascular-disease-management].
154. Donnan JR, Grandy CA, Chibrikov E, Marra CA, Aubrey-Bassler K, Johnston K, et al. Comparative safety of the sodium glucose co-transporter 2 (SGLT2) inhibitors: a systematic review and meta-analysis. BMJ Open. 2019;9(1):e022577.
155. National Institute for Health and Care Excellence. Dapagliflozin for treating chronic heart failure with reduced ejection fraction [TA 679]. 2021.
156. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. J Am Coll Cardiol. 2000;35(5):1245-55.
157. U.S. Food & Frug Administration. Medical Device Development Tools (MDDT). 2021.
158. Sullivan MD, Levy WC, Russo JE, Crane B, Spertus JA. Summary health status measures in advanced heart failure: relationship to clinical variables and outcome. J Card Fail. 2007;13(7):560-8.
159. Heidenreich PA, Spertus JA, Jones PG, Weintraub WS, Rumsfeld JS, Rathore SS, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. J Am Coll Cardiol. 2006;47(4):752-6.
160. Pokharel Y, Khariton Y, Tang Y, Nassif ME, Chan PS, Arnold SV, et al. Association of Serial Kansas City Cardiomyopathy Questionnaire Assessments With Death and Hospitalization in Patients With Heart Failure With Preserved and Reduced Ejection Fraction: A Secondary Analysis of 2 Randomized Clinical Trials. JAMA Cardiol. 2017;2(12):1315-21.
161. Joseph SM, Novak E, Arnold SV, Jones PG, Khattak H, Platts AE, et al. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in patients with heart failure with preserved and reduced ejection fraction. Circ Heart Fail. 2013;6(6):1139-46.
162. Spertus JA, Jones PG, Sandhu AT, Arnold SV. Interpreting the Kansas City Cardiomyopathy Questionnaire in Clinical Trials and Clinical Care: JACC State-of-the-Art Review. Journal of the American College of Cardiology. 2020;76(20):2379-90.
163. CV Outcomes I. The Kansas City Cardiomyopathy Questionnaire (KCCQ) 2021 [Available from: https://cvoutcomes.org/pages/3214].
164. National Institute for Health and Care Excellence. Single Technology Appraisal. Ivabradine for treating chronic heart failure [TA267]. Committee Papers. 2012.
165. National Institute for Health and Care Excellence. Single Technology Appraisal. Sacubitril valsartan for treating symptomatic chronic heart failure with reduced ejection fraction [TA388]. Committee Papers. 2016.
166. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013 [PMG9] 5.3 Measuring and valuing health effects. 2013

[Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#measuring-and-valuing-health-effects].
167. Ethgen O, Standaert B. Population–versus Cohort–Based Modelling Approaches. PharmacoEconomics. 2012;30(3):171-81.
168. Zaric GS. The impact of ignoring population heterogeneity when Markov models are used in cost-effectiveness analysis. Med Decis Making. 2003;23(5):379-96.
169. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2013 [Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf].
170. Hernández Alava M., Pudney S., Wailoo A. Estimating EQ-5D by Age and Sex for the UK. NICE DSU Report. 2022.
171. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
172. Evidera. Statistical Analysis Results for Inputs of the CEM for Empagliflozin in Chronic Heart Failure (Technical Report). Data on File.; 2021.
173. National Institute for Health and Care Excellence (NICE). Position statement on use of the EQ-5D-5L value set for England (updated October 2019). 2019.
174. Sullivan PW GV. EQ-5D Scores for Diabetes-Related Comorbidities. Value Health;19(8):1002-1008. 2016.
175. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. Med Decis Making. 2006;26(4):410-20.
176. 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.
177. Peasgood T BA, Mansell P, Elliott J, Basarir H, Kruger J. The Impact of Diabetes-Related Complications on Preference-Based Measures of Health-Related Quality of Life in Adults with Type I Diabetes. Medical Decision Making. 2016;36(8):1020-1033. doi:10.1177/0272989X16658660.
178. Pollard DJ BA, Dixon S, et al. Costeffectiveness of insulin pumps compared with multiple daily injections both provided with structured education for adults with type 1 diabetes: a health economic analysis of the Relative Effectiveness of Pumps over Structured Education (REPOSE) randomised controlled trial. BMJ Open 2018;8:e016766. doi:10.1136/bmjopen-2017-016766.
179. Eurostat. Harmonised Index of Consumer Prices (HICP). European Commission. Accessed December, 2021. [Available from: https://ec.europa.eu/eurostat/web/hicp].
180. Services NP. NHS Drug Tariff Costs March 2022. 2022.
181. National Health Service (NHS). National Schedule of NHS Costs - Year 2019-20. NHS trusts and NHS foundation trusts. Accessed December, 2021. [Available from: https://www.england.nhs.uk/national-cost-collection/].
182. Alva ML, Gray A, Mihaylova B, Leal J, Holman RR. The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84). Diabet Med. 2015;32(4):459-66.

183. McMurray JJV, Trueman D, Hancock E, Cowie MR, Briggs A, Taylor M, et al. Cost-effectiveness of sacubitril/valsartan in the treatment of heart failure with reduced ejection fraction. Heart. 2018;104(12):1006.
184. Personal Social Services Research Unit. Unit Costs of Health and Social Care. 2021.
185. National Health Service. National Schedule of NHS costs: 2018/2019 2019 [Available from: https://www.england.nhs.uk/national-cost-collection/].
186. Boehringer Ingelheim International GmbH. A phase III randomised, double-blind trial to evaluate efficacy and safety of once daily empagliflozin 10 mg compared to placebo, in patients with chronic Heart Failure with reduced Ejection Fraction (HFrEF). 2020.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Clarification questions

October 2022

File name	Version	Contains confidential information	Date
ID3945 empagliflozin EAG Clarification letter 30092022 IC LW [AIC]_Redacted	1.0	Yes	21 Oct 2022

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.

Section A: Clarification on effectiveness data

A1. Priority question. In EMPEROR-Preserved, the premature discontinuation rate, not due to fatal events, was 23%. In EMPEROR-Reduced, it was 16%. Please can you provide a clinical rationale for this variation in discontinuation rates?

Although discontinuation rates were higher in EMPEROR-Preserved, they are similar across both treatment arms, suggesting that any discontinuations were a feature of the population studied rather than a feature of the medication trialled. Overall, in EMPEROR-Reduced, the premature discontinuation rate, not due to fatal events, was 16% (empagliflozin arm) and 18% (placebo arm), which numerically are very similar. In EMPEROR-Preserved, the premature discontinuation rate, not due to fatal events, was 23% (empagliflozin arm) and 23% (placebo arm), again also numerically very similar. Furthermore, in both the EMPEROR-Reduced and the EMPEROR-Preserved trials, the overall adverse event rates were lower compared to placebo arm (76.2% in empagliflozin arm vs 78.2% in the placebo arm in EMPEROR-Reduced, and 85.9% in the empagliflozin arm vs 86.5% in the placebo arm in the EMPEROR-Preserved trial). ***The company believes the differences between the trials is largely due to the differences between the target population and***

baseline characteristics as well as the longer follow-up in the EMPEROR-Preserved trial.

The EMPEROR-Preserved population was slightly older on average at baseline than the EMPEROR-Reduced population (71 versus 67), which is likely to lead to a population with more comorbidities. For example, the average atrial fibrillation rate in EMPEROR-Reduced was 35-37% vs 50-51% in EMPEROR-Preserved, with hypertension present in 72% versus 90% of the trial population, respectively, both of which are associated with a higher risk of adverse events in the EMPEROR-Preserved population.

Finally, the study follow-up period should also be noted as this was longer in EMPEROR-Preserved (26.2 months) versus EMPEROR-Reduced (16 months) trial. Having a longer follow up (due to the need to accrue the statistically sufficient number of primary endpoint events) is also likely to increase the timeframe for discontinuations to occur and hence result in an increased rate of discontinuations in the EMPEROR-Preserved trial as compared with EMPEROR-Reduced.

In conclusion, the difference in discontinuation rates can be explained by both clinical factors such as age and comorbidity differences, as well as key elements of the trials themselves such as the difference in follow up time

A2. Priority question. The evidence assessment group (EAG) notes that the population with heart failure with preserved ejection fraction (HFpEF) from PULSE has different baseline characteristics, including different treatments received, than the patients in EMPEROR-Preserved. Participants also appear to have a lower likelihood of all outcomes presented in Table 48 of the company submission (CS) in PULSE compared to EMPEROR-Preserved. Please can the company provide a clinical rationale for these differences and, in particular, why EMPEROR-Preserved is more generalisable to clinical practice in the NHS.

While there were differences in the baseline characteristics and outcomes between EMPEROR-Preserved and PULSE, these are explained by the limitations in the PULSE study design. As a consequence, there is limited evidence that PULSE offers a better representation of a typical UK patient than EMPEROR-Preserved.

Comparison of treatment mix, baseline characteristics and outcomes between EMPEROR-Preserved and PULSE

Treatment mix

The EAG expressed concerns about the generalizability of EMEPROR-Preserved given the difference in baseline characteristics and outcomes. As highlighted, one of the concerns is regarding treatment mix; the company believes this can be explained by the way data is collected in both the trial and PULSE. CPRD data captures primary care prescribing only so will underestimate prescribing if this is happening outside of primary care.

Baseline characteristics

The prevalence of comorbidities appears to be lower in PULSE compared to the trial, which may also contribute to the observed differences. There are however many characteristics common to both studies that are broadly comparable, such as age, sex, body-mass-index, heart rate, SBP and eGFR (see Table 1). These similarities should be considered in light of the observation that there was significant missing data for some baseline characteristics in PULSE. Also, the largest group in PULSE was "unknown", where no EF was recorded, suggesting that a patient's diagnosis was not accurately categorised (73% [283,672/383,896] of the total population). Therefore, some of these patients in the unknown group might be HFpEF patients..

Table 1. Comparison of baseline characteristics between EMPEROR-Preserved and PULSE

Baseline characteristic	EMPEROR-Preserved	PULSE (Cohort recorded as having HFpEF) N=31,44	PULSE (EF not known) N=283,672
Age (mean)	72 years	72 years Percentage missing: NR	76.1 years
Sex (% male)	55%	48%	51%

		Percentage missing: NR	Percentage missing: NR
BMI (mean)	29.9	29.2 Percentage missing:31.6%	28.5 Percentage missing: 44.3%
Heart rate (mean)	70.3	74.4 Percentage missing: 44.4%	77.5 Percentage missing: 46.3%
SBP (mean)	132	133 Percentage missing: 8.5%	131 Percentage missing: 10.3%
eGFR (mean)	60	65 Percentage missing: 16.7%	62.7 Percentage missing: 18.5%
EF (mean)	54.3	54.7 Percentage missing: 90%	Not reported
NT-ProBNP (mean)	946	933 Percentage missing: 90.8%	1829 Percentage missing: 93%

The difference in HHF and CV-mortality rates observed in EMPEROR-R and PULSE is likely due inaccurate recording of events in PULSE rather than differences in patient characteristics. Unlike in the real-world, HHF and CV-mortality in EMPEROR-R were adjudicated by Committee according to a strict protocol. In the real-world, an elderly patient might be admitted to wards other than cardiology, and therefore HHF and CV-mortality may not be recorded as the primary reason for hospitalisation because general physicians and other specialists may not recognise the symptoms of acute HF. The difference between AC-mortality between the trial and PULSE was

because of inaccurate classification of patients. The largest group in PULSE was "unknown" which might have included HFpEF patients (73% [283,672/383,896] of the total population). Limitations in CPRD data have recently been highlighted in a recent clinical audit of medical records and SNOMED CT coding for 78 GP practices (864,194 population) in the UK for HF. Specifically, of 19,393 patients' records that were audited, the HF case finder identified 9,725 additional patients to be audited, of whom 2,916 patients with HFrEF required further codes (47% increase) [1].

For the reasons described above, although the PULSE data serve as a useful indicator that the populations are not completely dissimilar, the EMPEROR-Preserved data were considered the most suitable to inform the economic model. Overall, the company believes that the discussed differences between PULSE and EMPEROR-Preserved populations will not have an impact on the consideration of empagliflozin's cost-effectiveness.

A3. Priority question. Please present outcome data for the Kansas City Cardiomyopathy Questionnaire - clinical summary score (KCCQ-CSS), cardiovascular (CV) mortality, overall mortality, and occurrence of adjudicated HHF (first and recurrent) in the following subgroups:

- **People with heart failure ejection fraction (HF EF) 40% to <50%;**
- **People with HF EF 50% to <60%;**
- **People with HF EF ≥60%.**

Empagliflozin has demonstrated benefit across the EF spectrum and this is consistent with the MA. We recognise that the EAG would like to explore what outcomes could be expected by EF subgroups, and therefore these data are provided below [2], which is consistent with the marketing authorisation (MA) extension for empagliflozin for symptomatic chronic HF regardless of LVEF [3]. Please also note that the requested EF categories do not align with ESC categories and are, therefore, not necessarily of clinical relevance. There is ambiguity across the clinical community regarding the range of EF that should be used when categorising HF (see section B.1.3.1.1. Disease Overview – Classification). For simplicity, the company submission has defined the target population based on those

patients with symptomatic HF with EF >40% to cover the remaining population not assessed in the previous appraisal (TA773).

Nevertheless, the results below show that there was always a numerical improvement for patients in empagliflozin arm, for all subgroups (Table 2, Table 3). On some occasions, the differences were not statistically significant, but this can be expected since the different subgroups were not powered to show statistical significance, and they had small sample sizes.

Table 2. Outcome data for KCCQ-CSS scores by ejection fraction (Randomised set)

Endpoint	Empagliflozin 10 mg N / Change in clinical summary score	Placebo N / Change in clinical summary score	Mean change from baseline (95% CI)
KCCQ-CSS at 0-52 weeks			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
KCCQ-CSS at 0-148 weeks No week 148 data available, presenting from baseline to end of treatment visit			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; LVEF, left-ventricular ejection fraction.

Table 3. Outcome data for adjudicated HHF, CV mortality and all-cause mortality by ejection fraction (Randomised set)

Endpoint	Empagliflozin 10 mg N with event/N analysed	Placebo N with event/N analysed	Hazard ratio (95%CI)
Occurrence of adjudicated HHF (first and recurrent)			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]

Endpoint	Empagliflozin 10 mg N with event/N analysed	Placebo N with event/N analysed	Hazard ratio (95%CI)
CV mortality			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
Overall mortality			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; CV, cardiovascular; HHF, hospitalization for heart failure; LVEF, left-ventricular ejection fraction.

A4. Priority question. Please present the number of people for whom trial medication stopped for reasons other than death in the following subgroups:

- **People with HF EF 40% to <50%;**
- **People with HF EF 50% to <60%;**
- **People with HF EF ≥60%.**

The requested information is provided in Table 4 below. The company would like to emphasise that clinicians are not rigid in the use of EF subgroups and, further, that the subgroups requested above are not used in clinical practice. The company acknowledges that ejection fraction is often visually estimated from an echocardiogram and have a margin of error associated with them [4]. Thus, clinicians are in agreement that diagnosis and treatment based on EF is not well-defined or understood in the clinical practice.

Overall, approximately a quarter of patients across HF EF subgroups discontinued treatment for reasons other than death, which is comparable to the overall discontinuation rate of 23.3% (1,395 out of 5,988). These findings demonstrate that there is no evidence that EF affects the likelihood of discontinuing treatment. In addition, the benefit of empagliflozin on CV-death or hospitalisation for heart failure

has been demonstrated across the EF spectrum in the EMPEROR-Preserved and Reduced trials (Anker et al. 2021) [5].

Table 4. Medication discontinuations for reasons other than death in HF EF subgroups (Randomised set)

HF EF subgroup	Discontinuations / Sample size (%)
40% to < 50%	[redacted]
50% to < 60%	[redacted]
≥ 60%	[redacted]

Abbreviations: EF, ejection fraction; HF, heart failure.

A5. Priority question. A section of Table 16 in the CS is copied below.

1. Are the labels of the rows incorrect? Should:

- **"Value of >50%" be "Value of <50%";**
- **"Value of 50% to >60" be "Value of 50 to <60%".**

2. The numbers of people in the trial were very similar between these 3 groups defined by their baseline ejection fraction. Was this planned (e.g., did the trial have criteria for this at patient selection) and, if so, what was the reasoning for this?

Regarding A5.1, apologies for the oversight, indeed the labels contain errors. The labels should be: >40% to <50%; ≥50% to <60% and ≥60%.

Regarding A5.2, no, this was not planned. In fact, it was initially planned to recruit 35% to 50% of patients with LVEF >=50. In the clinical trial protocol it is stated:

"IRT will be used to aim for a trial population consisting of approximately 35% to 50% with an LVEF >50%. To ensure adequate enrolment of patients, the final decision on capping will be based on the recommendation from the executive steering committee during the recruitment period."

However, when recruiting all HFpEF patients as per the inclusion criteria, about two-thirds of the population had LVEF >=50% Table 5. The Executive Committee advised to keep recruiting without capping patients with higher LVEF. After this

decision and before any unblinding, the Trial Statistical Analysis Plan was updated to include subgroup analysis in groups of <50, >=50-<60 and >=60.

Table 5. Overview of patients' distribution per ejection fraction % (Randomisation set)

Baseline characteristic	Empagliflozin 10 mg	Placebo
Number of subjects	2,997	2,991
Left ventricular ejection fraction (EF)		
Mean	54.3+/-8.8	54.3+/-8.8
Value of >50%, N (%)	995 (33.2)	988 (33.0)
Value of 50% to >60, N (%)	1,028 (34.3)	1,030 (34.4)
Value of ≥60%, N (%)	974 (32.5)	973 (32.5)

Section B: Clarification on cost-effectiveness data

For any scenarios requested in Section B, please ensure these are implemented as user selectable options in the economic model so that these can be combined. Furthermore, if the company chooses to update its base-case results, please ensure that cost-effectiveness results, sensitivity and scenario analyses incorporating the revised base-case assumptions are provided with the response along with a log of changes made to the company base-case.

Model structure

B1. Priority question. Please explain why renal events were removed from the model (in comparison to the empagliflozin model used in TA773). Clinical expert opinion provided to the EAG noted that clinical events are equally relevant for the preserved ejection fraction population.

The EAG requested a clarification regarding the removal of renal outcome from the model. In EMPEROR-Preserved, renal outcome was a secondary, exploratory outcome, defined as time to the first event in the composite renal endpoint: chronic dialysis (with a frequency of twice per week or more for at least 90 days), renal

transplant, sustained reduction in eGFR from baseline of ≥40%, sustained eGFR <15 mL/min/1.73 m2 for patients with baseline eGFR ≥30 mL/min/1.73 m2, or sustained eGFR <10 mL/min/1.73 m2 for patients with baseline eGFR <30 mL/min/1.73 m2. A sustained reduction was determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (the first to last of the consecutive eGFR values).

Renal events were removed from the model because the difference between the two groups was not statistically significant, as shown in the trial's results (HR= 0.95 95%CI: 0.73–1.24). As correctly noted, renal events were included in the model used in TA773, but in the EMPEROR-Reduced trial this outcome was statistically significant (HR: 0.50; 95% CI, 0.32 - 0.77). Therefore, as this outcome was numerically favourable but statistically not significant for empagliflozin, it was decided to take a conservative approach and remove it from the model. Including the renal events will not affect the cost effectiveness results as the impact of safety on cost effectiveness is already largely reflected by inclusion of adverse events, and for which the impact is very minimal.

Clinical data used in the model

B2. Priority question. Please explain why only data on KCCQ-CSS collected in EMPEROR-Preserved at baseline; weeks 12; 32; and 52 were used in the model, when end of treatment (EOT past 148 weeks) data on KCCQ-CSS were collected.

A. **Please consider including the EOT data on KCCQ-CSS in the estimation of transition probabilities between KCCQ-CSS quartiles in the model.**

It is correctly stated that data from baseline to up to 52 weeks were collected and included in the model determining the transition probabilities. No data on KCCQ-CSS at 148 weeks were collected. EOT data were available but since EOT varies between different patients, the company believes that it should not be utilised in the model as this would lead to biased results. By adding EOT to the model, the company expects an increase in the uncertainty in the model outcomes and hence biased cost-effectiveness estimates.

A scenario was built in the CE the model ('Context' sheet, row 70) to test the impact of transition probabilities for KCCQ-CSS. In this scenario, at a defined point in time the proportion of patients in the KCCQ-CSS quartiles under the treatment arm was set equal to those proportions in the placebo arm at 5, 3, 2, and 1 years. For the trial population, the ICER increased from £14,428.65 to £16,985, £17,457, £17,187, and £16,139, when 5-year, 3-year, 2-year, and 1-year, time points were tested, respectively (Table 6).

Overall, it appears that the results supporting the cost-effectiveness of empagliflozin remain consistent even when more conservative scenarios for a waning treatment effect, in terms of KCCQ transition probabilities, are applied.

Table 6. Scenario analysis testing the impact of KCCQ transition probabilities on model's ICER

Description		ICER (£/QALY)	% change relative to base case ICER
Set the proportion of patients in the KCCQ-CSS quartiles under the empagliflozin arm equal to those proportions in the SoC arm at 5, 3, 2 and 1 year(s)	5 years	£16,139	11.85%
	3 years	£17,187	19.11%
	2 years	£17,457	20.99%
	1 year	£16,985	17.71%

Abbreviations: ICER, incremental cost-effectiveness ratio; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; QALY, quality-adjusted life-year; SoC, standard of care.

B3. Priority question. Please clarify which KCCQ-CSS data set was used to estimate the transition probabilities between KCCQ-CSS quartiles in the model (for example, was the mixed model for repeated measures [MMRM] analysis used, was the observed case on-treatment [OC-OT] population used, etc.).

The data used to produce the transition probability matrices were the observed patient KCCQ-CSS scores including after treatment discontinuation (OC-AD) from baseline to week 52. Missing values from visits within that interval were imputed using the last-observation-carried-forward (LOCF) method. This was a pre-defined approach, consistent to what has been applied in in TA773. The company believes that the OC-AD population was appropriately used for the analysis.

B4. Priority question. Please justify why the last observation carried forward (LOCF) method was deemed appropriate to handle missing data for KCCQ-

CSS scores in the estimation of transition probabilities across quartiles in the model.

The company believes that the assumptions around missingness do not impact the estimation of cost effectiveness. A key assumption of using the LOCF approach is that data are missing completely at random, or that there is no underlying systemic reason for why there are missing data at any timepoint. To estimate the KCCQ-CSS transitions in the base case, LOCF imputation was used, meaning that at weeks 12, 32, and 52, for patients still alive and followed up, if a value was missing the last observation was carried forward. The mean scores at weeks 12, 32, and 52 (Table 7) were very close between the imputed and non-imputed datasets, indicating lack of underlying bias.

In addition, as described below for question B5, the distribution of KCCQ-CSS score changes from baseline were very similar between the imputed and non-imputed datasets. Given the similarity between imputed and non-imputed KCCQ-CSS scores, the missing-at-random assumption of the LOCF approach was considered valid, and therefore this imputation method was used in the base case. The company would like to indicate that the same imputation method was used in the submission for HFrEF (TA773) and was accepted by the committee.

Table 7. Comparison of KCCQ-CSS score statistics for imputed and non-imputed data (Randomised set)

Without imputation	Empagliflozin 10mg		Placebo	
Visit	**N**	**Mean (SD)**	**N**	**Mean (SD)**
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]
With imputation	**Empagliflozin 10mg**		**Placebo**	
Visit	**N**	**Mean (SD)**	**N**	**Mean (SD)**
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]

*The higher number of observations at week 12 is due to records from patients with missing scores at baseline. These patients contributed data from week 12 onwards and were kept in the analysis.
Abbreviations: KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; SD, standard deviation.

B5. Priority question. Similarly to the data provided by the company after technical engagement in TA773, please fill the table below (twice) with the following KCCQ-CSS data from EMPEROR-Preserved:

A. The distribution of individual changes in mean KCCQ-CSS over time without imputed values.

B. The distribution of individual changes in mean KCCQ-CSS over time with imputed values (i.e. using the LOCF method described in the CS and used in the model).

The EAG would like more information on the KCCQ-CSS scores for imputed and non-imputed values. The company has provided the requested information below and would like to emphasise that the differences between the two arms in the imputed and without imputation KCCQ scores were very small (Table 88, Table 99). This increased the confidence to the conclusion that empagliflozin was associated with better KCCQ-CSS scores, and that the imputation method did not affect the estimation of transition probabilities and subsequently the cost-effectiveness results.

Table 8. Changes in KCCQ-CSS scores without imputation (Randomised set)

A: Without imputation		N*	Mean	SD	p-value	Min	P10	P25	Median	P75	P90	Max
Empagliflozin 10mg	Baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to week 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to week 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

A: Without imputation		N*	Mean	SD	p-value	Min	P10	P25	Median	P75	P90	Max
Empagliflozin 10mg	Change from baseline to EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*For N please provide the total number of patients with available observations on KCCQ-CSS data for each time point (i.e., baseline, week 12,32,52 and EOT). No data beyond week 52 were available. EOT change from baseline is presented, which is the last visit (up to week 52) where the patient had an observation and is the same in the imputed and non-imputed datasets.

Abbreviations: EOT, end of treatment; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; SD, standard deviation.

Table 9. Changes in KCCQ-CSS scores without imputation (Randomised set)

B: With imputation		N*	Mean	SD	p-value	Min	P10	P25	Median	P75	P90	Max
Empagliflozin 10mg	Baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to week 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to week 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Empagliflozin 10mg	Change from baseline to EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*For N please provide the total number of patients with available observations on KCCQ-CSS data for each time point (i.e., baseline, week 12,32,52 and EOT). No data beyond week 52 were available. EOT change from baseline is presented, which is the last visit (up to week 52) where the patient had an observation and is the same in the imputed and non-imputed datasets.

Abbreviations: EOT, end of treatment; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; SD, standard deviation.

B6. Priority question. Please provide the transition probabilities (equivalent to those reported in Table 28 of the CS) without imputed values (i.e., as observed in EMPEROR-Preserved).

A. Please provide a scenario where the observed transition probabilities without imputed values are used in the economic model.

The EAG would like to understand the impact that imputation had on the cost effectiveness results. Table 10 below reports the transition probabilities with and without imputed data. The transition probabilities from both approaches are similar, demonstrating that empagliflozin was associated with better KCC-CSS scores. The imputed probabilities should be considered more robust compared to the non-imputed, naïve approach, since the imputation method used was considered appropriate as stated above in response to question B4. Similarly, the company would like to highlight that the same imputation method was used in the submission for HFrEF (TA773) and was accepted by the committee. Finally, in response to the EAG's comment, the company programmed a scenario where transition probabilities without imputed values were used to populate the model. The results of this scenario analysis demonstrate that the deterministic ICER increases to £20,198.43/QALY compared to the base case deterministic ICER (£14,428.65/QALY) Table 11Table 11. The company believes that this approach, though, is less robust than the base case, in which imputed transition probabilities are used.

Table 10. Transition probabilities with and without imputed data (Randomised set)

			With imputation			Without imputation		
	From:	To:	Months 1 - 3	Months 4 - 8	Months 9+	Months 1 - 3	Months 4 - 8	Months 9+
Empagliflozin + SoC	1	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

			With imputation			Without imputation		
	From:	To:	Months 1 - 3	Months 4 - 8	Months 9+	Months 1 - 3	Months 4 - 8	Months 9+
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	4	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SoC	1	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	4	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SoC, standard of care.

Table 11. Summary results of QALYs per patient when using imputed data (base case) and non-imputed data (EAG scenario)

	With imputation			Without imputation		
	Empagliflozin + SoC arm	**SoC arm**	**Incremental**	**Empagliflozin + SoC arm**	**SoC arm**	**Incremental**
KCCQ-CSS 1st Quartile	0.68	0.72	-0.039	0.68	0.69	-0.014
KCCQ-CSS 2nd Quartile	0.96	1.00	-0.040	0.94	0.98	-0.036
KCCQ-CSS 3rd Quartile	1.12	1.09	0.036	1.12	1.13	-0.008
KCCQ-CSS 4th Quartile	1.69	1.58	0.116	1.73	1.63	0.104
Total QALYs	**4.27**	**4.17**	**0.098**	**4.29**	**4.22**	**0.069**

Abbreviations: KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score

Modelling of hospitalisations

B7. Priority question. The CS states that, "*the HHF* [hospitalisation for heart failure] *rates* [in EMPEROR-Preserved] *appeared to be relatively constant over time*". Please provide the data underpinning this conclusion.

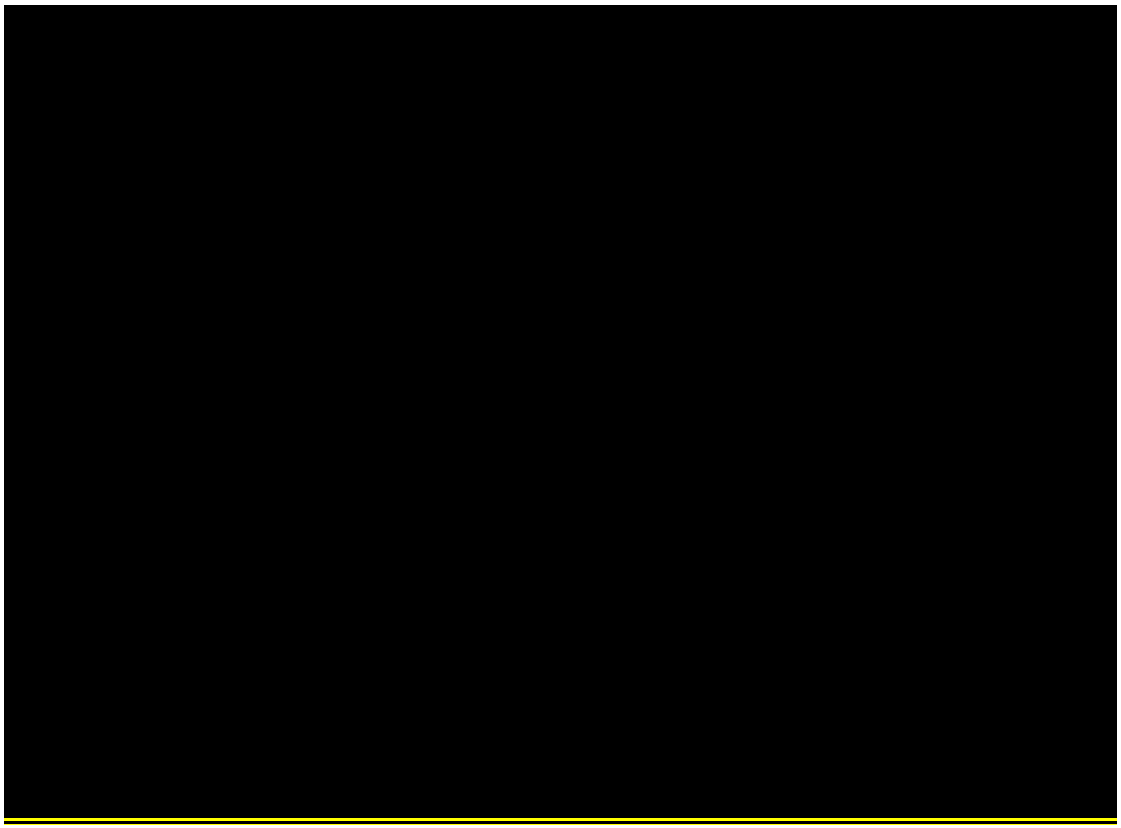
The EAG would like to understand better why the rates for hospitalisation were assumed to be constant. The assumption of a constant rate for heart-failure hospitalisations was based on visual assessment of the monthly HHF rate across follow-up, which appears constant over time as seen in Figure 1 further below. The HHF rate was calculated using patient level data from EMPEROR-Preserved trial (randomised set) and dividing the counts of HHF events each month (Figure 2) by the corresponding observation time contributed by patients (Figure 3). The observed spike in rates at the later stages of follow-up is due to fewer patients remaining in the trial at that point in time, causing few HHF events to have greater impact on the rate compared to earlier periods.

The constant HHF rate assumption effectively meant that time would be omitted as a predictor in the HHF equations. This approach is consistent with prior appraisals (TA267 (ivabradine) [6], TA388 (sacubitril valsartan) [7], TA679 (dapagliflozin) [8]), where it was assumed that the overall rate of HHF remained constant over the lifetime of the CE model. However, a separate equation was fitted that included a time predictor, which was estimated with a negative coefficient (Estimate = -0.012, *p*-value = 0.008), implying that HHF risk decreases over time. As this result was not clinically plausible, it was only retained as a scenario in the CE model, and the constant rate equation was used in the base case instead.

In anticipation of a request to explore the possibility that the assumption of constant rate for HHF has a significant impact on results as this was also requested for TA773, the company implemented a scenario in the submitted version of the model where the rate of hospitalisation changes every month. This scenario can be found in the "Context" sheet. Implementing this monthly increase in the rates of hospitalisation (KCCQ-quartile 1= 0.4%, KCCQ-quartile 2= 0.3%, KCCQ-quartile 3= 0.2%, KCCQ-quartile 4= 0.1%), resulted in an ICER of £13,860.49, which is lower compared to the base case ICER of £14,428.65.

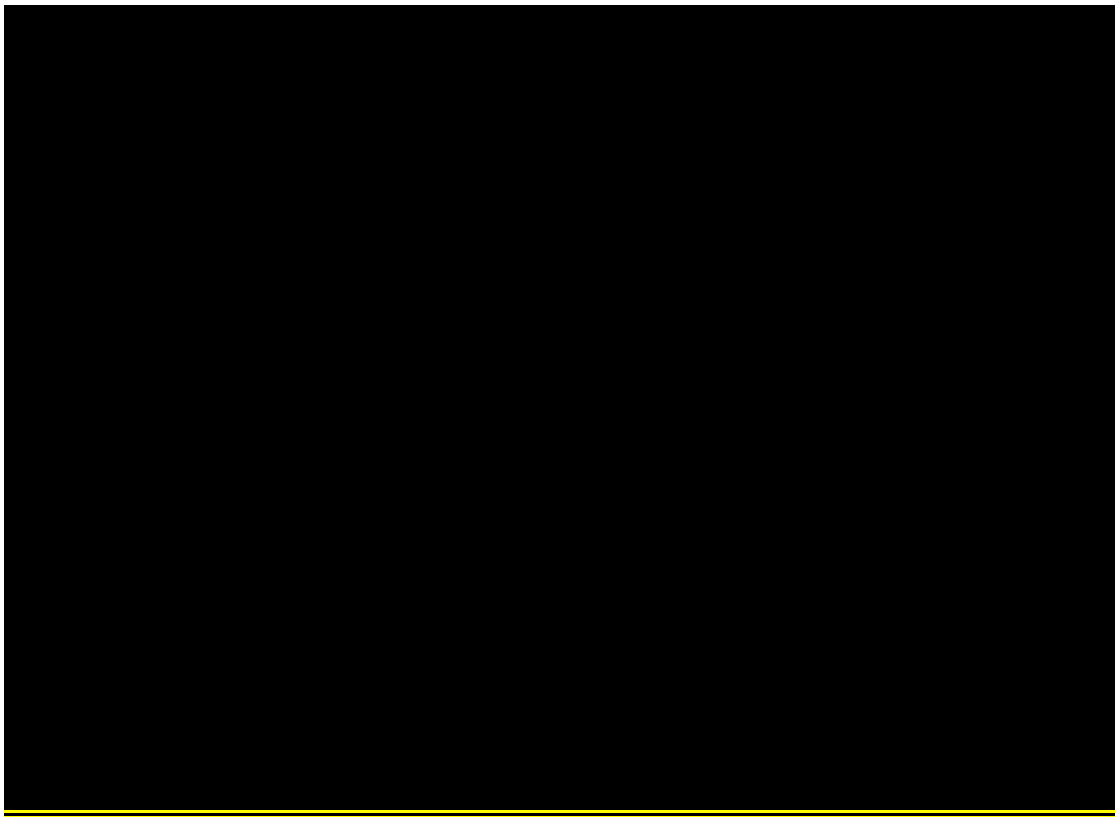
Given the above, the company believes that the modelling of a constant rate of HHF is appropriate, and it is unlikely that this choice has biased the cost-effectiveness results.

Figure 1: Monthly rate of HHF events over time for EMPEROR-Preserved patients in empagliflozin and placebo arms (Randomised set).



Abbreviation: HHF, hospitalization for heart failure.

Figure 2: Counts of monthly HHF events.



Abbreviation: HHF, hospitalization for heart failure.

Figure 3: Person-years of observation contributed per month.

B8. Priority question. Given how HHF events are very likely to be impacted by patients' age, please explain why age was not considered in the regression analysis used to estimate HHFs in the model (and only treatment and KCCQ-CSS states were used as covariates).

The model already submitted to the EAG included the functionality to select between reduced or extended risk equations to model the rate of HHF. This can be found in cell H67 of the "Clinical Inputs" sheet. When the "Reduced list" option is chosen, only treatment and KCCQ-CSS states are used as covariates. Other predictors (including age), that underwent a pre-specified variable selection process, were included in the extended set of equations to model HHF events only if found to be statistically significant. The variable selection process to arrive at the extended set of equations consisted of the following steps:

1. Univariate selection: Every potential predictor was added separately to the core model. Those found to be statistically significant at the 10% level were kept.

2. Multivariate backward selection: All predictors kept from step 1 were fitted to the model simultaneously, and a backward elimination process was used to remove non-influential predictors at the 10% significance level, until only significant predictors remained.

3. Multivariate forward selection: Predictors excluded in step 1 were introduced into the resulting model from step 2, one at a time. The predictor with the lowest p-value was kept. This process was repeated until no re-introduced predictor appeared significant at the 10% level.

4. Final selection: The resulting model from step 3 was finalised by removing predictors that had lost statistical significance at the 10% level after the re-inclusion of predictors.

The risk equations based on the reduced list using only treatment and KCCQ-CSS states as predictors were fitted to model outcomes for the ITT population and other subgroups, such as patients above or below 65 years of age. Both the extended and reduced risk equations consider age in the cost-effectiveness calculations (albeit in an indirect manner in the latter case). However, to test its impact directly, the age

predictor was added to the ITT reduced list risk equation and was found to be statistically non-significant (Coefficient estimate = [redacted], *p*-value = [redacted]). Therefore, the company believed that it is not meaningful to add this equation into the cost-effectiveness model to examine the impact on the ICER. Therefore, not having age as a predictor in the risk equations is justified and it is not expected to influence the model's results.

The company wants to highlight that the choice of using the reduced list of equations as base case in the model was based on the validation step that was conducted and described in section 'B.3.13.1.1 Internal validity' in the CS. Specifically, the rates of HHF, CV death and non-CV death observed during the average EMPEROR-Preserved trial follow-up of 26 months were compared with the economic model predictions over the same time horizon (

Figure 4). The reduced set of equations resulted in estimates that were closer to the estimates from the observed data, meaning that the model's predictions were more validated when using the reduced list of predictors compared to using the extended list of predictors. Thus, the reduced list of predictors was used as the base case.

Figure 4. Comparison of observed and model-predicted event rates in the ITT population: CV death, non-CV death, and HHF rates per 100 patient-years

Abbreviations: CV, cardiovascular; PY, person years; SoC, standard of care

B9. Priority question. Please fill the table below for the equivalent number of patients in the trial and in the model, the equivalent time period (for all arms in the trial and in the model), and for total HHF events (i.e., first and subsequent HHF together). Please include a corresponding table in the Excel model, linked to the model engine results.

The EAG requested for a table reporting the total number of HHF over 26 months and 3 years to be added in the model Table 12. Accordingly, a table was programmed into the model to obtain this information directly from the model's engine. This can be found in the "Context" sheet (cells K104:M105).

The results show that there are some differences in the number of events observed in EMPEROR-Preserved compared to those predicted in the model. However, the observed and predicted differences in events between empagliflozin and placebo are similar, indicating that the incremental results of the model are valid. Therefore, the company believes that this should be of minor consideration for the committee.

Table 12. Comparison of total number of HHF observed in the trial and those predicted in the cost-effectiveness model

	EMPEROR-preserved			CE model		
	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference
Total number of HHF (first and subsequent) over 26 months	■	■	■	■	■	■
Total number of HHF (first and subsequent) over 3 years	■	■	■	■	■	■

Abbreviations: CE, cost-effectiveness; HHF, hospitalization for heart failure; SoC, standard of care.

B10. Priority question. As discussed in the EAG review of the company's response to technical engagement (TE) of TA773, a more robust method for estimating HHF in the model would have been to use Kaplan-Meier (KM) data,

independent of KCCQ-CSS states. This would have allowed the rate of HHF to directly vary in the model (as opposed to the assumption of a constant rate) and would also have allowed first and subsequent hospitalisations to be modelled separately. In this appraisal, the same issues remain relevant as the company used a Poisson regression which assumed a constant risk of hospitalisation in the entire model (regardless of the time-varying element of KCCQ-CSS-linked HHF), and did not differentiate initial and subsequent hospitalisations in the model. Given that KM data on HHF in EMPEROR-Preserved shows a considerable difference in empagliflozin's effect on first and subsequent hospitalisations, can the company please:

A. Use KM data on first HHF to fit parametric models and extrapolate the rate of HHF events over time in the model according to technical support document (TSD) 14. The EAG's preference is for the KM data from EMPEROR-Preserved (independent on KCCQ-CSS) scores to be used, however, if the company's preference is to make HHF dependent on KCCQ-CSS, the EAG suggests that the company follows the same approach used by the company in TA773 which used, "*a counting process setup with start and stop times to create periods defined by the occurrence of each hospitalisation and/or changes in KCCQ-CSS quartiles. That is, a patient will have one record per change in KCCQ-CSS and per hospitalisation with start and stop times of the period defined by the time when these changes occur*."

B. Use the same approach as described in point A, however, to model subsequent HHF separately from first HHF.

The EAG has suggested the use of Kaplan-Meier data and parametric survival models on time to HHF in order to derive extrapolations and better assess the constant HHF rate assumption underlying the Poisson modelling approach in the base case. The company notes that parametric survival models are not typically applied for recurrent events; variations in the rate of event over time can be handled in the context of event count models (like Poisson) by allowing time as a predictor. A Poisson equation with time as a predictor (question B7) was fitted as a scenario analysis but demonstrated a clinically implausible negative coefficient (suggesting that HHF risk decreases over time), which was not appropriate as the base case.

The company notes that this also occurred in TA773 and was similarly considered implausible, as clinical experts confirmed that they would not expect the rate of HHF to decline over time. The observed patterns are likely attributable to the declining numbers of patients at risk and the disproportionate influence of events near the end of follow-up in the trial. To address the EAG's concern, the company implemented the proposed parametric approach on time to first HHF event, and considered a total of six distributions: exponential, Weibull, Log-Normal, Log-logistic, Gompertz, and Generalised Gamma. The fitted models had no other predictors apart from treatment arm. Their hazard estimates were extrapolated across the trial follow-up, as seen in Figure 5 below. With the exception of the exponential distribution, which assumes a constant hazard, all other distributions yielded extrapolations with decreasing or plateauing hazards, which are clinically implausible, and in agreement with the Poisson model that included time as a predictor.

These findings support the idea that the fitted Poisson model assuming a constant HHF risk represented the most clinically plausible choice. The assumption of a constant hazard in the long-term may not hold, and it is possible that rates may increase over time as patients progress and age, but as this increasing pattern could not be observed in the data, it is more appropriately examined through scenario analyses in the model.

As mentioned above in response to question B7, to explore the possibility that the assumption of constant rate for HHF may not be appropriate, the company already implemented a scenario where the rate of hospitalisation changes every month in the model submitted to NICE. This scenario can be found in the "Context" sheet. Implementing this monthly increase in the rates of hospitalisation (KCCQ-quartile 1= 0.4%, KCCQ-quartile 2= 0.3%, KCCQ-quartile 3= 0.2%, KCCQ-quartile 4= 0.1%), resulted in an ICER of £13,860.49, which is lower compared to the base case ICER of £14,428.65.

Given the results described above, the company believes that additional analysis on time from first to subsequent hospitalisation does not offer additional information and would not improve uncertainty of cost effectiveness. As time to subsequent hospitalisation analysis breaks randomisation as baseline is redefined as having

experienced an initial HHF, the insights from this analysis are potentially not valid or useful.

Overall, the company believes that there is substantial evidence to support the assumption of a constant HHF hazard over time. Nonetheless, even in the scenario that the hazard increases, empagliflozin remains a cost-effective treatment option.

Figure 5. Hazard extrapolations from parametric models fitted on time to first HHF events

Abbreviations: HHF, hospitalization for heart failure.

Mortality

B11. Priority question. Regarding the scenario included in the model and described in the CS, Table 50, Issue 7 - please disaggregate this analysis into two separate options (i.e. two independent drop-down menus in the model):

1. **A fitted curve to non-CV mortality for the placebo arm of the EMPEROR-Preserved trial with no direct treatment effect and with no indirect treatment effect through KCCQ-CSS predictors for the intention-to-treat (ITT) population.**
2. **A fitted curve to CV-mortality for the placebo arm of the EMPEROR-Preserved trial with no indirect treatment effect through KCCQ-CSS predictors for the ITT population.**

As per EAG's request, the company have programmed the aforementioned scenarios into the model in the "Clinical Inputs" worksheet. The model now has an option to choose to include or exclude KCCQ-CSS health state effect on cardiovascular (CV) and all-cause (AC) mortality, separately. The deterministic ICERs for each scenario are reported in Table 13 below. Similarly, there is the option to remove the treatment effect from both CV and AC mortality. As expected, the different scenarios led to some changes in the ICER compared to base case ICER; however, these differences are not substantial and therefore the conclusions regarding the cost-effectiveness of empagliflozin remain robust.

Table 13. deterministic ICER obtained in the base case and suggested scenarios

Assumption	Deterministic ICER (£/QALY)
Base case setting for all inputs	£14,428.65
Scenario 1: fitted curve to AC mortality with no direct treatment effect and with no indirect treatment effect through KCCQ-CSS predictors for the ITT population	£21,104.31
Scenario 2: fitted curve to CV mortality with no direct treatment effect and with no indirect treatment effect through KCCQ-CSS predictors for the ITT population	£16,112.98

Abbreviations: AC, all-cause; CV, cardiovascular; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; QALY, quality-adjusted life-years.

B12. Priority question. The company's base case predicts that at 10 years in the analysis, [redacted] of standard of care (SoC) patients are alive, while at 20 years in the model there are still [redacted] of SoC patients alive. Considering that:

A. Patients in EMPEROR-Preserved were 72 years at baseline, it seems clinically implausible that [redacted] of patients would still be alive at 92 years with preserved ejection fraction (EF), therefore, please run a scenario in the model where either these long-term extrapolations are validated by external data; or an adjustment to the tail of the survival curve is undertaken to reflect a more realistic survival prediction.

B. Clinical expert opinion provided to the EAG was that the preserved EF population in the UK is on average 80 years at baseline, and presents with considerable co-morbidities, please run a scenario in the model

where the baseline age for the UK population is reflected in terms of life expectancy in the long-term model.

The EAG questioned the model's predictions in term of patients' long-term survival based on clinical expert feedback. The company appreciates the clinicians' input regarding patients' long-term survival. As the long-term CV and non-CV mortality risk is based on parametric distributions, the model already includes the option to choose different distributions. In particular, generalised gamma and Gompertz lead to lower overall survival, which at 10 years in analysis is [redacted] for the former and [redacted] for the latter, and at 20 years is [redacted] for both. When selecting the generalised gamma and Gompertz models, the deterministic ICER is £14,472.56 for generalised gamma, and £17,552.60 for Gompertz, which is consistent with the conclusion for cost effectiveness similar to the base case deterministic ICER of £14,428.65 based on Weibull. Given that altering the distribution for mortality does not impact the conclusions for cost effectiveness, the company concludes that modelling of long-term survival is not a key contributor to model's results. The company maintains that Weibull is the most appropriate parametric distribution to model the long-term survival of patients, as this choice was made on pre-specified statistical criteria and visual inspection of all parametric models. The AIC and BIC statistics for the different parametric distributions can be found in Table 29 (page 112) of the CS, and the fit of the distributions can be seen in Figure 21 and Figure 22 (page 117).

As requested for part B of this question, a scenario was programmed into the model changing the baseline age of the ITT population from 71.89 to 80 years. This results in a decrease of the deterministic ICER from £14,428.65 (base case) to £14,021.13. Therefore, the company believes that baseline age of patients in the model does not alter the conclusion that empagliflozin is a cost-effective treatment.

B13. Priority question. Please fill the table below for the equivalent number of patients in the trial and in the model, the equivalent time period (for all arms in both the trial and the model) for CV deaths. Please include a corresponding table in the Excel model, linked to the model engine results.

The EAG requested a table reporting the total number of CV deaths over 26 weeks, and 3 years to be added in the model (see Table 14). The company was not certain if 26 months was required as well (as was the case in B9), so this time point has also

been added to the table in case it is of interest for the EAG. A table was programmed into the model to obtain this information directly from the model's engine. This can be found in the "Context" sheet (cells K110:M112).

As noted in the response to question B9, the results show that there are some differences in the number of events observed in EMPEROR-Preserved and those predicted in the model. Nevertheless, the observed and predicted differences in CV deaths between empagliflozin and placebo were similar, indicating that the incremental results of the model remain valid. Therefore, the company believes that this should be of minor consideration for the committee as it does not alter the conclusion of cost effectiveness for empagliflozin.

Table 14. Comparison of total number of CV deaths observed in the trial and those predicted in the cost-effectiveness model

	EMPEROR-preserved			CE model		
	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference
Total number of CV deaths at 26 weeks	■	■	■	■	■	■
Total number of CV deaths at 26 months	■	■	■	■	■	■
Total number of CV deaths at 3 years	■	■	■	■	■	■

Abbreviations: CE, cost-effectiveness; CV, cardiovascular; SoC, standard of care.

B14. Priority question. The EAG is unclear why the company decided to remove statistically significant predictors from the CV mortality equations (such as age, prior atrial fibrillation [AF], etc.). Therefore, please include a scenario in the model where the CV mortality equation includes all the statistically significant predictors.

The EAG questioned the predictor selection process that was used in mortality equations. Missing predictors were removed because they tested as statistically non-

significant in the pre-specified variable selection process, as described in Section 4.1 of the Statistical Analysis Report and in response to question B8 above.

Statistically significant predictors were included in the extended list risk equations (Table 15).

The model shared with the EAG includes the option to choose the extended list of predictors for the risk equations. When the extended predictor list risk equation is used to model mortality risk, the deterministic ICER is £16,955.14 as compared to the base case ICER of £14,428.65. Therefore, the company believes that despite using the alternative approach proposed from the EAG for the CV mortality equation, empagliflozin remains a cost-effective treatment option.

The company wants to emphasize that the choice of using the reduced list of equations as base case in the model was based on the validation step that was conducted and described in section 'B.3.13.1.1 Internal validity' in the CS, which is also mentioned above in response to question B8. Specifically, the rates of HHF, CV death and non-CV death observed during the average trial follow-up of 26 months were compared with the economic model predictions over the same time horizon. The reduced set of equations resulted in estimates that were closer to the estimates from the observed data, meaning that the model's predictions were more validated when using the reduced list of predictors compared to using the extended list of predictors (

Figure 4, question B8). Thus, the reduced list of predictors was used as the base case.

Table 15. CV mortality risk equations using extended predictor list and time-varyign KCCQ-CSS and NYHA health states, based on best-fitting (Weibull) distribution (Randomised set)

Formulation	KCCQ-CSS time-varying health states	
Predictors	Estimate	*p*-value
Distribution parameters (Weibull)		
Shape	[redacted]	[redacted]
Scale	[redacted]	[redacted]
Empagliflozin 10mg (ref: placebo)	[redacted]	[redacted]
Prior atrial fibrillation or flutter (ref: no, unknown)		
Yes	[redacted]	[redacted]
Age ≥ 65 (ref: age < 65)	[redacted]	[redacted]
Male (ref: female)	[redacted]	[redacted]
Race (ref: white, asian, multiple, pacific, unknown)		
Black	[redacted]	[redacted]
Native	[redacted]	[redacted]
Region (ref: Non Latin America)		
Latin America	[redacted]	[redacted]
Baseline KCCQ-CSS quartile (ref: Non-Q2 [55.7, 74])		
Q2 [55.7, 74)	-	-
NT-proBNP (log-scale)	[redacted]	[redacted]
Prior HF (ref: no)	[redacted]	[redacted]
Ischemic HF (ref: no)	[redacted]	[redacted]
Time since HF diagnosis (ref: 0-1 years)		
1-5 years	[redacted]	[redacted]
5+ years	[redacted]	[redacted]
EQ-5D-3L (standardized)	[redacted]	[redacted]
Updated KCCQ-CSS quartile (ref: Q1 [0, 55.7))		
Q2 [55.7, 74)	[redacted]	[redacted]
Q3 [74, 88)	[redacted]	[redacted]
Q4 [88, 100]	[redacted]	[redacted]
Updated NYHA class (ref: class III-IV)		
II	-	-
I	-	-

Abbreviations: CV, cardiovascular; EQ-5D-3L, EuroQoL - Five Dimensions - Three Levels; HF, heart failure; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; NT-proBNP, N-terminal pro B-type natriuretic peptide; NYHA, New York Heart Association.

Treatment discontinuation

B15. The model predicts that patients will stay on treatment for a mean duration of [redacted] years. Given that clinical expert opinion is that patients (who do not discontinue treatment) stay on empagliflozin for the rest of their lives; the baseline age of the model population (72 years); and the predicted survival for modelled empagliflozin patients of [redacted] years, it would appear that time on treatment is underestimated in the long-term model. Therefore, can the company please discuss the clinical plausibility of the estimated time on treatment in the model.

The EAG queried the plausibility of the model's predicted time to treatment discontinuation for empagliflozin. Time to treatment discontinuation was modelled using parametric distributions. The model includes the function to choose between different parametric models for time to treatment discontinuation. The base case applying the generalised gamma distribution resulted in a deterministic ICER of £14,428.65/QALY. If log-normal distribution is chosen, which is the parametric model leading to the longest time on treatment estimates, the time to treatment discontinuation is 5 years, which aligns with what would be expected based on the clinical expert's opinion. The deterministic ICER for this scenario reduces to £14,807.12/QALY, and thus not substantially different compared to base case, leading the company to conclude that time to treatment discontinuation is not an important factor affecting the cost-effectiveness results.

In addition, the model includes a functionality which allows the user to choose between including or excluding treatment discontinuation. It can be found in the "Clinical Inputs" tab at cell F149. In a scenario where treatment discontinuation is not included and patients stay on treatment throughout their lifetime, the deterministic ICER increases slightly from £14,428.65 (base case) to £15,125.77.

Based on the two scenarios above, the company believes that empagliflozin remains cost-effective independently of the approach used to model time to treatment discontinuation.

Quality of life

B16. Priority question. Clinical expert opinion provided to the EAG indicated that the assumption of a 12-month duration for the impact of HHF on patients' QoL is overestimated. The experts indicated that the average length of stay in the hospital for HHF for patients with preserved EF is 11 days (which isxx mean stay for HHF in EMPEROR-Preserved). Subsequently, one expert indicated that a reasonable assumption is that 1 day in hospital impacts patients' quality of life (QoL) for 1 week after discharge. The other clinical expert indicated that 6 months of impact (as a maximum) could also be plausible after discharge. Therefore, please conduct two alternative scenario analyses where:

A. **It is assumed that HHF events impact patients' QoL for [redacted] months after discharge.**
B. **It is assumed that HHF events impact patients' QoL for 6 months after discharge.**

The EAG noted correctly that the impact of HHF on HRQL lasts up to 12 months after the event in the model. The disutility for HHF is calculated from the utility equations, which use patient level data from EMPEROR-Preserved, similar to the other risk equations (e.g., for mortality) in the model. The company clarifies that the utility equation includes indicators for time since the hospitalisation rather than duration. More specifically, the equation includes terms for 0-1, 1-2, 2-4, 4-12 months from hospitalisation vs. not hospitalised ever or in past 12 months. The coefficient for 0-1 represents the change in utility in the first month after hospitalisation, 1-2 represents the change in the second months after hospitalisation, etc (shown in the CE model in the "Risk Equations – Lookup!" sheet). These describe the course of change in utilities over the year following the hospitalisation with patients returning to their pre-hospitalisation utility after one year. Therefore, length of hospitalisations is not taken into account in the calculations in any way. This is consistent with 12-month impact of HHF on patients' quality of life which was implemented in both TA773 and TA679, and for both appraisals this assumption was considered acceptable by the appraisal committee.

Published literature suggests that a 12-month duration for the disutility associated with an HHF is a reasonable assumption. A systematic literature review of ten studies as reported by Di Tanna et al 2021 [9] supports the assumption that a hospitalisation event impacts a patient's utility in the longer term. For example, Gorostiza et al 2015 reported that over a 12-month period post hospitalisation, patients demonstrated a lowered HRQoL in terms of mobility (81.7%) and usual activities (82.1%). It is further clinically plausible that the disutility due to HFF last 12 months, as demonstrated by the study by Vaduganathan et al 2022, which showed that for HFrEF and HFpEF patients experiencing a HHF event, KCCQ scores remain below pre-hospitalisation estimates within 12 months following the event (Figure 6).

Although the company believes that the approach taken to model HHF disutility is appropriate, the two scenarios requested from the EAG (i.e., impact of HHF lasts 2.75 and 6 months, respectively) were added in the model and can be found in the "Context" sheet. When the impact of hospitalisation is assumed to last 2.75 months, the HHF disutility is -0.0827, and when it lasts for 6 months, it is -0.1805. The ICER for the first scenario increases to £17,911.77/QALY and for the second £16,511.32/QALY, indicating that empagliflozin remains cost-effective even when the HRQoL impact of HHF is assumed to be shorter.

Given that an assumption of 12 months duration for HHF disutility has been accepted previously by NICE committee meetings and supported by the literature, the company is confident that the approach is appropriate. Even when more conservative approaches assuming shorter impact for HHF were implemented in the model, the conclusions for empagliflozin's cost-effectiveness did not change.

Figure 6. Mean Kansas City Cardiomyopathy Questionnaire clinical summary scores before and after a hospitalization for heart failure (Vaduganathan et al, 2022)

B17. Priority question. Please include age-related utility decrements throughout the model time horizon using the algorithm published by Ara and Brazier 2010.

The EAG requested the age-related utility decrements to be applied in the model. Accordingly, a scenario was programmed into the model as per the EAG's request. For this scenario, a multiplier was calculated based on cohort age and sex using the formula for general population EQ-5D reported in Ara and Brazier (2010) [10]. The multiplier was incorporated into the utility calculations in the model engine sheets for empagliflozin + SoC and SoC. In this scenario, the deterministic ICER slightly increased to £14,988.54/QALY (Table 16) compared to the base case deterministic ICER (£14,428.65/QALY). This very slight increase is driven by the slightly longer survival of patients in the empagliflozin + SoC arm compared with those on SoC. However, this scenario analysis has limited impact on the ICER and therefore should

be of minor consideration for the committee as the conclusions for cost effectiveness do not change.

Table 16. Deterministic cost-effectiveness results including age-related utility decrements using the algorithm published by Ara and Brazier (2010)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
SoC	£10,158.70	6.79	4.01	-	-	-	-
Empagliflozin + SoC	£11,566.10	6.87	4.11	£1,407.40	0.07	0.09	£14,988.54

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; SoC, standard of care; QALYs, quality-adjusted life year.

B18. Priority question. The CS states that, "*There were no statistically significant differences in EQ-5D* [Euroqol 5 Dimensions] *scores between the two treatment groups in the EMPEROR-Preserved trial, hence the treatment was not a predictor in the utility equation*." Furthermore, Table 32 in statistical appendix N shows that change in utility from baseline to week 100 in the trial was [redacted] for empagliflozin and [redacted] for placebo (mean difference of [redacted]) over 3 years. Therefore, can the company please:

A. **Compare the [redacted] utility gain seen in the trial over 3 years with the equivalent utility gain seen in the first 3 years of the economic model, and explain any discrepancies.**
B. **Reconcile the statement in the CS with the utility gain estimated in the model of [redacted]. The EAG is aware that the latter incorporates utility gains beyond transitions between KCCQ-CSS states (related to HHF), however, it notes that the impact of reducing HHF in the empagliflozin arm of the trial (and of changes in KCCQ-CSS scores) did not seem to be sufficient to show a statistical significant difference in patients' health-related quality of life (HRQoL) between treatment arms.**

The EAG queried the validity of the differences in utility gains as shown in the model. As can be seen in Table 17, the same difference between the two arms in utility gain of 0.02 over 3 years in the trial can also be seen in the results of the economic model. No discrepancies found between the models' prediction and the observed results in the trial in terms of utility gain differences between the two arms. This

increases the confidence of the company on the validity of the model's results, showing that empagliflozin can be considered cost-effective compared to SoC.

Another concern of the EAG is in regards to the utility gain differences at the end of the model's lifetime horizon. Utility gains in the model are influenced by multiple inputs, including: the time patients spent in each KCCQ state, with healthier states being related to higher utilities; adverse events, each of which have assigned a disutility which is triggered when a patient experiences this event; rate of hospitalisation due to heart failure, which functions similar to adverse event disutility. Based on the assessment of uncertainty and scenarios performed in response to the clarifications from the EAG, the company considers the lower risk of HFF as the main driver of utility gains for empagliflozin compared to SoC. The difference in the risk for hospitalisation due to heart failure between the two arms in the model (event rate per 100 person years is [redacted] for empagliflozin and SoC, respectively) is a reflection of the clinical data reported in the clinical trial (hazard ratio, 0.73; 95% CI, 0.61 to 0.88; P<0.001). The company provided a further validation check as detailed in Table 17, which shows that the predicted difference in utility gains found in the model is reasonable and reflects the increased clinical efficacy of empagliflozin compared to SoC.

Table 17. Utility results seen in year 3 of the economic model, disaggregated by health state

	Empagliflozin + SoC arm	SoC arm	Incremental
KCCQ-CSS 1st Quartile	[redacted]	[redacted]	[redacted]
KCCQ-CSS 2nd Quartile	[redacted]	[redacted]	[redacted]
KCCQ-CSS 3rd Quartile	[redacted]	[redacted]	[redacted]
KCCQ-CSS 4th Quartile	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]

Abbreviations: KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; QALY, quality-adjusted life-year; SoC, standard of care.

B19. Priority question. Please provide the following mapped EQ-5D-3L data (from the EQ-5D-5L data from EMPEROR-Preserved):

A. Average baseline EQ-5D-3L for both arms in the trial, together with respective number of observations, and statistical significance for the difference in utility at baseline across arms (as per table below).

B. Change from baseline in mean EQ-5D-3L, for both arms in the trial, together with respective number of observations, and statistical significance of changes, for all available time points (see table below).

The company has provided the requested information below and would like to emphasise that as expected (due to the randomisation performed) there are no statistically significant differences between the two arms at EQ-5D-3L at baseline Table 18Table 18. The results indicate the small but consistent differences between the utility gains for patients in empagliflozin compared to SoC arms Table 19. Overall, these results demonstrate the clinical efficacy of empagliflozin compared to SoC.

Table 18. Baseline EQ-5D-3L for empagliflozin and placebo (Randomised set)

	Empagliflozin 10 mg		Placebo		p-value
	N	Mean (SD)	N	Mean (SD)	
Baseline EQ-5D-3L	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: EQ-5D-3L, EuroQoL - Five Dimensions - Three Levels; SD, standard deviation.

Table 19. Change from baseline EQ-5D-3L for empagliflozin and placebo (Randomised set)

Change from Baseline	Empagliflozin 10 mg		Placebo		Difference (SE; p-value)
	N	Mean (SD)	N	Mean (SD)	
Week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 100	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 148	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SE, standard error; SD, standard deviation.

B20. Priority question. Please explain why non-statistically significant variables (such as KCCQ-CSS scores at baseline) were included as covariates in the HRQoL regression model (Table 36 in the CS). Please consider removing

non-statistically significant variables and rerunning the equation to estimate utility values in the model.

The EAG questions the predictors included in the risk equation for utilities. In the utility equations, all types of treatment-emergent adverse events were exempt from removal regardless of their statistical significance because their impact on utilities was of primary interest, as they were used as inputs for the model. Thus, the company cannot run a scenario with those predictors removed, as their coefficients are used to populate the model.

All levels of multi-level categorical predictors were retained if at least one level was statistically significant. This is true in the case of baseline KCCQ-CSS quartile in Table 20 below (Table 36 in the CS), where quartile 1 (KCCQ-CSS score of 0 to 55) is statistically significant at the 10% level, with a *p*-value of 0.042. It is also true in the case of the Region predictor, where levels "Latin America" and "Other" are statistically significant, and hence the non-significant levels "North America" and "Asia" are also retained. Therefore, the company maintains that a pre-specified and transparent approach was applied to populate the model with utility estimates.

Table 20. Health-related quality of life equation derived from EMPEROR-Preserved trial (Randomised set)

Covariate	Coefficient	SE	p-value
Distribution/Type	Linear Mixed Model		
Intercept	[redacted]	[redacted]	[redacted]
Demographics			
Age ≥65	[redacted]	[redacted]	[redacted]
Male (ref: Female)	[redacted]	[redacted]	[redacted]
Region			
Latin America	[redacted]	[redacted]	[redacted]
North America	[redacted]	[redacted]	[redacted]
Asia	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Baseline KCCQ quartile (ref: [88.02, 100])			
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]	[redacted]

Covariate	Coefficient	SE	p-value
KCCQ-CSS: 0 to 55 (Quartile 1)	[redacted]	[redacted]	[redacted]
Medical History			
Time Since HHF			
HHF: <1 month	[redacted]	[redacted]	[redacted]
HHF: 1 to <2 months	[redacted]	[redacted]	[redacted]
HHF: 2 to <4 months	[redacted]	[redacted]	[redacted]
HHF: 4 to <12 months	[redacted]	[redacted]	[redacted]
AE			
Urinary tract infection	[redacted]	[redacted]	[redacted]
Genital Mycotic Infection	[redacted]	[redacted]	[redacted]
Acute renal failure	[redacted]	[redacted]	[redacted]
Hepatic injury	[redacted]	[redacted]	[redacted]
Volume depletion	[redacted]	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]	[redacted]
Hypoglycaemic event	[redacted]	[redacted]	[redacted]
Bone fracture	[redacted]	[redacted]	[redacted]
Ketoacidosis*	[redacted]	[redacted]	[redacted]

*Not included in the base-case.

Abbreviations: AE, adverse event; CSS, clinical summary score; EQ-5D, EuroQol five dimensions; HF, heart failure; HHF, hospitalisation for heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; SE, standard error.

B21. Priority question. Please discuss the clinical plausibility of the considerably higher HHF-related disutility value estimated from the EMPEROR-Preserved population compared with the EMPEROR-Reduced population.

As the EAG noted, there is a difference in the 12-month disutility as estimated in EMPEROR-Preserved and EMPEROR-Reduced. The company believes that this difference is in line with other differences noted in terms of the target population and baseline characteristics of the two populations in the trials. In particular, the mean age in EMPEROR-Preserved is 71.8 and 71.9 for empagliflozin and placebo, and for EMPEROR-Reduced is 67.2 and 66.5 respectively. In addition, as noted also in question A1 above, EMPEROR-Preserved population was slightly older on average at baseline (71 versus 67), which is associated with higher comorbidities as would be expected in an older population. Further, the average atrial fibrillation in EMPEROR-

Reduced was 35-37% vs 50-51% in EMPEROR-Preserved, with hypertension in 72% versus 90% of the trial population, respectively, both of which are associated with a higher risk of adverse events in the EMPEROR-Preserved population. Furthermore, in EMPEROR-Reduced the overall adverse events were 76.2% in empagliflozin arm vs 78.2% in the placebo arm, and 85.9% in the empagliflozin arm vs 86.5% in the placebo arm in the EMPEROR-Preserved trial. Therefore, it appears that the observation that the population of EMPEROR-Preserved is older and has more adverse events is also reflected in the disutility for HHF.

It is also possible that there is a difference in the type of treatment that HFrEF and HF>40%EF patients are receiving. On some occasions patients with HF>40%EF could be more likely to be treated less aggressively by clinicians. This can be due to unconscious bias in the management from clinicians, since for patients who are older and with more comorbidities (as patients in EMPEROR-Preserved are), clinicians are more likely to discuss the End-of-Life criteria. This could also impact how patients perceive the HHF and how it affects their quality of life.

Nevertheless, the company believes that the HHF disutility used is not likely to have a substantial impact in the results showing the cost-effectiveness of empagliflozin, so it is potentially of less interest for the committee. This is reflected in the scenario described in question B16, where the impact of hospitalisation was assumed to last 2.75 months and 6 months, in which the ICER was £17,911.77/QALY and £16,511.32/QALY, respectively, indicating that empagliflozin remained cost-effective even when the HRQoL impact of HHF is assumed to be shorter.

B22. Priority question. Please discuss the clinical plausibility of the slightly higher KCCQ-CSS utility values (unadjusted for age) estimated from the EMPEROR-Reduced population compared with the EMPEROR-Preserved population.

The company believes that the difference in utilities between patients in EMPEROR-Reduced and EMPEROR-Preserved reflect the other differences noted in terms of the target population and baseline characteristics between the trials. As detailed in questions A1 and B20 above, the main differences between the two populations are baseline age and incidence of adverse events. In particular, the mean age in EMPEROR-Preserved is 71.8 and 71.9 for empagliflozin and placebo, and for

EMPEROR-Reduced is 67.2 and 66.5 respectively. Adverse event incidence rates were lower compared to placebo arm (76.2% in empagliflozin arm vs 78.2% in the placebo arm in EMPEROR-Reduced, and 85.9% in the empagliflozin arm vs 86.5% in the placebo arm in the EMPEROR-Preserved trial). It seems that, although in terms of heart failure symptoms patients in EMPEROR-Preserved would be expected to have higher utilities, the impact of age in conjunction with adverse events is the main driver of the mean utilities.

Costs and resource use

B23. Priority question. Please include a scenario analysis in the model reflecting the distribution of SoC treatments in the table below (advised by the EAG's clinical experts).

As requested, a scenario was programmed into the model using the suggested distribution of SoC treatments Table 21. The deterministic ICER decreased from £14,428.65 to £14,409.71. Given that this scenario analysis has limited impact on the ICER, it should be of minor consideration for the committee.

Table 21. Distribution of SoC treatments suggested by the EAG

Drug class	ITT Proportion
ACEi	40%
ARB	39%
ARNi	0%
MRA	37%
Beta blocker	86%
Loop or high ceiling diuretics	80%

Abbreviations: ACEi. angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor-blocker; ITT, intention-to-treat; MRA, mineralocorticoid receptor antagonists.

B24. Priority question. Please include a scenario analysis in the model reflecting the distribution of SoC treatments received by patients in the EMPEROR-Preserved trial (Table 10.4.4.1:1, page 98) in the clinical study

report (CSR), in order to match SoC costs with treatment effectiveness in the model.

The company believes that there was a misunderstanding on the distribution and source of SoC used in the model. The inputs used in the model are the same as those proposed by the EAG. The confusion potentially stems from the fact that Table 10.4.4.1: 1 of the CSR, which is the one referred in the EAG's question, sources the data from Table 15.1.4: 13 of the CSR, which is the table that the company used to extract the data. The only difference is that the percentages reported for ACE inhibitors (ACEi) and ARB were aggregated in Table 10.4.4.1: 1, but disaggregated in Table 15.1.4: 13. The company used the disaggregated percentages to allow for more appropriate costing.

B25. Priority question. Similarly to TA773, please include the following scenario analyses in the model:

A. **Use the following costs (and update to the current cost year where needed) from Table 3 in the Alva paper: fatal myocardial infarction £1,521, fatal ischaemic heart disease £3,766, and fatal stroke £3,954.**

B. **Estimate the weighted costs of CV-death by the proportion of events leading to CV deaths observed in EMPEROR-Preserved (Table 11.1.2.4.2:1, page 120 of the CSR).**

C. **Assume the cost of sudden cardiac death to be zero and alternatively;**

D. **Use the unit cost of £1,632 for all sudden cardiac deaths in the model corresponding to the total HRG costs for cardiac arrest (NHS Costs 2019-20) and update the cost as necessary.**

The EAG requested a number of scenarios to be included in the model, in which different costs, and their combinations, are applied. A scenario was programmed into the model to use the suggested cost estimates for fatal myocardial infarction, fatal ischaemic heart disease, and fatal stroke as requested in question B25 A. Overall, the deterministic ICER slightly increased from £14,428.65 to £14,501.34 without any impact on cost-effectiveness conclusions.

Regarding point B, C and D above, two additional scenarios were programmed into the model (separate from the above-mentioned point A). The results are reported below Table 22able 22. Overall, the deterministic ICER marginally increased and does not impact the conclusion regarding cost-effectiveness.

Table 22. Overview of ICERs from base case and scenario analyses

Assumption	Deterministic ICER (£/QALY)
Base case setting for all inputs	£14,428.65
Scenario 1: Use fatal myocardial infarction, fatal ischaemic heart disease, and fatal stroke costs as reported in the Alva paper.	£14,501.34
Scenario 2 (conservative approach): Estimate the weighted costs of CV-death by the proportion of events leading to CV deaths observed in EMPEROR-Preserved, assuming the cost of sudden cardiac death to be zero.	£14,837.39
Scenario 3 (non-conservative approach): Estimate the weighted costs of CV-death by the proportion of events leading to CV deaths observed in EMPEROR-Preserved, assuming the cost of sudden cardiac death to be £1,632 (total HRG costs for cardiac arrest [NHS Costs 2019-20]).	£14,703.84

Abbreviations: CV, cardiovascular; HRG, healthcare resource group; ICER, incremental cost-effectiveness ratio; NHS, National Health Service; QALY, quality-adjusted life-years.

B26. Priority question. Given the duration of HHF in EMPEROR-Preserved is: mean [redacted]; median [redacted]; and Q3 [redacted] and that the more severe cost code used for HHF (EB03A) is associated with a 53-day long hospitalisation, whereas the less severe cost code (EB03E) is associated with 13 days in hospital, please conduct a scenario analysis in the model where the cost of £2,062 (after appropriately inflated to the cost year) is used to calculate the cost of all HHFs events in the model.

The EAG requested a scenario where the cost code for less severe HHF (EB03E: £2,062.20) only is to be used as the cost of all HHFs events in the model. A scenario was programmed into the model to use the suggested cost estimate. Overall, the deterministic ICER slightly increased from £14,428.65 (base case) to £15,214.44 without any impact on cost-effectiveness conclusions.

Section C: Textual clarification and additional points

C1. Please confirm that the incremental cost-effectiveness ratio (ICER) reported in Table 46 of the CS for the time horizon of 20 years is correct. The EAG would expect the resulting ICER to be higher than the company's base case ICER.

The EAG is correct that there was an error with the ICER for 20 years in CS. The company apologises for this mistake. The correct ICER for 20 years follow-up is £14,583, so slightly increased from £14,428.65 (base case), as the EAG expected. The corrected ICER does not have an impact on the conclusions for the cost-effectiveness of empagliflozin.

References

[1] M. Kahn, A. Grayson, P. Chaggar, M. Ng Kam Chuen, A. Scott, C. Hughes and e. al, "Primary care heart failure service identifies a missed cohort of heart failure patients with reduced ejection fraction.," *Eur Heart J.,* 2021.

[2] J. Butler, M. Packer, G. Filippatos, J. P. Ferreira, C. Zeller, J. Schnee, M. Brueckmann, S. J. Pocock, F. Zannad and S. D. Anker, "Effect of empagliflozin in patients with heart failure across the spectrum of left ventricular ejection fraction," *European heart journal,* vol. 43, no. 5, pp. 416-426, 2022.

[3] T. A. McDonagh, M. Metra and M. Adamo, "ESC Scientific Document Group, Authors/Task Force Members:, ESC Scientific Document Group, ESC Scientific Document Group. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure," *Eur Heart J,* vol. 42, no. 36, pp. 3599-3726, 2021.

[4] W. O'Dell, "Accuracy of Left Ventricular Cavity Volume and Ejection Fraction for Conventional Estimation Methods and 3D Surface Fitting," *J Am Heart Assoc,* vol. 8(6), March 2019.

[5] S. D. Anker, J. Butler, G. Filippatos, J. P. Ferreira, E. Bocchi, M. Böhm, H.-P. Brunner–La Rocca, D.-J. Choi, V. Chopra and E. Chuquiure-Valenzuela, "Empagliflozin in heart failure with a preserved ejection fraction," *New England Journal of Medicine,* vol. 385, no. 16, pp. 1451-1461, 2021.

[6] National Institute for Health and Care Excellence (NICE), "Ivabradine for treating chronic heart failure [TA267]," 2012.

[7] National Institute for Health and Care Excellence (NICE), "Sacubitril valsartan for treating symptomatic chronic heart failure with reduced ejection fraction [TA388]," 2016. [Online]. Available: https://www.nice.org.uk/guidance/ta388.

[8] National Institute for Health and Care Excellence (NICE), "Dapagliflozin for treating chronic heart failure with reduced ejection fraction [TA679]," 2020. [Online]. Available: https://www.nice.org.uk/guidance/ta679.

[9] G. L. Di Tanna, M. Urbich, H. S. Wirtz, B. Potrata, M. Heisen, C. Bennison, J. Brazier and G. Globe, "Health state utilities of patients with heart failure: a systematic literature review," *Pharmacoeconomics,* vol. 39, no. 2, pp. 211-229, 2021.

[10] R. Ara and J. E. Brazier, "Populating an economic model with health state utility values: moving toward better practice," *Value in Health,* vol. 13, no. 5, pp. 509-518, 2010.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Professional 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 the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

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 13 pages.

NICE National Institute for Health and Care Excellence

About you

1. Your name	[redacted]
2. Name of organisation	UK Clinical Pharmacy Association – Heart Failure Committee
3. Job title or position	[redacted]
4. Are you (please select Yes or No):	An employee or representative of a healthcare professional organisation that represents clinicians? Yes A specialist in the treatment of people with this condition? Yes A specialist in the clinical evidence base for this condition or technology? Yes Other (please specify):
5a. Brief description of the organisation (including who funds it).	A membership organisation for pharmacy professionals, funded by membership fees
5b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] **If so, please state the name of manufacturer, amount, and purpose of funding.**	[Could not find appraisal matrix.]
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No

The aim of treatment for this condition

6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main goals of heart failure treatment are to : Relieve signs and symptoms - improve quality of life for patients Prevent hospital admission Prevent disease progression Reduce mortality Of particular emphasis for patients with heart failure with a mildly reduced (HFmrEF) or preserved ejection fraction (HFpEF), is to delay progression to heart failure with a reduced ejection fraction (HFrEF)
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	A reduction in hospital admission 20% (RRR) A reduction in CV death 20% (RRR) An improvement in scores related to quality of life using a validated patient tool – Kansas City Cardiomyopathy questionnaire (KCCQ) or Minnesota Living with Heart Failure Slower decline of renal function
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes HFpEF and HFmrEF accounts for half of all patients diagnosed with heart failure and is a growing concern. Management of these patients is a significant challenge, as there are no specific treatments to improve prognosis or clear way to diagnose. The evidence base and treatment options are predominantly for patients with HFrEF (ejection fraction <40%).

What is the expected place of the technology in current practice?

9. How is the condition currently treated in the NHS?	This appraisal considers two different phenotypes:

	HFmrEF (EF 41-49%) – No substantial RCT has been performed exclusively in HFmrEF. Some of the pharmacological treatment options for patients with HFrEF *may* be considered for this cohort of patients (European Society of Cardiology Guidelines, 2021). This includes ACE inhibitors/Angiotensin II receptor blockers/neprilysin inhibitors, beta-blockers and mineralocorticoid receptor antagonists. HFpEF (EF > 50%) – Treatment is focussed on managing patient comorbidities such as atrial fibrillation, diabetes, hypertension, kidney disease. Weight loss in obese patients and increasing exercise may improve symptoms and exercise capacity. Diuretics are provided to patients with all types of heart failure to reduce congestion. There is no evidence to advise non-pharmacological treatment (CRT or ICD therapy) in patients with HFmrEF or HFpEF.
9a. Are any clinical guidelines used in the treatment of the condition, and if so, which?	2018 NICE Chronic heart failure in adults: diagnosis and management NICE Guideline 106 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure European Heart Journal, Volume 42, Issue 36, 21 September 2021, Pages 3599-3726 https://doi.org/10.1093/eurheartj/ehab368
9b. Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The pathway of care varies. Diagnosis is unclear and the terminology of HFmrEF and HFpEF is not widely understood by other professionals. Not all specialist in HF agree on how to diagnose HFmrEF and HFpEF. Thereafter the treatment pathway may also differ. Some heart failure specialist services only see patients with EF<40%, therefore the increasing numbers of patients with HFmrEF and HFpEF poses a large burden to the NHS and particularly primary care, who may be managing these patients independently. These patients often have multiple presentations in A+E and are often admitted with fluid overload.
9c. What impact would the technology have on the current pathway of care?	The technology will provide a treatment option for patients where there is little or no evidence for any pharmacological treatment other than symptomatic relief. It may increase awareness of HFmrEF and HFpEF as more patients will be eligible for treatment.

	The technology might encourage commissioners to extend the scope of current heart failure services and provide more standardised pathways of care. This is essential to be able to deliver this urgently needed treatment.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	It is already used for HFrEF, type 2 diabetes mellitus so clinicians are relatively experienced prescribing it already. The dose is the same as other indications.
10a. How does healthcare resource use differ between the technology and current care?	The addition of an SGLT-2 inhibitor would require the patient to take one extra tablet a day (there is only one dose). Baseline bloods and a repeat check 2-4 weeks post initiation will be required similar to many other heart failure medications. No extra healthcare resource use is expected.
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	All clinical settings can use this technology – primary care, secondary care, specialist HF clinics – but it will most likely be initiated in primary care or a specialist HF clinic. It would also be used for patients admitted to hospital with HFmrEF and HFpEF – to optimise care prior to discharge.
10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Empagliflozin is already used in the management of type 2 diabetes mellitus, HFrEF and CKD. Many teams are already becoming upskilled to prescribe and monitor patients on empagliflozin. Further training and education may be required to improve the management of these patients.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes – there is very little existing medication or technology that treats patients with an EF >40% other than diuretics and the management of co-morbidities. The SGLT2i's are the first medication to offer a reduction in HF hospitalisation which has a huge impact on patient care and the NHS.
11a. Do you expect the technology to increase length of life more than current care?	Yes

11b. Do you expect the technology to increase health-related quality of life more than current care?	Yes
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Patients with Heart Failure (irrespective of EF), type 2 diabetes mellitus and CKD all have licenses to use empagliflozin. Improvements seen were in a broad range of patients with heart failure.

The use of the technology

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	The treatment is relatively easy to implement as it has been used for a number of years already in other indications. Baseline bloods will be required including U+E's, LFTs, SGLT2i and FBCs. The patient would then need their volume status reviewed prior to initiation. Patients should be reviewed between 2-4 weeks post initiation to ensure renal function stable and the patient has tolerated. There is no further dose titration required. Thereafter, monitoring should be done in line with normal management of the HF patient (unless the diabetic management needs further review). Prescribing could be in Primary care or Secondary care, although the diagnosis would likely need to be done in Secondary care and the first prescription may be better initiated in clinic.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these	No, the treatment will be ongoing indefinitely once initiated. The treatment would only be stopped if the patient developed significant side-effects.

include any additional testing?	
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Yes – it is expected that empagliflozin will reduce hospital admission for HF and improve the prognosis of patients with HFpEF and HFmrEF.
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes – there is currently no alternative evidence-based treatment which shows reduction in hospital admission and cardiovascular death. These patients are often poorly managed as many HF services exclude them from their services due to lack of capacity and underfunding. Patients often have multiple hospital admissions at present.
16a. Is the technology a 'step-change' in the management of the condition?	Yes – little alternative to treat this patient population as above
16b. Does the use of the technology address any particular unmet need of the patient population?	Yes – little alternative to treat this patient population as above
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The most common side effects of SGLT-2 inhibitors are: urinary tract infections, genital infections, normoglycaemic diabetic ketoacidosis, dizziness due to hypotension. We would recommend the SGLT-2 inhibitor is stopped immediately in any patient who develops normoglycaemic diabetic ketoacidosis (DKA) - and medical teams need to be aware of the rare possibility of this if the patient becomes unwell. The other side effects are usually less serious and settle when the medication is withheld.

	Patients should be advised of possible side effects when the medication is started so they know to seek medical attention should they develop any including providing a patient information booklet. They should also be counselled on “sick-day rules” and to withhold the medication if acutely unwell and at risk of dehydration e.g vomiting, diarrhoea, to reduce the risk of DKA. This is routine practice when SGLT-2 inhibitors are used for other licensed indications.

Sources of evidence

18. Do the clinical trials on the technology reflect current UK clinical practice?	EMPEROR-preserved (2021) showed empagliflozin reduced the combined risk of cardiovascular death or hospitalization for heart failure in patients with heart failure and a preserved ejection fraction, regardless of the presence or absence of diabetes. In June 2022 the MHRA approved a change in the licence of empagliflozin – it is now indicated for the treatment of symptomatic chronic heart failure and data from the EMPEROR-preserved trial was added to the SmPC. However this is not currently reflected in the NICE guidelines, so the use of SGLT-2 inhibitors is at the discretion of the clinician and funding is not agreed in many ICS/ICBs. Those familiar with treating this group of patients may choose to prescribe it, but those who see them less often are very unlikely at present to prescribe it.
18a. If not, how could the results be extrapolated to the UK setting?	Update the NICE guidance to facilitate prescribing empagliflozin in this patient cohort.
18b. What, in your view, are the most important outcomes, and were they measured in the trials?	Combined risk reduction in cardiovascular death and hospitalization for heart failure. Both measured in the trial. Alongside this, an improvement in patient quality of life, using a validated tool.

	In patients with HFpEF, SGLT-2 inhibition led to a 21% lower relative risk in the composite of cardiovascular death or hospitalization from heart failure. There was a 29% lower risk of hospitalization for heart failure. These results were consistent in those with or without diabetes, and across the range of preserved ejection fractions. Patients treated with empagliflozin had significant improvement in the KCCQ score (quality of life) versus placebo. This improvement was seen early in treatment and sustained for 1 year.
18c. If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
18d. Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	None known
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. How do data on real-world experience compare with the trial data?	So far patients initiated on an SGLT2i for HFrEF generally tolerate it well and feel a benefit from as early as 4 weeks. Elderly and frail patients often require diuretic dose cessation or reduction than those represented in the trial but still tolerate it well.

Equality

21a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
21b. Consider whether these issues are different from issues with current care and why.	N/A

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• Patients with HFrEF and HFpEF are a large population which have a poor quality of life and life expectancy and currently have no pharmacological treatment available with evidence of reduction in hospitalisation, improvement in quality of life or CV death • The prevalence of this patient cohort is expending as we have a growing older, frailer and obese population which will contribute to greater demands on hospital services and hospital admissions. • Empagliflozin was the first pharmacological therapy in a RCT to show a significant reduction in the primary end point to reduce HF hospitalisations and CV death • Empagliflozin has the ability to make a huge impact on how we manage this cohort of patients and hopefully lead to investment in services to manage these patients appropriately • Increased usage of empagliflozin in HF patients across the spectrum of EF will increase clinician knowledge and confidence to prescribe this treatment for all indications.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

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.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Patient expert statement

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 heart failure with preserved or mildly reduced ejection fraction or caring for a patient with chronic heart failure with preserved or mildly reduced ejection fraction. The text boxes will expand as you type.

In part 2 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).

NICE National Institute for Health and Care Excellence

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.

The deadline for your response is **5pm** on **Friday 28 October**. 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 heart failure with preserved or mildly reduced ejection fraction

Table 1 About you, chronic heart failure with preserved or mildly reduced ejection fraction, current treatments and equality

1. Your name	
2. Are you (please tick all that apply)	☐ A patient with chronic heart failure with preserved or mildly reduced ejection fraction? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with chronic heart failure with preserved or mildly reduced ejection fraction? ☒ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	
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 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 heart failure with preserved or mildly reduced ejection fraction? **If you are a carer (for someone with chronic heart failure with preserved or mildly reduced ejection fraction) please share your experience of caring for them**	I volunteer for Pumping Marvellous as a patient educator & speak to patients with HFpEF every day.
7a. What do you think of the current treatments and care available for chronic heart failure with preserved or mildly reduced ejection fraction on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	**The current treatments for HFpEF are currently very limited.**
8. If there are disadvantages for patients of current NHS treatments for chronic heart failure with preserved or mildly reduced ejection fraction (for example, how they are given or taken, side effects of treatment, and any others) please describe these	There are very few treatments available that offer benefits to prognosis, QoL or offer hope to patients with HFpEF.
9a. If there are advantages of empagliflozin 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?	**The advantages are that we need more medication choices for people with HFpEF** **QoL would probably be the most beneficial to patients.**

9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why? **9c. Does empagliflozin 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 empagliflozin over current treatments on the NHS please describe these. For example, are there any risks with empagliflozin? If you are concerned about any potential side effects you have heard about, please describe them and explain why	We need more choices & this drug offers more advantages than disadvantages. The fact that it can drop blood pressure is the only disadvantage I've come across.
11. Are there any groups of patients who might benefit more from empagliflozin 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	It's about choice. The more alternatives for patients the better.
12. Are there any potential equality issues that should be taken into account when considering chronic heart failure with preserved or mildly reduced ejection fraction and empagliflozin? 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	Lack of choice, research & prescribing inequalities It would be beneficial if primary care practitioners could prescribe these drugs to HF patients in the same way they do to diabetics.

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?	

Part 2: 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

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Patient expert statement

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 heart failure with preserved or mildly reduced ejection fraction or caring for a patient with chronic heart failure with preserved or mildly reduced ejection fraction. The text boxes will expand as you type.

In part 2 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).

NICE National Institute for Health and Care Excellence

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.

The deadline for your response is **5pm** on **Friday 28 October**. 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 heart failure with preserved or mildly reduced ejection fraction

Table 1 About you, chronic heart failure with preserved or mildly reduced ejection fraction, current treatments and equality

1. Your name	Nick Hartshorne-Evans
2. Are you (please tick all that apply)	☐ A patient with chronic heart failure with preserved or mildly reduced ejection fraction? ☒ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with chronic heart failure with preserved or mildly reduced ejection fraction? ☒ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	Pumping Marvellous Foundation
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 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 heart failure with preserved or mildly reduced ejection fraction? **If you are a carer (for someone with chronic heart failure with preserved or mildly reduced ejection fraction) please share your experience of caring for them**	I was diagnosed with Heart Failure in 2010 and have lived with it since. I am a recovered heart failure patient with reduced ejection fraction. I am however the Founder and CEO of the Pumping Marvellous Foundation, and we represent patients with **all types** of heart failure across our communities and the UK. The signs, symptoms, and disease burden of all types of heart failure are very similar. There is a system, treatment and care access and equity difference between HFrEF and HFmrEF and HFpEF
7a. What do you think of the current treatments and care available for chronic heart failure with preserved or mildly reduced ejection fraction on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	There are no guidelines or prognostically available treatments for people living with chronic HFpEF in the NHS. This is unacceptable and demonstrates the largest unmet need for patients living with heart failure. If the prevalence of HFpEF in the total UK population of all heart failure is 40% of 920,000 (2018 figures NICE) then there are just under 400,000 people in the UK at a severe disadvantage.
8. If there are disadvantages for patients of current NHS treatments for chronic heart failure with preserved or mildly reduced ejection fraction (for example, how they are given or taken, side effects of treatment, and any others) please describe these	There are no prognostically beneficial treatments for HFpEF patients There are no guidelines for HFpEF patients HFpEF patients access to Heart Failure Nurses and specialist MDT services is patchy at best. Commissioners of services do not commission services for HFpEF patients because of the lack of an evidence base in favour of HFrEF patients. HFpEF patients in the main are prescribed a diuretic for symptom relief and referred into Primary Care. Primary Care is not geared to treating or optimising patients with

	HFpEF. Many patients feel as though they are just left to wallow with nobody understanding how to help them. **The patient cohort for HFpEF is significant. If this was happening in Cancer there would be National outrage.**
9a. If there are advantages of empagliflozin 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 empagliflozin 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**	**There are no current treatments available to HFpEF patients in the NHS** **Mortality –** The primary endpoint of the study demonstrated statistically significant benefit on cardiovascular death and hospitilisation of heart failure. Both components contributed and if the trial was extended anecdotally, you may anticipate mortality benefit. **Hospital readmission –** There was a positive impact on hospital admissions and type of admission. For example, there was a 33% reduction in the total number of hospitalisations requiring intravenous diuretics (EMPRERROR – Preserved). Also, there was a sustained reduction in the risk and severity of heart failure events. **Quality of Life –** There was a statistically relevant benefit over the placebo arm when measured by KCCQ health questionnaire across all 3 domains CSS / TSS / OSS – the effect was apparent from just 3 months and sustained across the trial timeline. This positively impacted patients NYHA classification. (EMPEROR – Preserved) Each one of the endpoints are equally important to the variety of individual stakeholders. For the patient, quality of life is very important and has equal standing to Mortality. As there seemed not to be statistically benefit for Mortality there was a reduction in seriousness of heart failure events which has to have had a beneficial impact on the patient's well-being. The overriding advantage is that there are now treatments for people with HFpEF and as there was statistically relevant benefit across 2 domains, fundamentally this is important as it gives healthcare teams a treatment option for treating HFpEF and HFmrEF.

	Empagliflozin, without question, overcomes and address the current treatment drought.
10. If there are disadvantages of empagliflozin over current treatments on the NHS please describe these. For example, are there any risks with empagliflozin? If you are concerned about any potential side effects you have heard about, please describe them and explain why	There are no current treatments on the NHS. Empagliflozin is well tolerated with limited side-effects. I have no concerns about side effects as long as the patient is aware of them and they are dealt with by their healthcare team.
11. Are there any groups of patients who might benefit more from empagliflozin 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	All patient with all heart failure types benefit. Those with heart failure who do not have Type II Diabetes and reduced Kidney Function must benefit. The tablet is easy to take and should not disrupt the patients' other medications. It is well tolerated.
12. Are there any potential equality issues that should be taken into account when considering chronic heart failure with preserved or mildly reduced ejection fraction and empagliflozin? 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 More information on how NICE deals with equalities issues can be found in the NICE equality scheme	The system and process for prescribing may disadvantage and call into question whether all patients would have equal access and equity of opportunity to be prescribed. GP's know SGLT2i's very well, they have been prescribed without specialist involvement in Type II Diabetes for many years. **There should be no reason to refer for specialist reassessment or advice when prescribing SGLT2i's in Primary Care.** Referring for specialist assessment and or initiation is just another burden to the NHS where – Waiting times increase Specialist caseloads increase Patients suffer Time is important when prescribing HF medications therefore delay is detrimental to an already under invested population.

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

NICE National Institute for Health and Care Excellence

Part 2: 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

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Clinical expert statement

Thank you for agreeing to provide your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. 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 for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

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.

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.

NICE National Institute for Health and Care Excellence

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 replaced with the following text: 'academic/commercial in confidence information removed'. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for your response is **5pm** on **Wednesday 16 November**. 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: Treating chronic heart failure with preserved or mildly reduced ejection fraction and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Lisa Anderson
2. Name of organisation	St George's University Hospitals NHS Foundation Trust
3. Job title or position	Consultant Cardiologist and Heart Failure Lead
4. Are you (please tick all that apply)	☒ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with chronic heart failure with preserved or mildly reduced ejection fraction? ☒ A specialist in the clinical evidence base for chronic heart failure with preserved or mildly reduced ejection fraction or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	

8. What is the main aim of treatment for chronic heart failure with preserved or mildly reduced ejection fraction? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	The main goals of treatment for chronic heart failure with preserved or mildly reduced EF are to: Improve quality of life for patients Prevent hospital admissions Reduce cardiovascular mortality
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	A significant improvement in quality of life with a validated scoring tool. Significantly reduced hospital admissions. Significantly reduced cardiovascular mortality.
10. In your view, is there an unmet need for patients and healthcare professionals in chronic heart failure with preserved or mildly reduced ejection fraction?	Yes. Approximately half of patients with HF have a preserved or mildly reduced left ventricular ejection fraction (HFpEF/HFmrEF). There is a high symptom burden with frequent hospital admissions and increasing frailty as a result. Until now, clinical trials of new therapeutic approaches have been characterised by efficacy failure, and treatment options remain very limited.
11. How is chronic heart failure with preserved or mildly reduced ejection fraction currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	HFmrEF (EF 41-49%) – No RCT has been performed exclusively in this subgroup. However, because -EF in heart failure is a spectrum and -due to the large benefits seen in patients with more reduced EF, - and because many of the patients in this cohort are believed to be patients with recovering EF, the European Society of Cardiology Guidelines (2021) has made 2b recommendations (these drugs may be considered) for ACE inhibitors/Angiotensin II receptor blockers/neprilysin inhibitors, beta-blockers and mineralocorticoid receptor antagonists for treatment in this subgroup. HFpEF (EF > 50%) – Treatment is focussed on diuretic therapy and managing patient comorbidities such as atrial fibrillation, diabetes, hypertension, kidney disease. Weight loss in obese patients and increasing exercise may improve symptoms and exercise capacity.

12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	The healthcare resource use does not differ from current care. Following initiation, the vast majority of patients require only routine monitoring. A subgroup of more complex diabetic patients will require increased home blood glucose checks for 1 week after initiation and recheck HbA1C at 3 months. The technology will be used in all areas where patients are seen – specialist care, and primary and secondary care following recommendation from a HF specialist. This technology is already used in the management of HF patients with reduced ejection fraction and in type 2 diabetes and is also licensed for chronic kidney disease. Little investment, other than the writing of Local Guidelines for use, would be needed.
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Yes, I expect the technology to provide clinically meaning ful benefits compared with current care. Although a trend toward reduced cardiovascular mortality is seen, most of the effect on the primary end point was seen in reduced HF admissions. A highly significant improvement in the KCCQ QOL score was seen so I expect the technology to increase health related quality of life more than current care. This improvement was seen early (by 12 weeks) and persisted at 1 year.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No. The effect of empagliflozin on the incidence of primary outcome events was generally consistent across prespecified subgroups, including patients with or without diabetes at baseline.
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than	Heart failure admissions increased by 33% in the 5 years pre-pandemic with the largest increases in HFpEF admissions and HF is the commonest cause for

current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	hospital admission in those >65years. NHS Hospitals are at capacity and a treatment that has a positive impact on HF admissions will help HF patients, overstretched HF clinical teams as well as the wider health system.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	No additional testing is required before starting treatment and the treatment will be ongoing indefinitely once initiated.
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	Reduced hospital admissions will greatly impact quality of life for both patients and families.
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Yes. Until now all no evidence-based therapy has been available for HFpEF/HFmrEF patients. The therapy addresses the major unmet needs of reducing hospital admissions and improving quality of life.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Serious adverse events occurred in 1436 patients (47.9%) in the empagliflozin group and in 1543 patients (51.6%) in the placebo group. Patients are warned about the potential increase in genitourinary fungal infections and the need for sick day rules to reduce the risk of diabetic ketoacidosis.

20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Yes. The most important outcomes were measured in the trial (QOL, HF hospitalisations and CV death). I am not aware of adverse events not apparent in the clinical trials that have come to light subsequently.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
22. How do data on real-world experience compare with the trial data?	Since the publication of the data, it is likely that this medication has already been initiated for many admitted HFpEF patients. Many of these patients already meet other indications for SGLT2- (type 2 diabetes or CKD with proteinuria). The medication is well tolerated – in particular, given the frail, comorbid population, there is minimal effect on blood pressure or worsening of renal function.
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? 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.	No

Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. 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.	

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

At present, the disease trajectory and quality of life for patients with HFpEF and HFmrEF is poor.

There are currently no pharmacological treatment options shown to reduce hospital admission or improve quality of life for these patients

This technology will make a real and meaningful difference to NHS care for patients with HFmrEF and HFpEF

In the UK there are around 100,000 HF admissions annually, with a long length of stay (10 days mean), so a technology with an impact on reduced admissions will have wider benefits for an NHS system currently running at capacity.

Prevalence of HFmrEF and HFpEF is increasing in the UK, and these subgroups represent a large and growing proportion of heart failure admissions to hospital.

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 our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Clinical expert statement

Thank you for agreeing to provide your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature. 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 for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

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.

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.

NICE National Institute for Health and Care Excellence

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 replaced with the following text: 'academic/commercial in confidence information removed'. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for your response is **5pm** on **Friday 28 October**. 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: Treating chronic heart failure with preserved or mildly reduced ejection fraction and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Andrew Ludman
2. Name of organisation	British Cardiovascular Society
3. Job title or position	Consultant Cardiologist
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with chronic heart failure with preserved or mildly reduced ejection fraction? ☐ A specialist in the clinical evidence base for chronic heart failure with preserved or mildly reduced ejection fraction or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

8. What is the main aim of treatment for chronic heart failure with preserved or mildly reduced ejection fraction? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	Main aim depends on view point. Key aims from a healthcare provider perspective are to reduce hospital admission and cardiovascular mortality. From a patient perspective reduction in symptoms of breathlessness is very important.
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	Any reduction in hospital admission or mortality is welcome and is significant for that patient.
10. In your view, is there an unmet need for patients and healthcare professionals in chronic heart failure with preserved or mildly reduced ejection fraction?	Yes. There are few (if any) evidence based treatments in this condition.
11. How is chronic heart failure with preserved or mildly reduced ejection fraction currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	Treatment guidelines are written by the European Society of Cardiology as part of the overall heart failure guideline. The mainstay of treatment for HFpEF has been treatment of the contributing co-morbidities (e.g. hypertension, rate control of atrial fibrillation etc) as well as fluid balance management with diuretics. There is some evidence for spironolactone. The diagnostic pathway is defined via the investigation of heart failure NICE guideline in the UK. However the diagnosis is not always easy. The SGLT2i are really the first medication in this condition to demonstrate a significant benefit. Therefore this group of medications is likely to be adopted widely, with hopefully the same real-life benefit.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	The SGLT2i medications are already used for a number of indications within the NHS and so their use could be adapted safely and rapidly if approved.

• How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	There is likely to be a resource implication in terms of higher medication cost, albeit somewhat balanced by a reduction in hospital admission and the quality of life benefit around symptoms. I would suggest that empagliflozin could be used in line with SGLT2i for HFrEF which is prescribed in primary care following advice of a specialist heart failure team member. Alerting healthcare professionals to the new guidance and providing some education may be required.
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	There is no conclusive evidence of a decrease in overall mortality in the main current study of empagliflozin in HFpEF. Health related QoL is likely to be increased in comparison to current care with a reduction in the risk of heart failure worsening or hospitalisation and a decrease in symptoms (as measured by KCCQ score).
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	The majority of trial participants have a white ethnicity with smaller numbers of other ethnic groups. No clinical difference in response between groups has been detected. Further evaluation may allow confirmation of equal clinical effect in all.
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient	Straightforward usage for primary and secondary care professionals.

acceptability or ease of use or additional tests or monitoring needed)	
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	A diagnosis of heart failure with preserved or mildly reduced ejection fraction should be made. Symptomatic (NYHA II or greater).
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	No
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Yes this a step change in management. The first medication to show a meaningful difference in clinical outcomes for HFpEF. Patients with HFpEF have a significant unmet need in terms of treatments to improve symptoms, quality of life and reduce deterioration. The SGLT2i go someway towards this.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The incidence of adverse effects is similar to placebo. For empagliflozin a small increase in uncomplicated urinary infections was reported in the main study in this group of patients.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting?	Yes, the clinical trials reflect UK practice. The most important outcomes were assessed in the clinical trial.

• What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No additional adverse events have come to light.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
22. How do data on real-world experience compare with the trial data?	SGLT2i are used for a number of indications already and real world experience is similar to that presented in the trials. Patients and professionals are concerned about the risk of urinary infection and it is difficult to balance the relative risks/benefits around this.
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? 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. Please state if you think this evaluation could	Patients with HFpEF are often older, may have multiple medical problems and a higher degree of frailty and as such are often harder to reach with new medical innovations. Where possible specific evidence based recommendations for this group would be useful.

• exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. 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.	

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

SGLT2i (specifically empagliflozin and dapagliflozin) are already approved for treatment of heart failure with reduced ejection fraction.
There is robust clinical trial evidence of benefit for empagliflozin and dapagliflozin in the treatment of heart failure with preserved ejection fraction.
There are few if any other specific treatments for heart failure and preserved ejection fraction.
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 our privacy notice.

BMJ TAG

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction ID3945

STA Report

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135656.

Title:	Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction
Produced by:	BMJ Technology Assessment Group (BMJ-TAG)
Authors:	Steve Edwards, Director of Health Technology Assessment, BMJ-TAG, London Alex Allen, Senior Clinical Evidence Analyst, BMJ-TAG, London Mariana Bacelar, Principal Health Economist, BMJ-TAG, London
Correspondence to:	Steve Edwards, BMJ TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	24/11/2022
Source of funding:	This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135656.
Declared competing interests of the authors	No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The EAG would like to thank Dr James Gamble Consultant Cardiologist at Oxford University hospitals NHS Foundation Trust and Dr Geraint Morton Consultant Cardiologist at Portsmouth Hospitals University NHS Trust for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report. The EAG would also like to thank Dr Andrew Turley Consultant Cardiologist at South Tees NHS Foundation Trust for providing feedback on the clinical sections of the report.
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.
Report reference:	Edwards SJ, Allen A, Bacelar M. Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction: A Single Technology Appraisal. BMJ Technology Assessment Group, 2022

Contribution of authors:

Steve Edwards	Critical appraisal of the company's submission; validated the statistical analyses; provided feedback on all versions of the report. Guarantor of the report
Alexander Allen	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and drafted the summary, background and clinical results sections
Mariana Bacelar	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections

All authors read and commented on draft versions of the EAG report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACEI	Angiotensin-converting enzyme inhibitor
AE	Adverse event
AESI	Adverse event of special interest
AF	Atrial fibrillation or atrial flutter
AICc	Akaike Information Criterion with a correction for finite sample size
ARBs	Angiotensin receptor blockers
ARNI	Angiotensin receptor-neprilysin inhibitor
BMI	Body mass index
CEA	Cost-effectiveness analysis
CKD	Chronic kidney disease
CKD-EPI	Chronic kidney disease - Epidemiology collaboration equation
COPD	Chronic obstructive pulmonary disease
CPRD	Clinical Practice Research Datalink
CRM	Cardio-renal-metabolic
CRS	Cardio-renal syndrome
CSR	Clinical study report
CT	Computerised tomography
CV	Cardiovascular
DM	Diabetes Mellitus
EAG	External assessment group
NICE	National Institute of Health and Care Excellence
ECG	Electrocardiogram
eGFR	Estimated glomerular filtration rate
EF	Ejection fraction
ESC	European Society of Cardiology
EMA	European Medicines Agency
GP	General practitioner
HbA1c	Glycated haemoglobin
HES	Hospital Episode Statistics
HF	Heart failure
HFmrEF	Heart failure with mildly reduced ejection fraction
HFpEF	Heart failure with preserved ejection fraction
HFrEF	Heart failure with reduced ejection fraction

HHF	Hospitalisation for heart failure
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
IHD	Ischaemic heart disease
KCCQ-CSS	Kansas City Cardiomyopathy Questionnaire - clinical summary score
KCCQ-TSS	Kansas City Cardiomyopathy Questionnaire - total symptom score
KCCQ-OSS	Kansas City Cardiomyopathy Questionnaire - overall summary score
EF	Left ventricular ejection fraction
MA	Marketing authorisation
MACE	Major adverse cardiovascular events
MI	Myocardial infarction
MRA	Mineralocorticoid receptor antagonists
MRI	Magnetic resonance imaging
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NYHA	New York heart association
RCT	Randomised controlled trial
RS	Randomised set
RWE	Real-world evidence
SAE	Serious adverse event
SD	Standard Deviation
SBP	Systolic blood pressure
SGLT1	Sodium-glucose co-transporter-1
SGLT2i	Sodium-glucose co-transporter-2 inhibitor
SLR	Systematic literature review
SmPC	Summary of medicinal product characteristics
SoC	Standard of care
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
TS	Treated set
UK	United Kingdom

1 Executive summary

This summary provides a brief overview of the key issues identified by the evidence review group (EAG) as being potentially important for decision making. 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 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.5 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of NICE.

1.1 Overview of the EAG's key issues

Table 1. Summary of key issues

Issues	Summary of issue	Report sections
Issue 1	Estimation of KCCQ-CSS transition probabilities in the economic model	4.2.6.1, 4.2.6.2
Issue 2	The long-term effect of empagliflozin on patients' KCCQ-CSS scores	4.2.6.1, 4.2.6.2
Issue 3	Estimation of HHF in the economic model	4.2.6.3, 4.2.6.4
Issue 4	The impact of empagliflozin on patients' survival	4.2.6.7,4.2.6.8
Issue 5	The impact of the duration of HHF events on patients' quality of life	4.2.8,4.2.8.1
Issue 6	Overestimation of costs associated with HFF events and CV deaths	4.2.9.2,4.2.9.3,4.2.9.4,4.2.9.5

Abbreviations: KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score; HHF, hospitalisation for heart failure

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are the choice of transition probabilities between the KCCQ-CSS states of the model; the fact that the EAG considers the impact of empagliflozin on patients' survival to be uncertain; the suitability of KM HHF data to be used in the model; and the assumption around the duration of HHF events on patients' quality of life.

1.2 Overview of key model outcomes

National Institute for Health and Care Excellence (NICE) technology appraisals compare how much a new technology improves length (overall survival) and quality of life 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:

- Directly increasing the proportion of patients who remain in the better KCCQ-CSS states, which in its turn leads to better quality of life, and indirectly leads to better survival and lower hospitalisation rates.
- Directly decreasing the probability of patients being hospitalised for heart failure.
- Directly decreasing the probability of patients dying.

Overall, the technology is modelled to affect costs by:

- Its higher unit cost compared to standard of care (SoC) alone.
- Decreasing the probability of patients being hospitalised for heart failure.
- Decreasing the probability of patients having cardiovascular deaths.

The modelling assumptions that have the greatest effect on the ICER are:

- The transition probabilities used to estimate patients' distribution across the KCCQ-CSS states of the model.
- The assumptions made around the impact of empagliflozin on survival.
- The duration of the impact of HHF on patients' quality of life.

1.3 Summary of the EAG's clinical and economic key issues

The EAG's key issues on the clinical and cost-effectiveness evidence are given in Table 2 to Table 7.

Table 2. Issue 1. Estimation of KCCQ-CSS transition probabilities in the economic model

Report section	4.2.6.1, 4.2.6.2, 3.2.8.10
Description of issue and why the EAG has identified it as important	The EAG disagrees with the last observation carried forward (LOCF) imputation method used by the company to handle missing KCCQ-CSS data and prefers the use of raw observed KCCQ-CSS data from the trial. The number of observations for patients' KCCQ-CSS scores without imputations provides a robust sample size and is similar across treatment arms in EMPEROR-Preserved, suggesting that the data are well balanced. Given that the EAG is unclear when end of treatment occurred for patients with missing data, and the lack of any data to validate the underlying assumption of the LOCF method that the missing observations would produce identical KCCQ-CSS scores from the one captured in the previous data points, the EAG considers that the more robust approach is to use observed data without imputation.
What alternative approach has the EAG suggested?	To use the raw observed TPs between KCCQ-CSS derived from EMPEROR-Preserved.
What is the expected effect	The TPs derived with the two methods are substantially different, particularly

on the cost-effectiveness estimates?	the set of TPs used from month 9 onwards in the model. The scenario analysis using observed TPs between KCCQ-CSS quartiles leads to an increase in the company's base case ICER from £14,429 to £20,198 per QALY gained.
What additional evidence or analyses might help to resolve this key issue?	The EAG remains unclear how much data were imputed for each timepoint and for how long observations were carried forward using the LOCF method. The EAG asks that the company provides the distribution of the missing KCCQ-CSS observations, with a description of when end of treatment occurred for these patients. These data should help clarifying the number of missing observations and when these occurred (and thus for how many timepoints the LOCF imputed values were used).

Abbreviations: EAG, evidence assessment group; ICER, incremental cost-effectiveness ratio; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score; SoC, standard of care; TP, transition probability.

Table 3. Issue 2. The long-term effect of empagliflozin on patients' KCCQ-CSS scores

Report section	4.2.6.1, 4.2.6.2
Description of issue and why the EAG has identified it as important	The effect of empagliflozin on KCCQ-CSS (sustained by the combination of the proportion of patients in the better KCCQ-CSS states in the empagliflozin arm at month 8 and the low probability of disease progression for both SoC and empagliflozin arms in month 9+) leads to sustained treatment effect over time, which is unlikely to be clinically plausible. The company's assumption that empagliflozin patients experience SoC TPs after discontinuation is only partially conservative and leads to a sustained relative treatment effect for patients in KCCQ-CSS 4 in the model over time. Due to the company's model structure, this assumption impacts the benefits associated with empagliflozin on HHF and mortality, as these outcomes are dependent on patients' distribution across KCCQ-CSS states.
What alternative approach has the EAG suggested?	The company included a scenario analysis in the model which assumed that the TPs between KCCQ-CSS quartiles for patients on treatment are the same as those for patients off treatment from month 9 onwards in the model, which equates to the proportion of patients in each KCCQ-CSS state being the same in the empagliflozin and SoC arms of the model approximately 4 to 5 years after treatment initiation in the model. This scenario is likely to be overly pessimistic as it assumed that after 4 years empagliflozin does not impact patients' KCCQ-CSS anymore, even if patients are still on treatment.
What is the expected effect on the cost-effectiveness estimates?	The company's scenario increases the base case ICER to £32,482.
What additional evidence or analyses might help to resolve this key issue?	Longer follow-up data on the impact of empagliflozin on patients' KCCQ-CSS.

Abbreviations: ERG, evidence review group; HHF, hospitalisation for heart failure; ICER, incremental cost-effectiveness ratio; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary score; SoC, standard of care; TP, transition probability.

Table 4. Issue 3. Estimation of HHF events in the economic model

Report section	4.2.6.3, 4.2.6.4

Description of issue and why the EAG has identified it as important	In EMPEROR-Preserved, [redacted]. The results reported in the EMPEROR-Preserved CSR, not only indicate that the difference across arms in patients with first events was [redacted] than the difference in the number of patients with second events [redacted], but also that it is likely that empagliflozin does not provide a benefit in preventing subsequent hospitalisation events. By considering all events from the trial to have been first events in the model the company is, therefore, likely overestimating the benefit of empagliflozin on reducing HHF events. The EAG also notes that second events occurred "faster" in relation to first HHF events - out of those patients with 2 (or more) events, approximately [redacted] of patients had already experienced a second event at 3 years after the first event, whereas only about [redacted] of patients had experienced their first HHF event at 3 years. Therefore, the EAG's view remains that the company's approach does not appropriately capture the hospitalisations in EMPEROR-Preserved, both in the assumption of a constant hospitalisation rate; a constant treatment effect; and in the decision to not separate first and subsequent hospitalisation events. Given that KM data on HHF in EMPEROR-Preserved was available for first and subsequent hospitalisations, the EAG's view remains that using these data would have allowed the company to fit a parametric survival curve to the data and extrapolate into the model's time horizon without having to assume a constant rate of HHF and without having to assume a constant treatment effect with empagliflozin. The EAG notes that using survival curves would have still allowed the company to model HHF by KCCQ-CSS state. Finally, the EAG considers that not only the absolute number of HHF events in the model are considerably overestimated in relation to the events observed in the trial for the same period of time, as expected from the company's assumption of a constant rate of HFFs in the model, but the overestimation increases as time progresses in the model. Crucially, even though the observed and predicted differences in events between empagliflozin and placebo are broadly similar at 26 months and at 3 years (indicating less uncertainty in the incremental results of the model), the comparison of incremental estimated vs observed events at the last available time from the trial suggests that the model increasingly overestimates the benefit associated with empagliflozin.
What alternative approach has the EAG suggested?	The company reported fitting six distributions (exponential, Weibull, lognormal, log-logistic, Gompertz, and generalised Gamma) to time to first HHF KM data, with treatment arm as the only predictor. The company reported that the extrapolated hazard estimates were decreasing or plateauing hazards, which the company considered clinically implausible. The EAG notes that it has not seen any measures of fit (statistical or visual) to the different models used by the company, so it cannot assess the fit of the models, and the respective underlying hazard. Crucially, the company only fitted first HHF with parametric models, and did not undertake the same analysis for subsequent events. The EAG does not understand the company's decision, and notes that it is possible that the trend in the risk of subsequent events could have been different (for example, increasing),

	which could have added plausibility to the data analysis. Therefore, the EAG maintains its view that using HHF KM data from EMPEROR-Preserved is likely to provide a more robust approach to modelling time to HHF in the model, separately by first and subsequent HHF event.
What is the expected effect on the cost-effectiveness estimates?	The ICER will increase as the overall and incremental number of HHF decreases in the model.
What additional evidence or analyses might help to resolve this key issue?	The company should provide the additional analysis requested above, by first and subsequent event. The company should also provide measures of fit (statistical and visual) of the parametric used in the analysis and if needed, consider the use of different survival models.

Abbreviations: ICER, incremental cost-effectiveness ratio, KCCQ-CSS; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary scores; HHF, hospitalisation for heart failure

Table 5. Issue 4. The impact of empagliflozin on patients' survival

Report section	4.2.6.7,4.2.6.8
Description of issue and why the EAG has identified it as important	The EAG considers that the KM curves from EMPEROR-Preserved (both for overall and CV survival) do not provide sufficient evidence to substantiate empagliflozin having an impact on patients' survival compared to SoC patients. However, clinical expert opinion consistently reported the plausibility of empagliflozin reducing patients' CV-related mortality. The EAG disagrees with the company's approach of including KCCQ-CSS as predictors of survival in the all-cause mortality risk equations as these generate an indirect survival benefit for empagliflozin in the model. Given the KM overall survival data from EMPEROR-Preserved show that survival was similar in the two arms of the trial, the EAG considers that including a treatment benefit associated with empagliflozin on overall survival is unsubstantiated. The EAG also has some concerns regarding the company's inclusion of a treatment effect in the risk equations to estimate CV mortality in the model. Even though the EAG's clinical experts agreed with the clinical plausibility of empagliflozin being associated with a survival benefit on CV-related death, the EMPEROR-Preserved trial showed a non-statistically significant effect on this outcome. Furthermore, given the company's decision to include a treatment effect in the CV-mortality risk equations, the decision to do so through a proportional hazards model is problematic - considering the shape of the KM curves for CV mortality in EMPEROR-Preserved, it is implausible that a constant treatment effect would be observed throughout the trial (and extrapolated) period. The absolute and incremental number of CV deaths are overestimated in the model, with a trend suggesting that the overestimation increases as the model progresses. All-cause deaths are also overestimated in the model.
What alternative approach has the EAG suggested?	The EAG requested that the company included a scenario analysis in the model where both the direct and indirect treatment effect of empagliflozin on survival was removed from the model, for all-cause mortality and CV-related mortality, separately.

What is the expected effect on the cost-effectiveness estimates?	When the EAG assumed no survival benefit for empagliflozin on all-cause mortality, the model predictions in terms of number of deaths avoided in the empagliflozin arm at 3.5 years is more aligned with those observed in the trial at the same time than those predicted by the model when empagliflozin is assumed to have an impact on patients' overall. On the contrary, when the EAG assumed no survival benefit for empagliflozin on CV-related mortality, the model predictions considerably underestimated the number of CV deaths avoided in the empagliflozin arm of the trial, therefore, indicating that this might not be a plausible assumption. Removing the survival benefit associated with empagliflozin from the model leads to an increase in the final ICERs. The EAG caveats this analysis by the fact that it does not include the cost of non-CV deaths (please see point below).
What additional evidence or analyses might help to resolve this key issue?	The EAG considers it crucial that the committee validates the following assumptions: 1.Is empagliflozin likely to reduce CV mortality compared to SoC? a) If the answer to the above question is yes, then the EAG notes that there are two possible implications for non-CV deaths in this population: i. Either the CV deaths prevented in the empagliflozin arm do not translate into a reduction in overall mortality for these patients - given the age of the HFmrEF/HFpEF population, and the presence of several comorbidities, the proportion of patients who don't die of a CV cause die in a similar time frame of a non-CV cause, therefore suggesting that there will be more non-CV deaths for empagliflozin than SoC patients – in this case, it is crucial that the cost of a CV and non-CV death is incorporated into the economic analysis, to assess the cost-effectiveness of "replacing" an equal number of CV deaths with non-CV deaths for empagliflozin patients. Or; ii. The CV deaths prevented in the empagliflozin arm translate into a reduction in overall mortality for these patients vs SoC patients – the proportion of patients who don't die of a CV cause end up dying much later of a non-CV cause, and a similar proportion of empagliflozin and SoC patients die of non-CV causes in the shorter term. b) If the answer to question 1 is no, then there is no difference in CV or non-CV deaths for empagliflozin and SoC patients.

Abbreviations: CV, cardiovascular; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary scores; KM, Kaplan-Meier; OS, overall survival; SoC, standard of care; UK, United Kingdom.

Table 6. Issue 5. The impact of the duration of HHF events on patients' quality of life

Report section	4.2.8,4.2.8.1
Description of issue and why the EAG has identified it as important	The EAG considers that the impact of HHF on patients' quality of life is overestimated in the model given the company's assumption that all HHFs in the model impact patients' quality of life for 1 year after the event. Clinical expert opinion provided to the EAG indicated that a reasonable assumption is that 1 day in hospital impacts patients' quality of life for 1 week after discharge. The other clinical expert indicated that 6 months of

	impact (as a maximum) could also be plausible after discharge.
What alternative approach has the EAG suggested?	The EAG requested that the company conducted two alternative scenario analyses where: 1. It was assumed that HHF events impact patients' QoL for 2.75 months after discharge (corresponding to an impact of 11 weeks after being hospitalised for 11 days, the mean hospitalisation time in EMPEROR-preserved for heart failure). 2. It was assumed that HHF events impact patients' QoL for 6 months after discharge.
What is the expected effect on the cost-effectiveness estimates?	The ICER for the first scenario increases to £17,912 per QALY gained and to £16,511 for the second scenario (compared to the base case ICER of £14,429).
What additional evidence or analyses might help to resolve this key issue?	None.

Abbreviations: CV, cardiovascular; EMPEROR-R, EMPEROR-Reduced; ERG, evidence review group; HFrEF, heart failure with reduced ejection fraction; ICER, incremental cost-effectiveness ratio; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire clinical summary scores; KM, Kaplan-Meier; SoC, standard of care; UK, United Kingdom.

Table 7. Issue 6. Overestimation of costs associated with HFF events and CV deaths in the model

Report section	4.2.9.2,4.2.9.3,4.2.9.4,4.2.9.5
Description of issue and why the EAG has identified it as important	The mean duration of HHF in EMPEROR-Preserved was 11 days; with a median of 8 days; and Q3 of 13 days. The EAG is therefore, concerned with the company's use of the more severe cost code (EB03A), associated with a 53-day long hospitalisation, which is likely to overestimate HHF costs in comparison to EMPEROR-Preserved. The EAG also disagrees with the use of the chosen estimates from Alva *et al.* 2015 to estimate the cost of CV deaths as these relate to the added costs on hospitalisations due to T2DM complications. Finally, the EAG considers that the costs of CV death are further overestimated given that 46% of deaths in EMPEROR-Preserved were sudden cardiac deaths (likely to be associated with no additional cost), and that the company did not consider these in the weighted costs of CV death.
What alternative approach has the EAG suggested?	Given the less severe cost code (EB03E) is associated with 13 days in hospital, the EAG asked that the company conducted a scenario analysis in the model where the cost of £2,062 was used to calculate the cost of all HHFs events in the model. The EAG also asked that the company used the alternative estimates provided in Table 3 of Alva *et al.*, which reported the absolute cost of CV fatal events. Finally, the company also included two additional options in the model: a conservative approach where the cost of sudden cardiac death was zero, decreasing the overall costs of CV death to £1,452; and a second option where the total HRG costs for cardiac arrest of £1,632 was used, leading to overall costs of CV death of £2,345.
What is the expected effect on the cost-effectiveness estimates?	Decreasing the costs associated with HHF events and CV deaths leads to an increase in the ICERs.

What additional evidence or analyses might help to resolve this key issue?	None.

Abbreviations: EMPEROR-R, EMPEROR-Reduced; ERG, evidence review group; HHF, hospitalisation for heart failure; QALY, quality adjusted life year; SoC, standard of care; UK, United Kingdom; type-2 diabetes mellitus (T2DM)

1.4 Other key issues: summary of the EAG's view

As reported in Table 22, SoC + empagliflozin results in a mean 4.51 increase in KCCQ-CSS score over 52 weeks and SoC alone offers a 3.18 increase. The mean change (95% CI) in KCCQ-CSS score for empagliflozin compared to placebo was 1.32 (0.45 to 2.19) and this is statistically significant. However, as detailed in Section 3.2.8.10, a 5-point change in the KCCQ score is commonly considered to be a clinically significant difference in health status in people with HF. Therefore, the comparative effectiveness did not approach what would be considered to be a clinically significant difference.

1.5 Summary of EAG's preferred assumptions and resulting ICER

The EAG's preferred assumptions are:

1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.
2. Assuming that empagliflozin does not have an effect on overall survival.
3. Using the age-related decrements from Ara and Brazier 2010.
4. Using the EB03E code to cost HHF in the model, rather than using a weighted mean cost derived from national FCEs for more severe HHFs events.
5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).

In addition to the assumptions listed above, the EAG conducted one set or alternative scenarios:

6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.
7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.

Finally, the EAG conducted the alternative combined scenarios described above with the following:

8. Assuming that empagliflozin does not have an effect on CV related deaths (directly or indirectly), or on overall survival.

Depending on the assumptions made, the EAG-preferred ICER falls within a wide range of [redacted]. Given how similar the deterministic base case results are to the probabilistic results, the EAG did not present probabilistic results of the exploratory analysis conducted.

The company's scenario analysis of assuming that the TPs between KCCQ-CSS quartiles for patients on treatment are the same as those for patients off treatment from month 9 onwards has a considerable impact on the EAG's cumulative ICER, ranging from [redacted], depending on the assumptions used.

Therefore, the EAG recommends that the committee validates the following assumptions:

1. The impact of empagliflozin on patients' survival (both overall and CV-related, as explained in Table 5 above).
2. The duration of the impact of empagliflozin on patients' KCCQ-CSS scores.
3. The duration of the impact of HHF events on patients' quality of life.

Table 8. EAG's preferred model assumptions

Preferred assumption	Incremental costs	Incremental QALYs	ICER (change from company base case
Company base case	[redacted]	[redacted]	£14,429
Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.	[redacted]	[redacted]	[redacted]
Assuming that empagliflozin does not have an effect on overall survival (directly or indirectly) but has an effect on CV mortality.	[redacted]	[redacted]	[redacted]
Using the age-related decrements from Ara and Brazier 2010.	[redacted]	[redacted]	[redacted]
Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.	[redacted]	[redacted]	[redacted]
Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).	[redacted]	[redacted]	[redacted]
Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.	[redacted]	[redacted]	[redacted]
Assuming HHF impact patients' QoL for 6 months year after hospitalisation.	[redacted]	[redacted]	[redacted]
Assuming that empagliflozin does not have an effect on overall survival (directly or indirectly) and does not have an effect on CV mortality.	[redacted]	[redacted]	[redacted]
EAG's assumptions combined when HHF impacts patients' QoL for 11 weeks after	[redacted]	[redacted]	[redacted]

hospitalisation and empagliflozin has an effect on CV mortality			
EAG's assumptions combined when HHF impacts patients' QoL for 6 months after hospitalisation and empagliflozin has an effect on CV mortality	[redacted]	[redacted]	[redacted]
EAG's assumptions combined when HHF impacts patients' QoL for 11 weeks after hospitalisation and empagliflozin does not have an effect on CV mortality	[redacted]	[redacted]	[redacted]
EAG's assumptions combined when HHF impacts patients' QoL for 6 months after hospitalisation and empagliflozin does not have an effect on CV mortality	[redacted]	[redacted]	[redacted]

Abbreviations: EAG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

2 Introduction and background

2.1 Introduction

This report contains an assessment of the company submission (CS) submitted for the Single Technology Appraisal (STA) of empagliflozin (Jardiance®, Boehringer Ingelheim) for treating chronic heart failure (HF) with preserved or mildly reduced left ventricular ejection (LVEF).

The European Society of Cardiology (ESC) defines chronic HF with reduced LVEF (HFrEF) as a LVEF of 40% or less; chronic HF with mildly reduced LVEF (HFmrEF) as a LVEF between 41% and 49%; and chronic HF with preserved LVEF (HFpEF) as a LVEF of 50% or more.[1] The NICE guideline Chronic heart failure in adults: diagnosis and management [NG106] defines HFrEF as a LVEF of 40% or less but does not specify a mildly reduced LVEF category.[2] This report will use the ESC definitions for reduced, mildly reduced, and preserved LVEF.

2.2 Background

Empagliflozin was previously recommended by the National Institute for Health and Care Excellence (NICE) for treating HFrEF [TA773].[3] It is to be used as an add-on to optimised standard care with:

- an angiotensin-converting enzyme (ACE) inhibitor or angiotensin 2 receptor blocker (ARB), with a beta blocker and, if tolerated, a mineralocorticoid receptor antagonist (MRA), or
- sacubitril valsartan with a beta blocker and, if tolerated, an MRA.

The company seeks to expand that recommendation for empagliflozin for people with HFmrEF or HFpEF.

Within Section B.1.3.1.1 of the company submission (CS), the company provides an overview of HF signs, symptoms, and underlying pathologies. Overall, based on advice from its clinical experts, the External Assessment Group (EAG) considers the CS to present an accurate overview. There were, however, a number of additional clinical factors the EAG's experts noted.

The EAG's clinical experts stated that at diagnosis when a person's LVEF can stay the same as at diagnosis but people commonly cross from any one category to any other during the course of their disease.

A measurement of a person's LVEF is normally done by using an echocardiogram. The EAG's clinical experts highlighted that there is significant intraobserver, inter observer and inter study variability in LVEF measurement. A person's LVEF could measure as much as 10% higher or lower when repeated

the next day and therefore, a person may move between categories with repeated scans or when scanned by different personnel. The EAG's clinical experts stated that this is a limitation for defining a person's HF using echocardiogram to measure their LVEF.

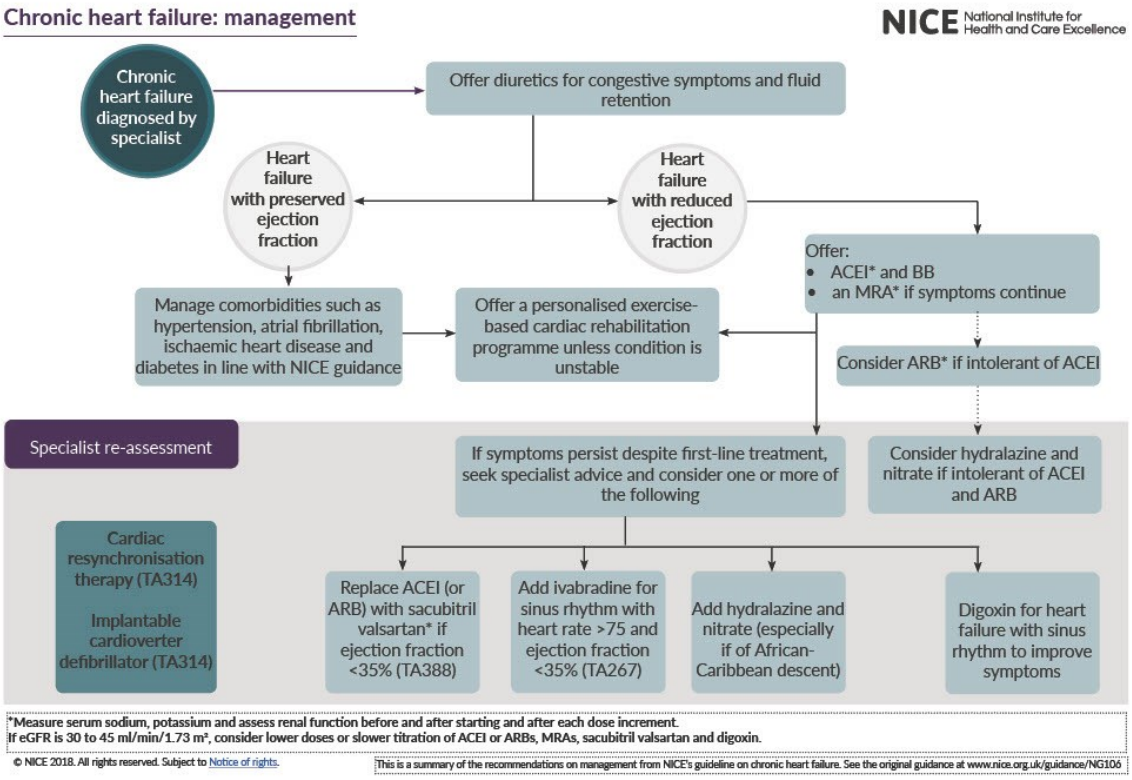
The clinical experts advised that people with HFrEF are a more homogeneous pathophysiological group with similar final common pathophysiological pathways that are treated using the same classes of drugs. This is less the case for people with HFpEF who are a heterogenous pathophysiological group. In between these two groups are people with HFmrEF where the underlying pathophysiology can be less clear. Many of these people align with those with HFrEF and respond to the set of treatments used in that group. Alternatively, they are sometimes more similar to people with HFpEF for whom there are limited HF treatments.

2.2.1 Treatment of people with heart failure and positioning of empagliflozin

Figure 1 below is taken from Chronic heart failure in adults: diagnosis and management [NG106] and presents the England and Wales treatment pathway for people with HF as it stood in 2018.[4] SGLT2 inhibitors, empagliflozin and dapagliflozin, were recommended as add-on treatments for people with HFrEF in 2021.

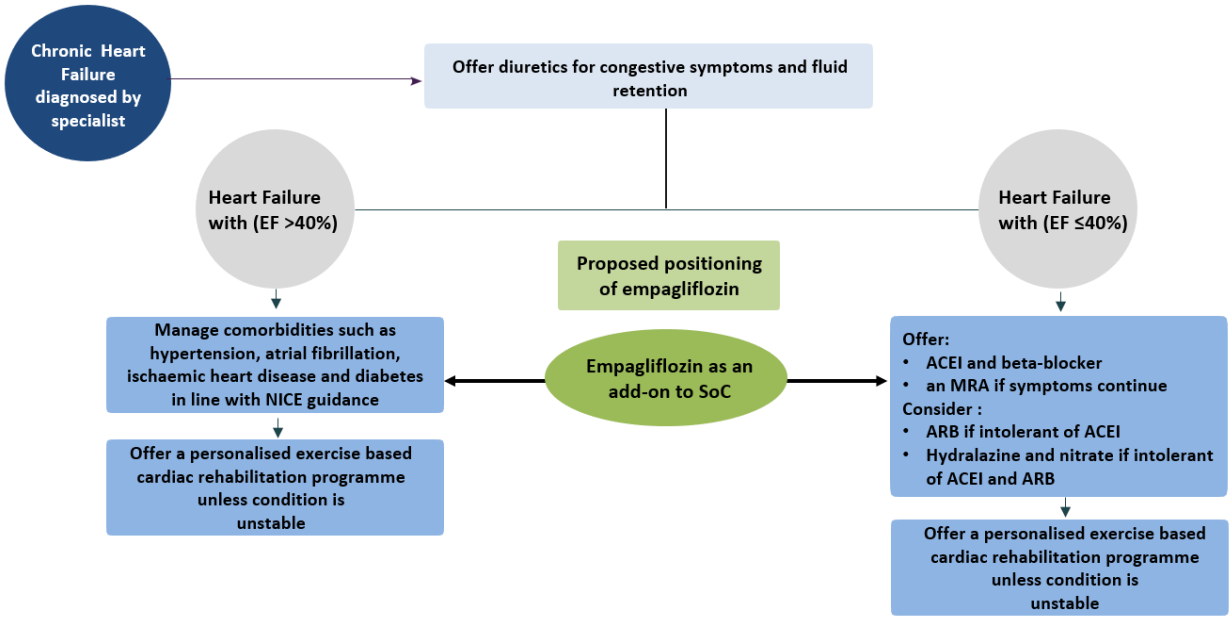
All people with HF are offered diuretics for congestive symptoms and fluid retention. At that point a person's LVEF determines the treatment pathway they take. There are several pharmacotherapy options for people with HFrEF and as mentioned above, this now includes SGLT2 inhibitors. People with HFmrEF/HFpEF have no specific HF treatment after they are offered diuretics. Management of these people involves treating comorbidity that cause and/or exacerbate heart failure including hypertension, AF, diabetes and obesity. This is because drugs which improved prognosis in HFrEF, such as angiotensin-converting enzyme (ACE) inhibitors, angiotensin II receptor blockers (ARBs), beta blockers, mineralocorticoid receptor antagonists (MRAs), and angiotensin receptor-neprilysin inhibitor (ARMI) have not been clearly shown to improve prognosis in people with HFmrEF/HFpEF.[5-8]

Figure 1. Treatment pathway in chronic heart failure management [NG106]



The company's proposed treatment pathway, including empagliflozin, for all people diagnosed with HF, is reproduced from the CS in Figure 2 below. It does not mention dapagliflozin which is also currently recommended as an add-on treatment for people with HFrEF.[9]

Figure 2. Proposed positioning of empagliflozin in NICE treatment pathway for chronic HF (reproduced from CS, Figure 4)



Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blockers; EF, ejection fraction; MRA, mineralocorticoid receptor antagonist.

In June 2022, the Medicines and Healthcare Products Regulatory Agency (MHRA) granted marketing authorisation for the use of empagliflozin as a treatment for adults with HFmrEF/HFpEF. This approval expanded the indication of empagliflozin to include all adults with HF. The treatment regimen is a fixed 10 mg oral dose once daily and the mode of action is detailed in Table 3 in Section B.1.2 in the CS.

The company has positioned empagliflozin as a single HF treatment, outside of diuretics, that can be offered to all people with HFmrEF/HFpEF. The EAG's clinical experts have highlighted that treatment options for this patient population are severely limited and, as such, there is an unmet need for an effective treatment to be made available.

Within the CS, the company states for people with HFrEF, there is an issue with people receiving empagliflozin based on perceived ambiguity in the recommendation for TA773. That is, "Start empagliflozin for treating symptomatic heart failure with reduced ejection fraction on the advice of a heart failure specialist." The company asserts that this has resulted in an additional referral to secondary care being required. The first obtains a diagnosis and the second for treatment initiation with empagliflozin. The company recommends that prescribing of empagliflozin could be better facilitated in primary care.

The EAG's clinical experts advised that a diagnosis and categorisation of HF should usually be made in secondary care, with treatment decisions made by the patient's multidisciplinary team (MDT). Prescribing of empagliflozin could take place in primary care, following a recommendation from the MDT.

2.3 Critique of the company's definition of the decision problem

The company provides a summary of the final scope[10] issued by NICE, together with their rationale for any deviation from the final scope. The decision problem is addressed in Table 2 in Section B.1.1 in the CS.

Table 9. Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope	EAG comment
Population	The table in the final scope states the population to be adults with symptomatic chronic heart failure with left ventricular ejection fraction of 40% or more.	Adults for the treatment of symptomatic chronic heart failure (EF >40%).	Not applicable	The company uses the EMPEROR-Preserved trial as the source of its clinical effectiveness data. Clinical experts advising the EAG, consider the population within EMPEROR-Preserved to be younger, but outside of age, broadly representative of clinical practice in the UK. See Sections 3.2.1 and 3.2.2 for further discussion of the eligibility criteria and baseline characteristics of the trial participants.
Intervention	Empagliflozin in combination with standard care (including loop diuretics and symptomatic treatments for co-morbidities).	Empagliflozin as an add-on to established clinical management. Empagliflozin does not replace established clinical management.	Not applicable	The intervention in the EMPEROR-Preserved trial was empagliflozin in combination with standard care. The clinical experts advising the EAG felt the standard care treatment were a reasonable representation of that in England and Wales. However, they were concerned that people in both groups were using [redacted] medications. Further discussion can be found in Section 3.2.2.
Comparator(s)	Established clinical management without empagliflozin, including but not limited to loop diuretics,	Established clinical management.	Not applicable	The comparator in the EMPEROR-Preserved trial was placebo in combination with standard care. See

	calcium-channel blockers, amiodarone, and anticoagulants.			the cell above for comments on the standard care treatment.
Outcomes	The outcome measures to be considered include: • symptoms of heart failure • hospitalisation for heart failure • all-cause hospitalisation • mortality • cardiovascular mortality • kidney function • adverse effects of treatment • health-related quality of life	"Same"	Not applicable	All of the outcomes in the NICE final scope were reported in the EMPEROR-Preserved trial. This was the primary source of data for the economic model. The EAG notes that the primary outcome from EMPEROR-Preserved is the composite of the combined risk of CV death or HHF. While the combined outcome isn't used in the economic model, the two individual outcomes (CV death and HHF) are used.
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.	"Same"	Not applicable	The clinical evidence was taken primarily from the EMPEROR-Preserved trial and the population matched that stated in the NICE final scope. Further discussion can be found in Sections 3.2.1 and 3.2.7.

	The availability of any commercial arrangements for the intervention, comparator, and subsequent treatment technologies will be taken into account.			
Subgroups to be considered	Not included in the draft scope	No subgroups were considered separately in the economic analysis		The company does not provide any cost-effectiveness results by subgroup in the CS. However, the company does provide the primary outcome for 15 subgroups (see Section B.2.7) and used this to justify a similar treatment effect across subgroups. See further discussion in Section 3.2.7.
Special considerations, including issues related to equity or equality	Not included in the draft scope	Broad prescribing of SGLT2i in primary and secondary care for HF, regardless of EF, could reduce the inequality in terms of access to heart failure care in the UK	The CS details the socio-economic equality considerations linked to CV disease. The CS also reflects on the recommendation, initiation, and prescribing of empagliflozin within the HF population. People in lower socio-economic classes utilise secondary care less often, their opportunity to access HF medications would also be lower, if they are solely prescribed in secondary care. BI support the UK Government's and NICE's commitment to the reduction of health inequalities, reiterated in the recent NICE five-year strategy publication.	The EAG considers matters relating to equality to be considerations for the committee. With regards to diagnosis and treatment of HF, please see Section 2.2.1.

			Principle 9 of NICE's Social Value Judgments states that due regard must be given to reducing inequalities. It states that equality should be considered in relation to the nine protected characteristics in the Equality Act 2010 (age, disability, gender reassignment, race, religion or belief, sex, sexual orientation, marriage and civil partnership, pregnancy and maternity) and socio-demographic factors. Further, the COVID-19 Marmot review aims to reduce the widened gap in health inequalities and build a fairer society post pandemic. Broad prescribing of SGLT2i across primary and secondary care can support the reduction in disparity in terms of access to HF care across socio-economic groups within the UK. Together with TA773, this appraisal further supports this objective by providing a treatment option for those patients regardless of EF	
Abbreviations:				

3 Clinical effectiveness

3.1 Critique of the methods review

The company presented the methods of the systematic literature review (SLR) in Appendix D of the company submission (CS), and the EAG's critique is presented in the table below. Appendix D states a SLR was conducted to identify randomised controlled trials (RCTs) or SLRs in adults with chronic HF (LVEF >40%).

Table 10. A summary of the EAG's critique of the systematic literature review

Systematic review step	Section of CS in which methods are reported	EAG's assessment of robustness of methods
Data sources	Appendix D.1.1	**The EAG considers the sources and dates searched to be comprehensive.** Databases searched: • Embase; MEDLINE; MEDLINE In-Process; the Cochrane Library (CENTRAL) The original searches carried out in 2020 included RCT filters for all databases except MEDLINE In-Process. Later searches, 8 July 2021 and 7 July 2022, did not utilise RCT filters and additionally searched the Cochrane Database of Systematic Reviews. Conference proceedings (2018 to 2021): • American Heart Association (AHA), European Society of Cardiology Congress (ESC), American College of Cardiology (ACC) Bibliographies of key systematic reviews and meta-analyses were screened to ensure that initial searches captured all the relevant studies.
Search strategies	Appendix D.1.1.3	**The EAG is satisfied that the company's searches have identified all evidence relevant to the decision problem.** The electronic database searches were performed from database inception to date of the SLR, without any time limit. The electronic databases were searched using a combination of MeSH/EMTREE terms, and free-text terms for both disease and intervention facets. All free-text terms were limited to abstracts, titles and keyword headings.
Inclusion criteria	Appendix D.1.1.4 (Table 15)	**The EAG considers it likely a relevant study was excluded.** The eligibility criteria matched the target population. Criteria for the intervention and comparator were wider than that specified in the NICE scope. The outcomes contained small differences to the NICE scope but the CS stated that their list was tentative rather than exhaustive. Records were limited to English language studies. A reference list of all records excluded at full text review was provided. However, 17 'included' RCTs were effectively excluded without reasoning provided. One of these studies, EMPERIAL-Preserved, appeared to meet the

		SLR's inclusion criteria but was not presented. This is discussed further in Section 3.1.1.
Screening	Appendix D.1.1.4	**The EAG considers the reporting of methods for screening to be adequate.** Title/abstract screening and full text screening was conducted by two independent reviewers. Any discrepancy was resolved by a third, independent reviewer.
Data extraction	Appendix D.1.1.5	One reviewer extracted the data from the included full-text articles. All extracted data were quality checked against the original source article by the second reviewer.
Tool for quality assessment of included study or studies	Appendix D.1.2.3 and D.5 (Table 13)	**The EAG agrees with the company's choice of quality assessment tool of RCTs.** Study quality was assessed using recommendations given in the NICE manufacturer's submission template.

Abbreviations: CENTRAL, Cochrane Controlled Register of Trials; EAG: External Assessment Group

3.1.1 Inclusion criteria

From the SLR, 18 RCTs were stated to be included and this contained the EMPEROR-Preserved trial, which was the primary source of data for the economic model. The methodology and results of the EMPEROR-Preserved trial were presented in the CS but summaries and results from the remaining 17 RCTs 'included' in the SLR were not presented. Therefore, 17 RCTs were effectively excluded from the SLR but no explicit reasoning was given. It is good practice within SLRs to provide reasons for exclusion when the full text of a paper has been assessed and a lack of transparency in the process leads to a risk that relevant studies have been excluded.

Of the 17 RCTs effectively excluded from the review, 16 did not utilise empagliflozin as part of the intervention and so it is unlikely they addressed the NICE decision problem. However, EMPERIAL-Preserved[11] met the SLR's PICO but was effectively excluded. Section B.2.1.1, separate to, the excluded studies list, states the primary outcome of EMPERIAL-Preserved was not relevant for the decision problem and the quality of life secondary endpoint measured using PROs is not recommended by the NICE reference case. It was a 3-month trial and included Kansas City Cardiomyopathy Questionnaire (KCCQ) outcomes and safety outcomes, both of which were relevant for this STA. The company should have included this study and presented the KCCQ and safety data, either for use alongside the EMPEROR-Preserved data in the economic model or for use in validating 3-month KCCQ and safety data from the EMPEROR-Preserved trial.

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation

In this section, the EAG critiques the EMPEROR-Preserved trial as the primary source of data for the economic model. The methods and baseline characteristics of participants are presented in Section B.2.3 of the CS, plan for analysis and CONSORT flow diagram in B.2.4, critical appraisal of the trial in B.2.5, clinical effectiveness results in B.2.6, and subgroup analysis in B.2.7. As noted in Section 3.1.1, the EMPERIAL-Preserved trial was relevant but effectively excluded from the SLR.

The EAG's critique of the design, conduct and internal validity of EMPEROR-Preserved is presented in Table 11 below.

Table 11. A summary of the EAG's critique of the design, conduct and analysis of EMPEROR-Preserved

Aspect of trial design or conduct	Section of CS in which information is reported	EAG's critique
Randomisation	B.2.3, CSR	**Appropriate** Randomisation was performed by using a permuted block design with a computer pseudo-random number generator. Randomisation was stratified by EF (<50%, ≥50%) and >66% of enrolled patients had EF ≥50%.
Concealment of treatment allocation	NA	**Appropriate** An Interactive Response Technology System was used to determine treatment assignment.
Eligibility criteria	B.2.3	**Appropriate** The inclusion criteria for EMPEROR-Preserved, adults for the treatment of symptomatic chronic heart failure (EF >40%), matched the population stated in the NICE final scope. The EAG's clinical experts indicated the trial's inclusion criteria linked to a minimum NT-proBNP and maximum BMI criteria were not a cause for concern. This is expanded upon in Section 3.2.1.
Blinding	B.2.3	**Appropriate** EMPEROR-Preserved was a double-blind study. The participants, treating physicians, and independent external clinical event committees were blinded.
Baseline characteristics	B.2.3.2 Table 16	**Appropriate** The baseline characteristics were similar between treatment groups. The EAG's clinical experts considered the characteristics to be broadly appropriate for the population they see in clinical practice. They did comment that the age of the people in the trial was lower than that seen in their clinical practice and there may be a lower proportion of people with HF and comorbid with type 2 diabetes and/or hypertension. Within the trial were roughly an equal number of people with an EF from 40% to <50%, 50% to <60%, and ≥60%. The EAG's clinical experts did not indicate this was an unreasonable reflection of what might be found in

		clinical practice. At clarification, the company stated that no capping was used to influence the LVEF of people recruited to the trial.
Dropouts	B.2.4.1 Table 17	**Some concerns** The median length of treatment for participants in EMPRERO-Preserved was 26 months. Of the people in the empagliflozin group, 23.2%, and 23.4% in the placebo group discontinued medication for reasons other than death. Discontinuations due to adverse events were 11% in the empagliflozin group and 10% in the placebo group. The EAG asked for clarification on why these numbers were higher in EMPEROR-Preserved than EMPRERO-Reduced given the similarity of populations and treatments. The company's clarification response (CR) stated the differences between the trials was due to the differences between the recruited population and the longer follow-up in the EMPEROR-Preserved trial. This is further discussed in Section 3.2.3.
Statistical analysis		
Sample size and power	B.2.4.1.3 Table 17	**No concerns** The CS stated that there was a target of 841 primary outcomes to achieve a power of 90% for a two-sided test with α=0.05 and detect a hazard ratio of 0.80. By the end of the trial, there were 926 primary outcomes.
Handling of missing data	B.2.4.1.3 Table 17	**Some concerns** The study handled missing data in several primary analyses using either a multiple imputations framework or multivariate Cox regression models. Both methods impute data that is matched the data that was collected assuming that data are missing at random. The EAG has concerns these two methods for managing missing data do not take into account participants who withdrew from the study due to ineffectiveness of the intervention. Last observation carried forward (LOCF) was used to handle missing data for KCCQ-CSS scores. Again, this method imputes data that matches the data that was collected assuming that data are missing at random. In several cases including the subgroup analysis and adverse events, no imputation was performed. The concerns are further discussed in Section 3.2.3.
Outcome assessment	B.2.3.1.5, B.2.3.1.7, B.2.4.1, B.3.4.1, CSR 9.5.1	**Some concerns** The outcome assessors were blinded to the treatment group of the participants. The adjudicated primary endpoints were comprehensively defined in Table 15 of the CS. EQ-5D-5L was assessed at baseline and weeks 12, 32, 52, 100, and 148. KCCQ, the efficacy outcome utilised in the economic model, was assessed at 52 weeks but was not assessed again until the two EOT visits. This is of concern to the EAG as due to this being an event-driven trial, the EOT outcome could be prior to 52 weeks in some participants and more than three years in others. This is further detailed in Section 3.2.5.
Analysis for estimate of effect	B.2.6.2.9, B.2.4.2,	**Appropriate** Most of the study's reported outcomes were analysed using the randomised set (RS) and a number used the treated set (TS). Both groups appear very closely matched in terms of their composition. The primary report of deterioration of renal function was in people who were "on-treatment". This is per-protocol analysis where the comparison of treatment groups includes only those people who completed the treatment originally allocated when the outcome was assessed. Per protocol analysis

		is not thought to represent the real-life situation and is likely to show an exaggerated treatment effect.[12] However a secondary analysis reports the outcome in the randomised set.

Abbreviations: ALT: alanine aminotransferase; CHD: chronic hepatitis delta; CSR: clinical study report; EAG: evidence review group; HDV: hepatitis delta virus; HRQoL: health-related quality of life; ITT: intention-to-treat; NICE: National Institute of Health and Care Excellence; PCR: polymerase chain reaction

3.2.1 Eligibility criteria

The people included in the trial met the decision problem stated by NICE. They were adults with HFmrEF and HFpEF. In Section 9.3.1 of the CSR, the full inclusion criteria are detailed. A notable criterion was participants having had no prior measurement of LVEF ≤40% under stable conditions. EF measurements can vary by as much as 10% when repeated the next day and this inclusion criterion limited the number of people who were HFrEF participating in the trial. Empagliflozin was shown to be effective in the HFrEF (LVEF ≤40%) population in the EMPRERO-Reduced trial[13], and including a proportion of these people in the trial would have biased the result in favour of empagliflozin.

Participants were required to have a NT-proBNP of more than 300 pg/mL for people without atrial fibrillation (AF), or more than 900 pg/mL for people with AF. NT-proBNP is most often used to diagnose or rule out heart failure and indicates how much stress the heart is under. The NICE guideline, Chronic heart failure in adults: diagnosis and management (NG106),[2] recommends people with NT-proBNP of 400 pg/mL or higher are referred for specialist assessment. Thus several people with a slightly lower level would have been recruited for the trial. The clinical experts advising the EAG assessed this to be an "inclusive" criterion and unlikely to rule out a relevant section of the HF population.

People were excluded if they had a BMI of 45 kg/m^2 or greater at baseline. No reasoning for this criterion was provided in the CS. The EAG's clinical experts speculated that this is likely to be a more comorbid group and there could be a concern that the positive effects of empagliflozin treatment would be harder to discern if this group had been included. However, they did not have any specific concerns that this population would react differently to empagliflozin treatment than people with a BMI under 45 kg/m^2.

3.2.2 Baseline characteristics

The population recruited in the trial were adults with HFmrEF and HFpEF. The EAG's clinical experts commented that the trial participants are representative of people with HF seen in England and Wales.

The people recruited to the trial had a mean age of 72 years old and the EAG's clinical experts indicated this is younger than the people in their practice who are more often in their 80s. The experts understood it was difficult to recruit older people into clinical trials and the subgroup analysis provided by the company appeared to indicate recruiting younger people did not necessarily favour empagliflozin.

In the trial, 81% of people had a New York Heart Association (NYHA) functional class II and 18% class III. The EAG's clinical experts considered this to be a reasonable reflection of the relevant England and Wales HF population.

The family background of people in the study was 76% White, 4% Black, 14% Asian, 6% other. This is broadly representative, although in England and Wales, the proportion of Asian people is lower and White people is higher. The EAG's clinical experts were not aware of any evidence or reasoning why family background would impact the effectiveness of empagliflozin.

The time since diagnosis for the trial participants was slightly higher than the EAG's clinical experts see in practice. This is because they more commonly see people at diagnosis and people are discharged back to primary care for treatment initiation and prescribing. The number with HF for over 10 years were few and this was as expected as people diagnosed with HFmrEF/HFpEF tend to be over 75 years old, with comorbidities, and many will not survive for more than 10 years. The EAG's clinical experts felt that the participants recruited for the trial were a better reflection of HF care in England and Wales than limiting recruitment to people who were recently diagnosed.

The population was comorbid and details of the CV history and diabetes status is presented in Table 12 below. Overall, the EAG's clinical experts considered these proportions to be a reasonable reflection of people with HFmrEF and HFpEF in England and Wales. One expert stated that the proportion of people of people with HFmrEF/HFpEF and hypertension or type 2 diabetes (T2D) would be slightly lower in the England and Wales. Subgroup analysis presented in Figure 3 in Section 3.2.7 found little difference in the primary outcome for people living with or without T2D and the

EAG's clinical experts were not aware of any reason why small variations in comorbidity would influence the estimate of the effectiveness of empagliflozin.

The EAG's clinical experts also noted that it is common practice in HF trials to make hospitalisation for HF in the past 12 an inclusion criterion. This is a method used to "enrich" the trial population with people who have been hospitalised before and who are therefore more likely to be hospitalised again. This was not done in the EMPEROR-Preserved trial.

Table 12. Comorbidities at baseline (adapted from Table 16 in the CS)

CV history, N (%)		
	Empagliflozin (N=2,997)	Placebo (N=2,991)
Hospitalisation for HF in ≤12 months	699 (23.3)	670 (22.4)
Atrial fibrillation	1,543 (51.5)	1,514 (50.6)
Hypertension	2,721 (90.8)	2,703 (90.4)
Diabetes status		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
With diabetes	1,466 (48.9)	1,472 (49.2)
Abbreviations:		

The baseline medications participant's received were detailed in Table 16 in Section B.2.3.2 of the CS. They have been reproduced in Table 13 below. The clinical experts advising the EAG considered that the baseline medications reflected the people recruited into the study. A slightly higher proportion of people with hypertension participated in the trial than would be expected in England and Wales's clinical practice. They are treated with ACEI and beta-blockers, and consequently the number of people using ACEI and beta-blockers might be slightly lower in clinical practice. Also, [redacted] of people in each group were treated with [redacted]. [redacted] does not have marketing authorisation for use in people with HFmrEF/HFpEF, and people who had previously been measured as HFrEF were excluded from the study. The EAGs clinical experts explained that this could be due to [redacted] having marketing authorisation in people with HFmrEF in the USA.

Table 13. Baseline medications (reproduced from CS, Table 16)

HF medication, N (%)		
	Empagliflozin (N=2,997)	Placebo (N=2,991)
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

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

ACEI: Angiotensin-converting enzyme inhibitors, ARB: angiotensin II receptor blockers, ARNI: angiotensin receptor/neprilysin inhibitor

PULSE was a large retrospective observational study of the burden of chronic HF, in England, and was used again to validate the outcomes predicted from the economic models. It was based on existing data from the UK Clinical Practice Research Datalink (CPRD) linked to hospital episode statistics (HES) and office for national statistics (ONS) mortality data. Table 14 below details the 'background treatments' for people in EMPEROR-Preserved and PULSE. The EAG noted much higher proportions of people were using each medication class in EMPEROR-Preserved than in the PULSE (LVEF>40%) group and requested an explanation at clarification. The company stated that background treatment in PULSE used CPRD and this captures primary care prescribing but will underestimate prescribing if it happens outside of primary care. The company also stated that the largest group in PULSE had "unknown" LVEF which will have included people with HFmrEF/HFpEF. Correct categorisation of this group would have provided a more accurate treatment profile of the HFmrEF/HFpEF population.

Table 14. Background treatments for HF in EMPEROR-Preserved vs PULSE (a CPRD study) (reproduced from CS, Table 9)

Treatment Arm	EMPEROR-Preserved (Combined Groups)	PULSE (unknown group)	PULSE (LVEF>40%)
N	9718 (100.0)	[redacted]	[redacted]
HF medication [(N), %]			
ACEi/ARB	7305 (75.2)	[redacted]	[redacted]
Beta-blocker	8700 (89.5)	[redacted]	[redacted]
Diuretic	8708 (89.6)	[redacted]	[redacted]
MRA	4905 (50.5)	[redacted]	[redacted]
Sacubitril/valsartan	861 (8.9)	[redacted]	[redacted]
Ivabradine	331 (3.4)	[redacted]	[redacted]
Digoxin	NR	[redacted]	[redacted]
Hydralazine/nitrate	282 (2.9)	[redacted]	[redacted]

The EAG asked the company for clarification on whether there was a recruitment strategy linked to participant's LVEF outside of recruiting people with LVEF ≥40%.

The EMPEROR-Preserved trial protocol states the aim to recruit a trial population consisting of approximately 35% to 50% with an LVEF >50% with a mechanism to cap enrolment based on a recommendation from the executive steering committee. The mean LVEF in the trial was 54% and participants were recruited to the trial with a slightly higher EF than was originally aimed. The company stated in their response at the clarification stage that the Executive Committee advised to keep recruiting without capping participants with higher LVEF. However, prior to unblinding, the Trial Statistical Analysis Plan was updated to include subgroup analysis in three LVEF groups: <50%, ≥50% to <60%, and ≥60%.

A breakdown of the participants into the three LVEF subgroups is presented in Table 15 below. The clinical experts advising the EAG stated the numbers in each LVEF category in the trial were not an unreasonable reflection of what could be found in clinical practice.

Table 15. Breakdown of participants by baseline left ventricular ejection fraction

Baseline characteristics	Empagliflozin 10 mg	Placebo
Left ventricular ejection fraction (LVEF)		
Value of <50%, N (%)	995 (33.2)	988 (33.0)
Value of 50% to <60, N (%)	1,028 (34.3)	1,030 (34.4)
Value of ≥60%, N (%)	974 (32.5)	973 (32.5)

3.2.3 Dropouts

EMPEROR-Preserved was an event-driven trial and people who were recruited earlier in the trial were treated and followed up for longer than those who were recruited later. The length of time participants were observed during the study was detailed in the clinical study report and has been reproduced in Table 16 below. [redacted] of participants were treated for at least one year and [redacted] were treated for at least 2 years. The median length of observation for trial participants was 26.2 months.

Table 16. Observational period up to the end of planned treatment period – randomised set (reproduced from CSR, table 10.5: 1)

	Placebo	Empagliflozin 10 mg	Total
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

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

Discontinuations for reasons other than death were 23.2% in the empagliflozin group and 23.4 % in the placebo group. Discontinuations due to adverse events were 11% in the empagliflozin group and 10% in the placebo group. Dropouts, not due to fatal events, in the EMPEROR-Reduced trial which assessed empagliflozin in the HFrEF population, were 16% in the empagliflozin group and 18% in the placebo group. The EAG requested clarification on the higher discontinuation rate in EMPORER-Preserved.

The company's clarification response stated there were a higher proportion of dropouts in EMPORER-Preserved as compared to EMPEROR-Reduced due to the differences between the recruited populations and the longer follow-up in the EMPEROR-Preserved trial. The EMPORER-Preserved participants were older than those in EMPEROR-Reduced, 71 years old versus 67 years old, respectively. Additionally, they were more comorbid with 50% to 36% living with AF and 90% to 72% with hypertension. The company state that these conditions are both associated with a higher risk of adverse events and consequently discontinuation of treatment. The company also notes that the median study treatment period was longer in the EMPEROR-Preserved (26.2 months) versus EMPEROR-Reduced (16 months) trial. Having a longer treatment period increases the timeframe for discontinuations to occur and, hence, results in an increased rate of discontinuations in the EMPEROR-Preserved trial as compared with EMPEROR-Reduced.

The EAG also requested clarification on the number of people in three LVEF subgroups who discontinued medication and Table 17 below was included in the company's response. Similar proportions discontinued in each LVEF subgroup.

Table 17. Medication discontinuations for reasons other than death in HF EF subgroups (Randomised set)

HF LVEF subgroup	Discontinuations / Sample size (%)
40% to < 50%	[redacted]
50% to < 60%	[redacted]
≥ 60%	[redacted]

3.2.4 Handling of missing data

Table 17 in Section B.2.4.1 of the CS sets out the framework for data management and patient withdrawals used in EMPEROR-Preserved. Multiple methods were used for primary analysis and

sensitivity analyses. This includes the Multiple imputations framework, multivariate Cox regression models, and last observation carried forward (LOCF). All three techniques share the same assumption that data are missing at random and impute data points that reflect the data that was collected. The EAG is concerned the assumption does not account for participants who withdrew from the study due to the ineffectiveness of the intervention as the data points imputed will mirror the effectiveness seen in the participants who did not discontinue treatment.

For the primary outcome, using the multiple imputations framework, imputations were performed for 172 patients with incomplete data (84 empagliflozin and 88 placebo). This equates to 3% in each treatment group. The EAG is reassured by a similar dropout rate in both groups in the trial. That said, any remaining bias introduced by this method of accounting for missing data is likely to favour the more effective treatment. However, given the small proportion of patients affected, the EAG does not consider this likely to have a major impact on outcomes reported from the trial.

The EAG requested data with and without imputation for KCCQ-CSS at each time point. This is further discussed in Section 3.2.8.10.

The CS also indicates there was no imputation of data for safety analyses but it does specify those who discontinued treatment due to adverse events.

3.2.5 Outcome assessment

EMPEROR-Preserved was an event-driven RCT, which was powered to detect a hazard ratio of at least 0.8. A power calculation was made based on how many people to recruit, how likely the event of interest was, how many events were required, and the period people would have to be followed to achieve this. This led to the same end-of-treatment date for all participants irrespective of when they were recruited. Thus, people who were recruited early were treated for over three years and people recruited towards the end of the trial received treatment for less than one year. The median treatment period of people in the trial was 26 months.

The outcome assessors were blinded to the treatment group of the participants and the adjudicated primary endpoints were comprehensively defined in Table 15 of the CS. Table 18 below has been adapted from the trial flow chart in the EMPEROR-Preserved CSR. People had 16 scheduled visits, 10 in person and 6 by telephone, until week 148. Visits then continued every 12 weeks until the EOT visit and EOT+30 days visit.

Adverse events were assessed at [redacted] visits. Trial endpoints, including CV mortality and HHF, at [redacted] from visit [redacted]. Quality of life via EQ-5D-5L was assessed [redacted] times in the first year and then at [redacted], and at the [redacted]. KCCQ was assessed [redacted] in the first year, including at [redacted], but was not assessed again until the [redacted].

The EAG expected KCCQ to be measured at later points similar to that done for the EQ-5D outcome. Measures at later time points could have provided further evidence of the longevity of empagliflozin's effectiveness. KCCQ measured at the [redacted] visit was not a consistent time point for participants and the length of treatment at that point depended on when the person entered the trial.

Table 18. Outcomes trial flow chart (adapted from CSR, Table 9.5.1)

Trial period	Screening	Randomised Treatment Period														Follow-Up Period	
Visit	1	2	3	4	5 PC	6	7 PC	8	9 PC	10	11 PC	12	13 PC	14	15 PC	16	EOT	FU
Trial week	-3	1	4	12	22	32	42	52	64	76	88	100	112	124	136	148	EOT	+30 days
Assmt of endpoints			▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮
Adverse events	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮	▮
KCCQ		▮		▮		▮		▮									▮	▮
EQ-5D		▮		▮		▮		▮				▮				▮	▮	▮

From Visit 8 and onwards, on-site visits were to be scheduled every 24 weeks until end of trial. Patients who prematurely discontinued trial medication performed the EOT visit and the follow-up visit, and then continued with scheduled visits until the trial was stopped. For patients not willing to attend scheduled visits, telephone calls were to be made regularly to document any occurrence of outcome events and vital status. After 148 weeks, visits were to be repeated with same intervals as from Week 64 onwards.

PC: phone call

Assmt: assessment

3.2.6 *Analysis for estimate of effect*

Most of the EMPEROR-Preserved reported outcomes were analysed using the randomised set (RS). Several outcomes were measured in the treated set (TS) such as renal outcomes, including eGFR, and KCCQ. This was all people who were dispensed study medication and were documented to have taken at least one dose of the investigational treatment. It was unclear whether EQ-5D-5L was measured in this group. Figure 6 in the CS states that one person randomised to the intervention group did not start treatment and two people in the placebo group did not start treatment so it would appear the TS and RS are closely matched.

Deterioration of renal function was reported in people who are "on-treatment". This is per-protocol analysis where the comparison of treatment groups includes only those people who completed the treatment originally allocated when the outcome was assessed. This analysis is not thought to represent a real-life situation and it is likely to show an exaggerated treatment effect.[12] However a secondary analysis of the deterioration of renal function was also reported using the randomised set.

In most cases, the outcomes were reported from baseline and during periods of the study until a final measure towards the end of the study. Outcomes linked to mortality (CV or all-cause), HHF, and the primary composite outcome were reported every 3 months from baseline until 36 months. The efficacy outcome utilised in the economic model was the KCCQ-CSS outcome and it is reported as 52 weeks in Table 22 in Section B.2.6.2.9 of the CS. It was also measured at the end of treatment (EOT) visit and this was reported by the company at the clarification stage.

3.2.7 *Subgroups*

Figure 17 in the CS reports on 15 pre-specified subgroup analyses using the primary outcome, risk of CV death or HHF. This has been reproduced below in Figure 3. No imputation was done for covariates included in treatment by subgroup interaction terms. The EAG notes that carrying out so many bivariate sensitivity analyses can produce spurious results and the subgroup analyses should be interpreted with caution.

Among the conclusions drawn by the company from these subgroup analyses was that empagliflozin was similarly effective in people who are diabetic versus people who are non-diabetic, people with baseline eGFR ≥60 mL/min/1.73m^2 versus people with eGFR <60 mL/min/1.73m^2, and people with

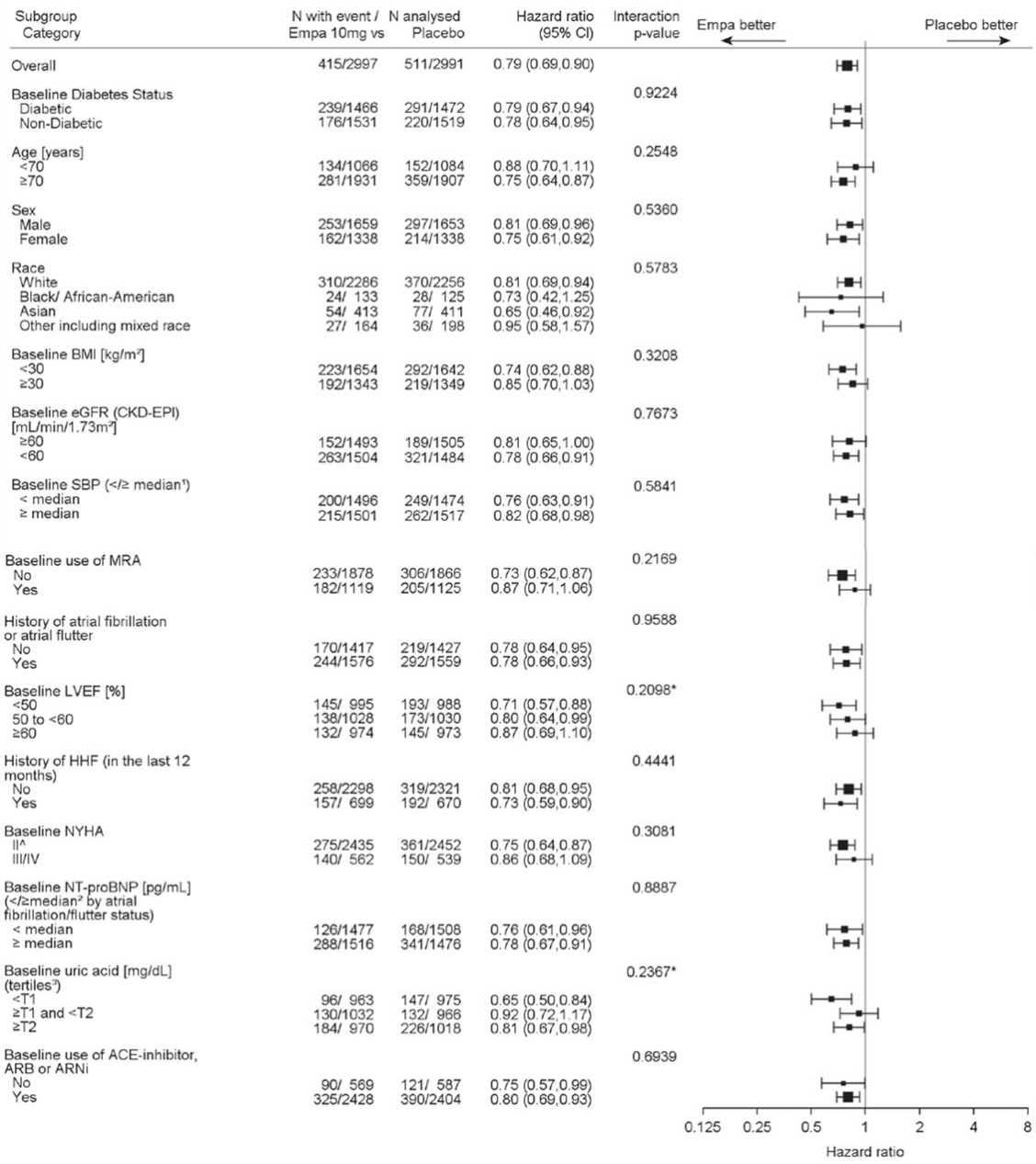
above the median NT-proBNP versus people below the median NT-proBNP in the trial. The EAG noted subgroup analyses that indicated variations in effectiveness between subgroups. Empagliflozin was numerically more effective in people who were 70 years or older than those under 70 years old. The EAG's clinical experts advised that the people they see in clinical practice tended to be closer to 80 years old and were not concerned about the numerical reduction in effectiveness in the younger population. Empagliflozin also appeared to be numerically more effective in people with a BMI under 30 kg/m^2 than those with a BMI of 30 kg/m^2 or over at baseline. As stated in Section 3.1.1, the EAG's clinical experts speculated that people living with obesity are a comorbid group and the positive effects of empagliflozin treatment may be harder to discern in a comorbid population.

Subgroup analysis was carried out for people in three LVEF categories. Empagliflozin was numerically more effective in people with a baseline LVEF >40% to <50% than in people with an LVEF from 50% to <60%. Likewise, it was numerically more effective in people with an LVEF from 50% to <60% than those with an LVEF ≥60%. In people with HF and an LVEF ≥60%, the estimated hazard ratio of 0.87 with confidence intervals that span the line of no effect could be interpreted as not showing a clinically meaningful benefit. A benefit of this analysis was that there was an apparent linear reduction in effect across three subgroups but the study was not powered to detect a significant difference between treatments in smaller subgroups. The clinical experts advising the EAG reflected on this variation in effectiveness across subgroups defined by LVEF. They assessed the numerical reduction in effectiveness in people with HFpEF may be a function of the heterogenous pathology of the population. They concluded that empagliflozin is effective in this subgroup. A clinical expert stated that most departments do not quote LVEF when it is above 55% as that is considered normal.

Subgroup analysis in three LVEF categories was provided by the company in their response to clarification questions. It is further discussed in the relevant results section below:

- Kansas City Cardiomyopathy Questionnaire (KCCQ) score;
- Total HHF (first and recurrent);
- Cardiovascular mortality;
- All-cause mortality.

Figure 3. Primary outcome of EMPEROR-Preserved in pre-specified subgroups (reproduced from CS, Figure 17)



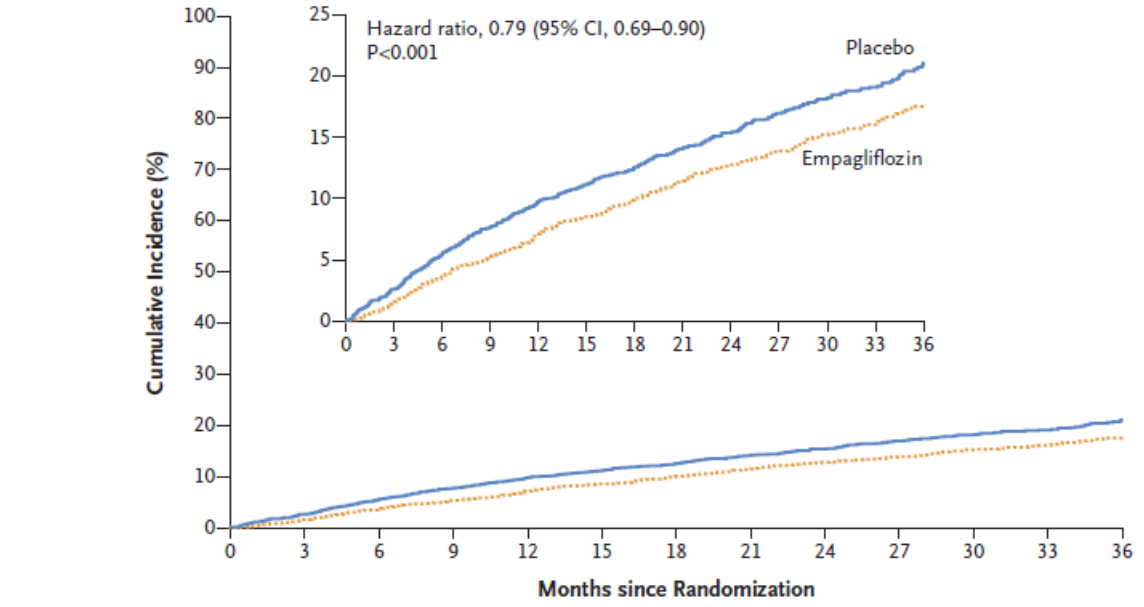
3.2.8 *Clinical effectiveness results from EMPEROR-Preserved*

Summaries of the primary and secondary outcomes with the EAG comment are presented in this section.

3.2.8.1 *Primary outcome: combined risk of CV death or HHF*

The results demonstrated significantly that fewer people in the empagliflozin group (415/2997 people, 13.8%). experienced an event compared to in the placebo group (511/2991 people, 17.1%). The company presented Figure 4, the cumulative incidence plot of CV death or first HHF, considering non-CV death as a competing risk. The EAG notes from the plots for the individual outcomes of CV death and HHF that the benefit for empagliflozin appears to be predominantly driven by HHF events.

Figure 4. Mean cumulative incidence plots of CV death or first HHF (Reproduced from CS, Figure 7)



The company conducted a Cox regression of the data for all randomised people adjusted for age, baseline eGFR (CKD-EPI)cr, region, sex, treatment, and baseline diabetes status which showed a clinically important reduction in the risk of CV death or HHF with empagliflozin compared to placebo (HR 0.79; 95% CI: 0.69 to 0.90, p<0.001). In addition, the company conducted three sensitivity analyses of the primary endpoint:

- Multiple imputation analysis addressing incomplete data for primary endpoints in 172 participants;
- Results unadjusted for covariates;
- Sub-distribution hazard ratio adjusted for non-CV death as a competing risk in RS (Fine-Gray model).

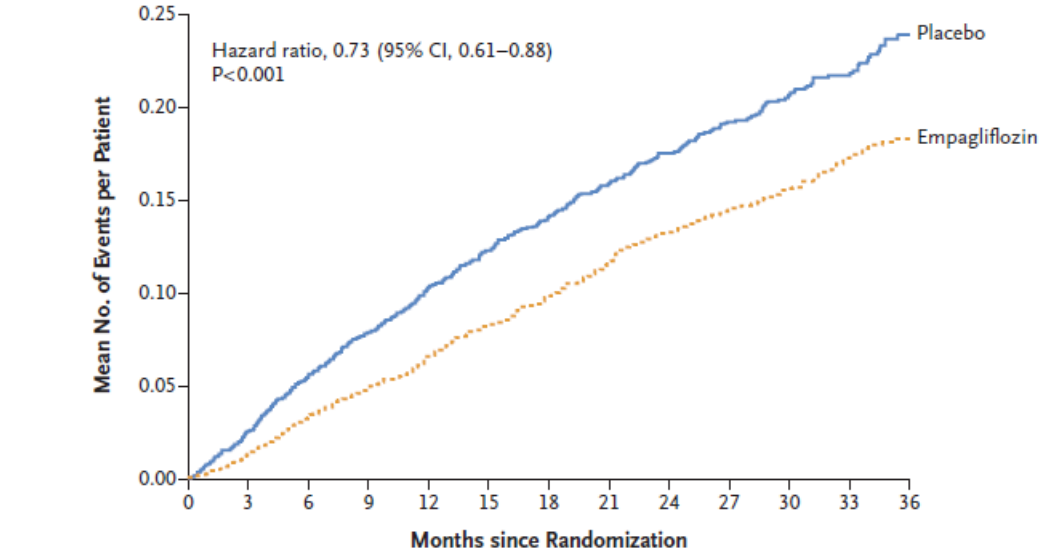
The EAG notes that the results from the sensitivity analyses were all consistent with the results of the primary analysis (CS, Table 19).

This combined outcome is not utilised in the economic model but its two constituent parts are.

3.2.8.2 Total HHF (first and recurrent)

The total number of HHF event was significantly lower in the empagliflozin group than in the placebo group with 407 events and 541 events, respectively. Figure 5 shows the mean cumulative incidence plot of total HHF over time.

Figure 5. Mean cumulative incidence plots of total HHF (first and recurrent) (Reproduced from CS, Figure 8)



The company conducted an analysis using a joint frailty model with CV death as a competing risk which showed a clinically important reduction in the risk of HHF (first and recurrent) with empagliflozin compared to placebo (HR 0.73; 95% CI: 0.61 to 0.88, p<0.001). In addition, the company conducted five sensitivity analyses:

- Parametric joint gamma frailty model considering CV death as a competing risk;
- Joint frailty model considering all-cause mortality as a competing risk;
- Negative binomial model;
- Negative binomial model without covariate adjustment;

- Cox regression for time to first adjudicated HHF.

The EAG notes that the results from the sensitivity analyses were all consistent with the results of the primary analysis (CS, Table 20). First and recurrent HHF was an input for the economic model.

The EAG requested total HHF (first and recurrent) data in three LVEF subgroups at the clarification stage. The company's response is presented in Table 19 below. A [redacted] for empagliflozin was found for the LVEF 40% to <50% and LVEF 50% to <60% subgroups. There was [redacted] between empagliflozin and placebo in the LVEF ≥60% group. This should be interpreted with caution as the study was not powered to detect a significant difference between treatments in smaller subgroups.

Table 19. Outcome data for total HHF (first and recurrent) by ejection fraction (Randomised set)

Endpoint	Empagliflozin 10 mg	Endpoint	Empagliflozin 10 mg
Occurrence of adjudicated HHF (first and recurrent)			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
Abbreviations:			

3.2.8.3 Deterioration of renal function

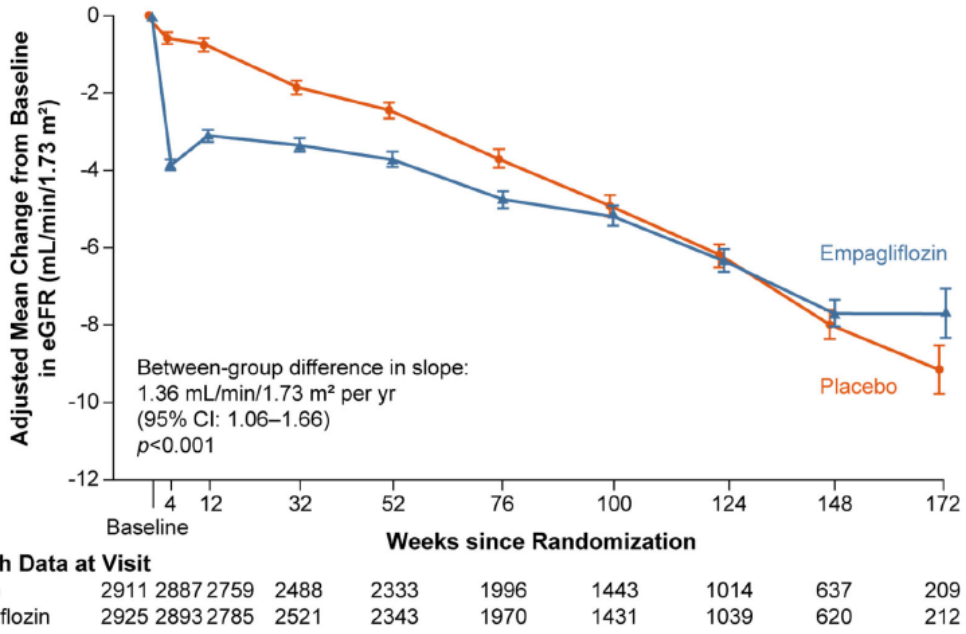
The primary analysis was per-protocol analysis where the comparison of treatment groups includes only those people who completed the treatment originally allocated when the outcome was assessed. The mean slope of change in eGFR (mL/min/1.73 m^2) from baseline was slope was -1.25 ± 0.11 mL/min/1.73 m^2 per year in the empagliflozin group and -2.62 ± 0.11 mL/min/1.73 m^2 per year in the placebo group. The estimated between-group difference in mean slope was 1.36 ± 0.30 mL/min/1.73 m^2 per year. Figure 6 shows an initial drop in eGFR seen at the start of the treatment with empagliflozin and the company states this is a reversible functional change in intrarenal haemodynamics commonly observed with SGLT2 inhibitors and is not associated with an excess risk of acute kidney injury. After the initial drop, the reduction of eGFR in the empagliflozin group is slower than that in the placebo group. After 124 weeks, the groups reached parity in eGFR, and from 148 weeks a benefit can be seen for empagliflozin. The clinical experts advising the EAG agreed with the company's assessment of the long-term benefit of empagliflozin for change in eGFR. They indicated longer treatment data would better support the conclusion but also were aware that dapagliflozin, an SGLT2 inhibitor, has been recommended as an option for treating chronic kidney

disease (CKD). The EMPA-KIDNEY trial for empagliflozin in people with CKD has been stopped early. A company press release indicates this was due to "clear positive efficacy."[1]

A further analysis was undertaken in the RS. The adjusted mean eGFR change from baseline to follow-up was 2.4 (1.6, 3.2) mL/min/1.73m^2 per year for empagliflozin versus placebo. This would appear to be a larger effect that that seen using the "on-treatment" data.

This outcome was not used in the economic model. The renal input considered for the economic model was the time composite renal outcome and this is discussed in Section 3.2.8.4.

Figure 6. Change in the estimated glomerular filtration rate, based on the TS and measurements up to one day after the last intake of study medication (reproduced from CS, Figure 9)



3.2.8.4 Time to composite renal outcome

The composite renal endpoint was comprised of chronic dialysis (with a frequency of twice per week or more for at least 90 days), renal transplant, sustained reduction in eGFR from baseline of ≥40%, sustained eGFR <15 mL/min/1.73 m^2 for patients with baseline eGFR ≥30 mL/min/1.73 m^2, or sustained eGFR <10 mL/min/1.73 m^2 for patients with baseline eGFR <30 mL/min/1.73 m^2. The risk of the composite renal endpoint was similar between the empagliflozin and the placebo treatment

[1] https://investor.lilly.com/news-releases/news-release-details/jardiancer-phase-iii-empa-kidney-trial-will-stop-early-due-clear

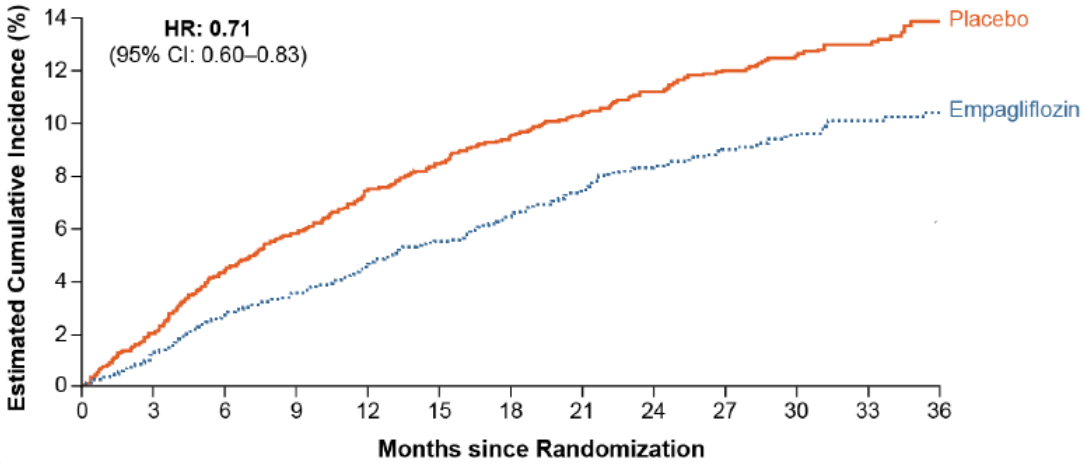
groups (HR 0.95; 95% CI: 0.73 to 1.24, p<0.001). The analysis was undertaken using a Cox regression model including covariates age, baseline eGFR (CKD-EPI)cr, region, baseline diabetes status, sex, baseline EF, and treatment.

This composite renal outcome was not used in the economic model for this submission but it was used by the company in a prior submission for empagliflozin in the HFrEF population.[3] The EAG queried the omission at the clarification stage and the company response stated that in the EMPEROR-Reduced trial[14] there was a statistically significant benefit in this outcome and in EMPEROR-Preserved there was only a numerical benefit. Therefore, the company decided to take a conservative approach and not include the outcome in the economic model for this submission. They also stated that including the composite renal outcome would not affect the cost-effectiveness results as the impact of safety is largely reflected by the inclusion of adverse events, for which the impact is minimal.

3.2.8.5 Time to first adjudicated HHF

Statistically Significantly fewer people experienced the event of first adjudicated HHF in the empagliflozin group (259 of 2,997, 8.6%) compared to the placebo group (352 of 2,991, 11.8%). Figure 7 is the estimated cumulative incidence of the first adjudicated HHF, considering all-cause mortality as a competing risk. The risk of adjudicated HHF was significantly reduced with empagliflozin treatment versus placebo (HR 0.71; 95% CI: 0.60 to 0.83). The analysis was undertaken using a Cox regression model including covariates age, baseline eGFR (CKD-EPI)cr, region, baseline diabetes status, sex, baseline EF, and treatment. This outcome was not used in the economic model.

Figure 7. Time to the first adjudicated HHF (reproduced from CS, Figure 11)

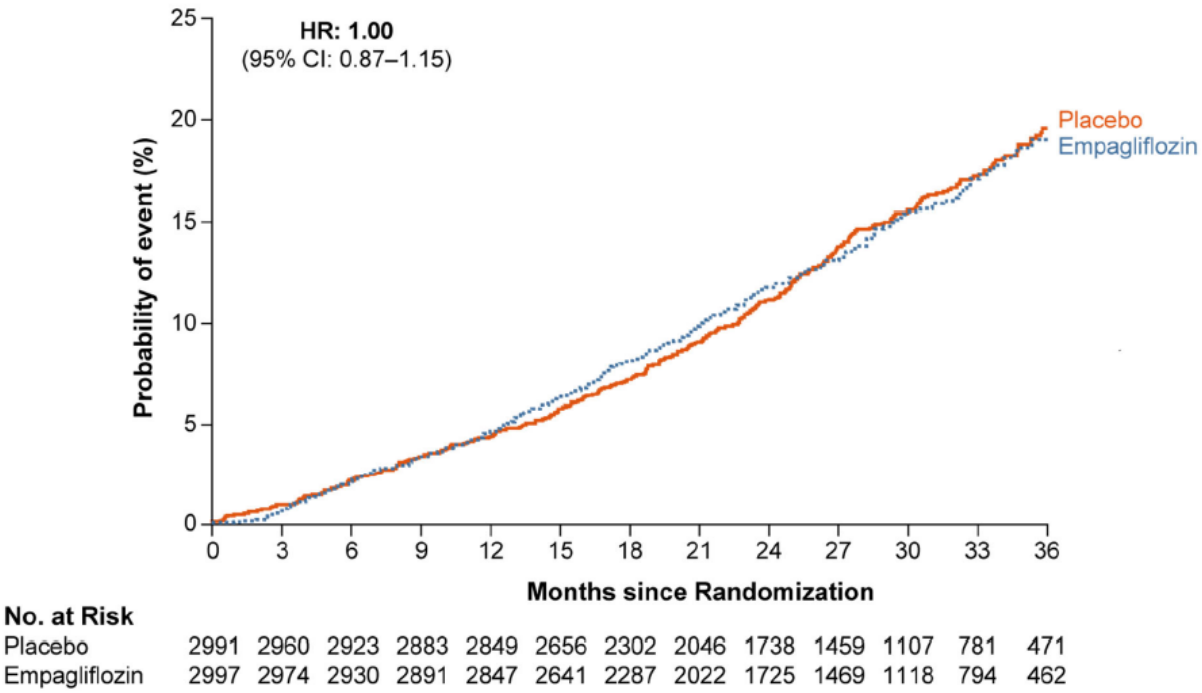


No. at Risk													
Placebo	2991	2888	2786	2706	2627	2424	2066	1821	1534	1278	961	681	400
Empagliflozin	2997	2928	2843	2780	2708	2491	2134	1858	1578	1332	1005	709	402

3.2.8.6 All-cause mortality

The Kaplan-Meier estimate of time to all-cause mortality in the RS is shown in Figure 8. Death from any cause occurred in 422 patients (14.1%) in the empagliflozin group and 427 patients (14.3%) in the placebo group. Cox regression of time to all-cause mortality data for all randomised patients did not show a difference between treatment groups (HR 1.00; 95% CI: 0.87 to 1.15). All-cause mortality was used by the company as an input in the economic model.

Figure 8. Kaplan-Meier estimate of time to all-cause mortality in all randomised patients (reproduced from CS, Figure 12)



The EAG requested all-cause mortality data in three LVEF subgroups at clarification to assess the effectiveness of empagliflozin across three LVEF subgroups. This is presented in Table 20 below. A [redacted] for empagliflozin was found for the LVEF 50% to <60% subgroup and a [redacted] for placebo in the LVEF ≥60% subgroup. There was a [redacted] between empagliflozin and placebo in the LVEF 40% to <50%. This should be interpreted with caution as the study was not powered to detect a significant difference between treatments in smaller subgroups. The EAG also notes there is no linear relationship between LVEF and overall mortality across the subgroups.

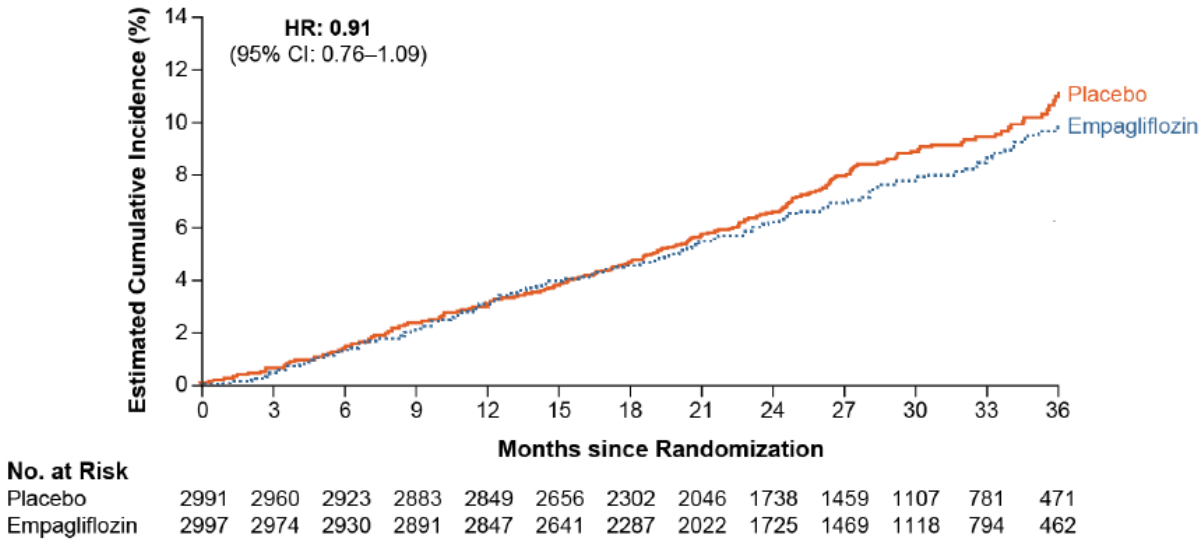
Table 20. Outcome data for all-cause mortality by left ventricular ejection fraction (randomised set)

Endpoint	Empagliflozin 10 mg N with event/N analysed	Placebo N with event/N analysed	Hazard ratio (95% CI)
Overall mortality			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
Abbreviations:			

3.2.8.7 *Cardiovascular mortality*

Most of the deaths recorded during the study were CV-related, due to sudden cardiac death or HF. Adjudicated CV death occurred in 219 people (7.3%) in the empagliflozin group and 244 people (8.2%) in the placebo group. There was a reduced risk of CV death in the empagliflozin relative to placebo (HR 0.91; 95% CI: 0.76 to 1.09) but it was not statistically significant. Figure 9 is the cumulative incidence of adjudicated CV death in the RS, considering non-CV death as a competing risk. The company used this outcome as an input in the economic model.

Figure 9. Cardiovascular death (reproduced from CS, Figure 13)



The EAG requested CV mortality data in three LVEF subgroups at the clarification stage. This is presented in Table 21 below. A [redacted] for empagliflozin was found for the LVEF 40% to <50% and LVEF 50% to <60% subgroups. There was [redacted] between empagliflozin and placebo in the LVEF ≥60% group. This should be interpreted with caution as the study was not powered to detect significant differences between treatments in smaller subgroups.

Table 21. Outcome data for CV mortality by left ventricular ejection fraction (randomised set)

Endpoint	Empagliflozin 10 mg N with event/N analysed	Placebo N with event/N analysed	Hazard ratio (95% CI)
CV mortality			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
Abbreviations:			

3.2.8.8 Time to onset of diabetes mellitus (DM) in patients with pre-DM

The onset of diabetes mellitus (DM) in people with pre-DM occurred in 120 of 1,001 people in the empagliflozin group (12.0%) and 137 of 979 people (14.0%) in the placebo group. There was a reduction in risk of the onset of DM with empagliflozin compared to placebo but it was not statistically significant (HR 0.84; 95% CI: 0.65 to 1.07). The company did not use this outcome as an input in the economic model.

3.2.8.9 First and recurrent all-cause hospitalisation

All-cause hospitalisation occurred in 42.4% (1,271/2,997) of people in the empagliflozin group and 44.8% (1,340/2,991) in the placebo group. The risk of recurrent all-cause hospitalisation was reduced with empagliflozin treatment compared to placebo, but it was a small reduction and did not reach statistical significance (HR 0.93; 95% CI: 0.85 to 1.01). The analysis used a joint frailty model that accounted for the dependence between recurrent all-cause hospitalisation and all-cause mortality. A Cox regression of first all-cause hospitalisation showed a small but statistically significant reduction (2.33%) in risk of with empagliflozin compared to placebo (HR 0.92; 95% CI: 0.85 to 0.99). The company did not use these outcomes as inputs in the economic model.

3.2.8.10 Kansas City Cardiomyopathy Questionnaire (KCCQ) score

The KCCQ asks questions on the frequency and severity of symptoms, the physical and social limitations associated with those symptoms, and persons' perceptions of the impact of their symptoms and function on their quality of life. It is scored on a 0-100 scale and higher values indicate a better health status. A 5-point change in the KCCQ score is commonly considered to be a clinically significant difference in health status in people with HF.[15]

The company reports the KCCQ overall summary score (OSS), total symptom score (TSS), and clinical summary score (CSS). The TSS quantifies the symptom frequency and severity, CSS includes the TSS, and the physical function domain, and OSS includes the TSS, CSS, quality of life, and social function.

In the CS, the company reported the KCCQ score change from baseline in the TS using per-protocol analysis at week 52. This is reported in Table 22 below and all three scores, CSS, OSS, and TSS, show a small but statistically significant benefit for empagliflozin versus placebo. None of the point estimates attained a 5-point change in the KCCQ score which is commonly considered to represent a clinically meaningful difference in health status. However, as highlighted in Table 32 in Section

4.2.6.2, 10% of people had an increase of 23.4 in KCCQ-CSS at 12 weeks and 10% of people had a decrease of 12 in KCCQ-CSS at 12 weeks. This is an indication that empagliflozin offers a large benefit in KCCQ-CSS for many people in this the HFmrEF and HFpEF population.

Table 22. KCCQ score at 52 weeks (adapted from CS, Table 22)

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
QoL measured by KCCQ at 52 weeks, TS		
Change in clinical summary score at 52 weeks (±SE)	4.51±0.31	3.18±0.31
Adjusted mean change from baseline (95% CI)	1.32 (0.45 to 2.19)	
Nominal p-value	0.0028	
Change in overall summary score at 52 weeks (±SE)	5.03±0.30	3.66±0.31
Adjusted mean change from baseline (95% CI)	1.37 (0.52 to 2.21)	
Nominal p-value	0.0015	
Change in total symptom score at 52 weeks (±SE)	5.89±0.34	3.95±0.34
Adjusted mean change from baseline (95% CI)	1.94 (1.01 to 2.88)	
Nominal p-value	<0.0001	
The clinical summary score on the Kansas City Cardiomyopathy Questionnaire ranges from 0 to 100, with higher scores indicating a better quality of life. Analysis of PRO data with a MMRM was based on the treated set (TS) and using on-treatment values only.		

The KCCQ-CSS score at weeks 12; 32; and 52 outcomes were utilised as inputs in the economic model. The company indicated missing values from visits were imputed using the last-observation-carried-forward (LOCF) method. As stated in Section 3.2.4, the EAG is concerned data were imputed under the assumption that data were missing at random. The EAG is unclear how much data were imputed for each timepoint and for how long observations were carried forward. KCCQ was measured at [redacted], [redacted], [redacted], [redacted], and [redacted]. It is unclear if, for example, KCCQ scores measured at [redacted] were carried forward through LOCF to visit [redacted]. Details of how much evidence were imputed at each timepoint and from what timepoint the imputed evidence came would help resolve this concern.

At the clarification stage, the KCCQ-CSS data was requested, with and without imputation, at all time points where it was collected. The company response is presented below in Table 23.

The company provided the KCCQ-CSS outcome at 12 weeks, 32, weeks, and 52 weeks with imputation and without imputation. The imputation did [redacted] the mean difference between treatment groups, but where a difference is seen, [redacted] the empagliflozin group. The EAG is concerned that the company reported an identical number of people in the dataset with imputation and the dataset of people without imputation. This means it is not clear from the response how much data were imputed for these calculations. Table 16 indicates that [redacted] people in the empagliflozin group and [redacted] people in the placebo group were being followed at 52 weeks and Table 23 indicates KCCQ data of [redacted] people in the empagliflozin group and [redacted] people in the placebo group were collected. Therefore, it would appear that data were imputed for [redacted] people missing KCCQ data in the empagliflozin group and [redacted] people missing KCCQ data in the placebo group. Analyses using LOCF are of questionable veracity[16] due to the data missing at random assumption, and the analysis at 52 weeks appears to impute at least [redacted] of the values in each treatment group. For this reason, the EAG considers the data without imputation to be a truer estimate of the treatment effect.

The company also provided what are stated to be EOT results in the clarification response, reproduced below in Table 23. The EAG understands EOT measurements to be taken in all people who were being observed when the study ended. People who prematurely discontinued trial medication performed the EOT visit and the follow-up visit, and then continued with scheduled visits until the trial was stopped.

2997 participants were randomised to the empagliflozin group and 422 of the group died during the study and could not provide EOT measurements. The CONSORT diagram (Figure 6, in the CS) indicates 84 had incomplete follow-up for the primary endpoint and this gives an indication of how many people were lost to EOT follow-up. Therefore, EOT outcome data could be expected in the remaining 2491 participants in the study, but the company response only included data for [redacted] participants in the empagliflozin group. There is a similar discrepancy in the EOT reporting in the placebo treatment group.

In addition to the unclear reporting stated above, a footnote in Table 23 states "[redacted]." This implies no KCCQ data was collected after 52 weeks but this is not in line with the Outcomes trial flow chart in the CSR (Table 9.5.1).

The EAG does not consider the EOT results reported by the company to be interpretable due to uncertainty on whose data is being reported and at what time point this occurred. The EAG agrees with the company's clarification response that stated using this data in the economic model would lead to biased results.

Table 23: KCCQ score with and without imputation (adapted from clarification response)

Timepoint with/without imputation		N	Mean	SD	p-value
Without imputation					
Empagliflozin 10mg	Change from baseline to week 12	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
With imputation					
Empagliflozin 10mg	Change from baseline to week 12	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
Without imputation					
Empagliflozin 10mg	Change from baseline to week 32	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
With imputation					
Empagliflozin 10mg	Change from baseline to week 32	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
Without imputation					
Empagliflozin 10mg	Change from baseline to week 52	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
With imputation					
Empagliflozin 10mg	Change from baseline to week 52	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
Without imputation					
Empagliflozin 10mg	Change from baseline to EOT	[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]
No imputation was performed for data at the EOT timepoint					
Empagliflozin 10mg	Change from baseline to EOT				
Placebo					

Abbreviations: EOT change from baseline is presented, which is the last visit (up to week 52) where the patient had an observation and is the same in the imputed and non-imputed datasets.

The EAG requested KCCQ-CSS data, at 52 weeks and EOT, in three LVEF subgroups at clarification. This is presented in Table 24 below. At 52 weeks, the study found a [redacted] of empagliflozin in all 3 subgroups but it was only [redacted] in the LVEF 50% to <60% group. In the LVEF ≥60% subgroup, the effect size was [redacted] with 95% confidence intervals indicating a [redacted] about the effect.

The EOT results indicate a [redacted] for empagliflozin in all 3 subgroups and it was [redacted] in two of these; LVEF 40% to <50% and LVEF 50% to <60%. In the LVEF ≥60% subgroup, the benefit was [redacted] and the effect size was [redacted] with 95% confidence intervals indicating [redacted] about the effect. This should be interpreted with caution as the study was not powered to detect significant differences between treatments in smaller subgroups.

Table 24. Outcome data for KCCQ-CSS scores by left ventricular ejection fraction (randomised set)

Endpoint	Empagliflozin 10 mg N with event/N analysed	Placebo N with event/N analysed	Hazard ratio (95% CI)
KCCQ-CSS at 0-52 weeks			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
KCCQ-CSS at baseline to end of treatment visit			
LVEF 40% to <50%	[redacted]	[redacted]	[redacted]
LVEF 50% to <60%	[redacted]	[redacted]	[redacted]
LVEF ≥60%	[redacted]	[redacted]	[redacted]
Abbreviations:			

3.2.8.11 Further exploratory secondary endpoints

"Further exploratory secondary endpoints" are reported in the CS and reproduced below in Table 25. The EAG notes a [redacted] of people in the empagliflozin group had a myocardial infarction (MI), and similarly, a slightly [redacted] had a stroke. These were not [redacted] results and the variation between groups was small. The clinical experts advising the EAG indicated that they were not aware of any biological reason to link these adverse outcomes to treatment with empagliflozin. They also noted that other trials of SGLT2 inhibitors, including those of empagliflozin, had not found a relative increase in the occurrence of MI or stroke. The outcomes are not used in the economic model.

Table 25. Summary of further exploratory secondary endpoints from EMPEROR-Preserved study (adapted from Table 22 in the CS)

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
Time to adjudicated MI (fatal or non-fatal), RS		
Patients with MI, N (%)	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
HR vs placebo (95% CI)	[redacted]	

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
Nominal p-value	[redacted]	
Time to adjudicated stroke (fatal or non-fatal), RS		
Patients with stroke, N (%)	[redacted]	[redacted]
Ischaemic	[redacted]	[redacted]
Haemorrhagic	[redacted]	[redacted]
Unclassified	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
HR vs placebo (95% CI)	[redacted]	
Nominal p-value	[redacted]	
Patients with fatal stroke	[redacted]	[redacted]
Time to new onset of Afib, as ECG finding or as AE, RS		
Patients without baseline or history of Afib[a], N (%)	[redacted]	[redacted]
Patients with new onset of Afib, N (%)	[redacted]	[redacted]
Incidence rate per 100 years at risk	[redacted]	[redacted]
HR vs placebo (95% CI)	[redacted]	
Nominal p-value	[redacted]	
Blood pressure (mm Hg) changes from baseline to week 52 (mm Hg), RS		
Systolic blood pressure change (±SE)	-1.8±0.3	-0.6±0.3
Adjusted mean difference (95% CI)	-1.2 (-2.1 to -0.3)	
p-value	0.01	
Diastolic blood pressure change (±SE)	-0.9±0.2	-0.7±0.2
Adjusted mean difference (95% CI)	-0.2 (-0.7 to 0.3)	
p-value	0.46	
HbA1c (%) change from baseline to week 52, RS patients with diabetes		
Adjusted mean change from baseline (±SE)	-0.16±0.02	-0.03±0.02
Adjusted mean difference (95% CI)	-0.19 (-0.25 to -0.14)	
p-value	<0.0001	
Body weight (kg) change from baseline to week 52, RS		
Adjusted mean change from baseline (±SE)	-1.39±0.09	-0.11±0.09

Endpoint	Empagliflozin (N=2,997)	Placebo (N=2,991)
Adjusted mean difference (95% CI)	-1.28 (-1.54 to -1.03)	
p-value	<0.0001	
Haematocrit (%) change from baseline to week 52, RS		
Adjusted mean change from baseline (±SE)	1.94±0.07	-0.41±0.07
Adjusted mean difference (95% CI)	2.36 (2.17 to 2.54)	
p-value	<0.0001	
NT-proBNP (pg/mL) change from baseline to week 52, RS		
Adjusted median change from baseline (IQR)	-29 (-335 to 263)	-9 (-286 to 322)
Adjusted geometric mean ratio	0.95 (0.91 to 0.99)	
p-value	0.01	
Uric acid (mg/dL) change from baseline to week 52, RS		
Adjusted mean change from baseline (±SE)	-0.90±0.03	-0.10±0.03
Adjusted mean difference (95% CI)	-0.80 (-0.88 to -0.72)	
p-value	<0.0001	
Abbreviations: AE, adverse event; Afib, atrial fibrillation; CI, confidence interval; ECG, electrocardiogram; HbA1c, glycated haemoglobin; HR, hazard ratio; IQR, interquartile range; KCCQ, Kansas City Cardiomyopathy Questionnaire; MI, myocardial infarction; QoL, quality of life; RS, randomised set; SE, standard error; TS, treated set. Note: Plus-minus values are means ± SE		

3.2.8.12 Adverse events

Median exposure to study medication was approximately 23 months in both treatment groups, with 84% of patients treated for at least 1 year. An overall summary of adverse events (AEs), taken from the CS, is presented in Table 26 below. This indicated numbers of serious adverse events (SAEs) were similar between groups and the most frequent were cardiac failure, atrial fibrillation, pneumonia, and acute kidney injury. All other SAE were reported in fewer than 3.0% of participants per treatment group. 19.1% of people in the empagliflozin group and 18.4% of people in the placebo group had an AE leading to discontinuation of study medication. The EAG is assured by the similar proportions discontinuing in each treatment group.

Table 26. Overall summary of AE in the TS (reproduced from CS, Table 23)

Category of AE	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Number of patients in the TS, N (%)	2,996 (100.0)	2,989 (100.0)

Category of AE	Empagliflozin 10 mg, N (%)	Placebo, N (%)
Patients with any AE	2,574 (85.9)	2,585 (86.5)
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]
Severe	[redacted]	[redacted]
Investigator-defined drug-related AE	[redacted]	[redacted]
AE leading to discontinuation of study medication	571 (19.1)	551 (18.4)
Serious AE	1,436 (47.9)	1,543 (51.6)
Serious AE		
Resulting in death	[redacted]	[redacted]
Life threatening	[redacted]	[redacted]
Persistent or significant disability/incapacity	[redacted]	[redacted]
Requires or prolongs hospitalisation	[redacted]	[redacted]
Congenital anomaly or birth defect	[redacted]	[redacted]
Other medically important serious event[a]	[redacted]	[redacted]

Abbreviations: AE, adverse event; TS, treated set.

Note: Percentages calculated using total number of patients per treatment as the denominator. A patient may be counted in more than one seriousness criterion.

[a]Other medically important serious event was defined as any important medical event (when based upon appropriate medical judgement) which might jeopardise the patient and might require medical or surgical intervention to prevent one of the other serious outcomes included in the definition of serious adverse events shown in the table above. Examples of such events could be intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias, or convulsions that do not result in hospitalisation or development of dependency or abuse.

Source: EMPEROR-Preserved CSR, Table 15.3.1.1 (139)

3.2.8.13 PULSE

The company used the PULSE study to validate the event rates projected by the model, this is reproduced from the CS in Table 27 below. The model-predicted rates of CV mortality and HF hospitalisation were [redacted] than that seen in PULSE and the EAG requested an explanation at the clarification stage.

The company explained that this was likely due to the inaccurate recording of events in PULSE rather than a challenge to the validity of EMPEROR-Preserved data and/or the model. PULSE gathered 'real-world' data and in EMPEROR-Preserved outcomes were adjudicated by Committee according to a strict protocol. The company argued that HHF and CV mortality would be under-reported in PULSE because general physicians and other specialists may not recognise the symptoms of acute HF. This

would not have occurred in EMPEROR-Preserved where the outcomes were adjudicated and the study protocol was prepared to ensure these were correctly identified. The company also stated that 73% of people in PULSE had "unknown" EF which will have included a number with HFmrEF/HFpEF. Correct categorisation of this group would have provided HHF and CV-mortality outcome data that was a more accurate reflection of the HFmrEF/HFpEF population.

Table 27. Model-predicted vs observed rates per 100 patient-years in PULSE (reproduced from CS, Table 48)

Characteristics	PULSE SoC (Events per 100 patient-years)	Model simulation Pulse - Placebo (60 months) (Events per 100 patient-years)
Non-CV mortality	[redacted]	[redacted]
CV mortality	[redacted]	[redacted]
All-cause mortality	[redacted]	[redacted]
HF hospitalisation	[redacted]	[redacted]
Abbreviations:		

3.3 Conclusions of the clinical effectiveness section

The company presented an accurate introduction detailing the signs, symptoms, and underlying pathologies of people diagnosed with HF. There were a number of additional clinical factors the EAG's experts noted. After diagnosis, many people stay in the same LVEF category, though a person's LVEF can improve due to HF treatment. They also stated that, compared to people with HFrEF, people with HFmrEF/HFpEF are a more heterogenous pathophysiological group and do not have a set of underlying conditions that are treated with the same drugs. Therefore, treatment options for this patient population are severely limited and, as such, there is an unmet need for an effective treatment to be made available.

The company's submission addressed the decision problem stated in the final scope issued by NICE. The EAG considers the literature searches conducted for the SLR to be adequate to find all studies relevant to the decision problem. However, there were 18 studies stated to be included in the SLR but only the EMPEROR-Preserved trial was presented. Thus, 17 studies were effectively excluded without any reasoning given for this exclusion. One study, EMPERIAL-Preserved, was an investigation into empagliflozin in the HFmrEF/HFpEF population and investigated relevant outcomes such as KCCQ scores and safety. This was a 3-month study and should have been presented in this submission.

The ERG considers EMPEROR-Preserved, the key study informing the clinical effectiveness of empagliflozin, to be a well-designed and well-conducted RCT, with an overall low risk of bias and high internal validity. The eligibility criteria for the study were in line with the decision problem. The EAG noted people who were previously measured to have an LVEF ≤40% were excluded from the study and this was a conservative decision as inclusion of this population would have biased the results in favour of empagliflozin.

The EAG's clinical experts commented that the trial participants in EMPEROR-Preserved are broadly representative of people with HFmrEF/HFpEF seen in England and Wales. The experts noted that the people recruited to the trial had a mean age of 72 years old and this is younger than the people in their practice who are more often in their 80s. Also, the proportion of people with HFmrEF/HFpEF and hypertension or T2DM would be slightly lower in the England and Wales population than were present in the study.

The study groups were well-balanced for background medications. The EAG's clinical experts commented that the baseline medications reflected the people recruited to the study. For example, a slightly higher proportion of people with hypertension participating in the trial led to different baseline medications. People with hypertension can be treated with ACEI and beta-blockers, and consequently the number of people using ACEI and beta-blockers might be slightly lower in clinical practice. In each group, [redacted] of people were treated with [redacted]. [redacted] does not have marketing authorisation for use in people with HFmrEF/HFpEF in the UK but it does have marketing authorisation in the HFmrEF population in the USA.

Discontinuations for reasons other than death were 23.2% in the empagliflozin group and 23.4 % in the placebo group. Discontinuations were lower in the EMPEROR-Reduced trial of empagliflozin in the HFrEF population and the company indicated the higher proportions in EMPEROR-Preserved were due to the higher age of the population recruited and that the HFmrEF/HFpEF population is more comorbid. They also commented that EMPEROR-Preserved was a longer trial with a median treatment period of 26 months as compared to 16 months in EMPEROR-Reduced.

The analysis used to account for missing data utilised several different methods including the multiple imputations framework, multivariate Cox regression models, and last observation carried forward (LOCF). All three techniques share the same assumption that data are missing at random and impute data points that reflect the data that was collected. The EAG is concerned the

assumption does not account for participants who withdrew from the study due to the ineffectiveness of the intervention. The EAG considers unadjusted data to be a truer reflection of the treatment effect than data with imputed values that make the missing at random assumption.

The primary composite outcome of CV death or HHF demonstrated a statistically significant benefit for empagliflozin compared to placebo ($p<0.001$). The EAG notes from the plots for the individual outcomes of CV death and HHF, that the benefit for empagliflozin appears to be predominantly driven by fewer HHF events than placebo. This primary outcome was not used in the economic model although its composite parts were.

Empagliflozin had a statistically significant benefit for total HHF (first and recurrent) in the primary analysis ($p<0.001$) and five sensitivity analyses. There was no significant difference between empagliflozin and placebo for CV mortality or all-cause mortality ($p=0.3$ and $p=0.$ 0.9893, respectively) and both are included in the company's base case economic model. There was no significant difference between empagliflozin and placebo for the time to the composite renal outcome ($p=0.7243$). This outcome was not used in the economic model and the company indicated renal outcomes are adequately incorporated into the model through adverse events data and that the effect of using the outcome in the model would be negligible as there was a numerical benefit for the empagliflozin group.

The KCCQ-CSS score was the clinical outcome that captured disease severity and progression. It was measured at weeks 12; 32; 52; and EOT. In the CS, the company reported a statistically significant for empagliflozin over placebo in KCCQ-TSS, KCCQ-CSS, and KCCQ-OSS scores at 52 weeks ($p<0.0001$, $p<0.0001$, $p<0.0001$, respectively). However, none of the outcomes reached a 5-point change in the KCCQ score that is commonly considered to be a clinically significant difference in health status in people with HF.

The data at weeks 12; 32; 52 was used in the economic model and the company accounted for missing data by using LOCF. This technique assumes data are missing at random and the EAG do not agree that this is a reasonable assumption as people will have left the study due to the ineffectiveness of treatment. The company provided unadjusted data at the clarification stage and this was similar to the to the adjusted data. However, in all cases where the adjusted data varied from the unadjusted data, it showed an increased comparative benefit for empagliflozin. The EAG considers the unadjusted data to be a less biased reflection of the outcome.

The adverse events outcome was reported in the treated set (TS) and were similar between treatment groups. Serious adverse events were experienced by 51.6% of participants in the placebo group and 47.9% of people in the empagliflozin group. One category where there was some variation between treatment groups was investigator-defined drug-related AE and they were experienced by [redacted] in the empagliflozin group and [redacted] in the placebo group.

The company used the PULSE study to validate the event rates projected by the model. The model-predicted rates of CV mortality and HF hospitalisation were notably higher than that seen in PULSE. The company argued that HHF and CV mortality would be under-reported in PULSE because general physicians and other specialists may not recognise the symptoms of acute HF. The company also argued that most of the people in PULSE had unknown LVEF and this may have led to less representative HF groups defined by LVEF.

4 Cost effectiveness

Empagliflozin is currently recommended for the treatment of chronic heart failure with reduced ejection fraction (HFrEF) in the NHS (NICE appraisal TA773). The company's approach to the present submission was to use the same cost utility model, however, utilizing the pivotal EMPEROR-Preserved trial instead of EMPEROR-Reduced as the main source of clinical evidence to inform the economic model.

The company's approach also relied on testing what it considered to be the key uncertainties in the economic evidence identified by the EAG in TA773: whether the cost utility model accurately predicted the rate of deaths; hospitalisations; and treatment discontinuation compared to the rate observed in EMPEROR-Reduced (and in this case EMPEROR-Preserved).

Nonetheless, the EAG notes that in TA773 the committee focused on the comparison of empagliflozin with dapagliflozin, whereas the relevant comparator for the current submission is standard of care (SoC). Furthermore, the former appraisal relied on the company's Butcher indirect treatment comparison, while the latter relies on different statistical methods. Therefore, the EAG report provides a description of the company's approach in its entirety, along with its appropriateness in the population with chronic HF with mildly reduced LVEF (HFmrEF) and chronic HF with preserved LVEF (HFpEF), hereafter referred to as (HFmrEF/HFpEF.

chronic heart failure with preserved ejection fraction (HFpEF), and only draws comparisons to TA733 where relevant to the discussion.

4.1 EAG comment on the company's review of cost effectiveness evidence

The company conducted a systematic literature review (SLR) to identify existing economic evaluations for the treatment of chronic HFmrEF/HFpEF. The company used the same SLR protocol as that described in TA733, with the difference that only studies for patients with an EF<40% were included in the PRISMA diagram. Full details of the process and methods used are described in Appendix G of the CS.

In summary, following the abstract and full-text screening process, no relevant studies were identified for inclusion in the SLR and therefore no quality assessment was conducted.

The SLR of cost-effectiveness studies, described in Appendix G, did not identify any suitable economic evaluations in the chronic HF with EF >40% population. Therefore, the cost-effectiveness model for economic evaluation of empagliflozin + SoC for chronic HF patients with EF >40% builds on the modelling approach previously accepted by the NICE committee for empagliflozin + SoC for patients with EF ≤40% (TA773) and the economic model submitted for dapagliflozin in HFrEF to NICE, which the EAG considers appropriate.

4.2 Summary and critique of company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 28 summarises the EAG's appraisal of the company's economic evaluation against the requirements set out in the NICE reference case checklist for the base-case analysis, with reference to the NICE final scope outlined in Section 2.

Table 28. NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Yes.
Perspective on costs	NHS and PSS	Yes.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Yes.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes.
Synthesis of evidence on health effects	Based on systematic review	Yes.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Yes.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes.

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	Yes.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Yes.

Abbreviations: EAG, evidence review group; NHS, national health service; PSS, personal social services; QALY, quality adjusted life year

4.2.2 Population

The population in the base case economic model consists of the ITT population from EMPEROR-Preserved, which included adults with chronic HFmrEF/HFpEF, with or without diabetes.

Clinical expert opinion provided to the EAG was that the HFmrEF/HFpEF population in the UK is on average older than that in EMPEROR-Preserved (80 years vs 72 years at baseline) and presents with considerable comorbidities. During clarification, the EAG requested that the company undertook a scenario in the model where the baseline age for the UK population was reflected. This scenario had a negligible impact on the final ICER.

4.2.3 Interventions and comparators

The intervention included in the economic model was empagliflozin formulated as a 10 mg tablet taken once a day, in addition to SoC. For simplicity, hereafter, the EAG refers to the intervention as empagliflozin.

SoC was modelled as a basket of drugs used in first-line heart failure care. The assumed proportions of each drug used in SoC is given in Table 29. The clinical experts advising the EAG agreed with the drugs included in the company's basket of SoC treatments, with the exception of the use of angiotensin receptor-neprilysin inhibitors (ARNis), which the clinical experts indicated would not be used in UK clinical practice for HFmrEF/HFpEF patients. The experts also indicated that the proportion of loop or high ceiling diuretics used in the UK is higher at 80%. During clarification, the company conducted a scenario analysis in the model to reflect the EAG's clinical experts' views and the impact on the final ICER was negligible.

Table 29 – Composition of Standard of Care

Drug category	Proportion in ITT population
ACEi	40%

ARB	39%
ARNi	2%
MRA	37%
Beta blocker	86%
Loop or high ceiling diuretics	68%

Abbreviations: ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor-neprilysin inhibitor; BB, beta-blockers; MRA, mineralocorticoid receptor antagonist; SGLT2i, sodium-glucose co-transporter-2 inhibitor

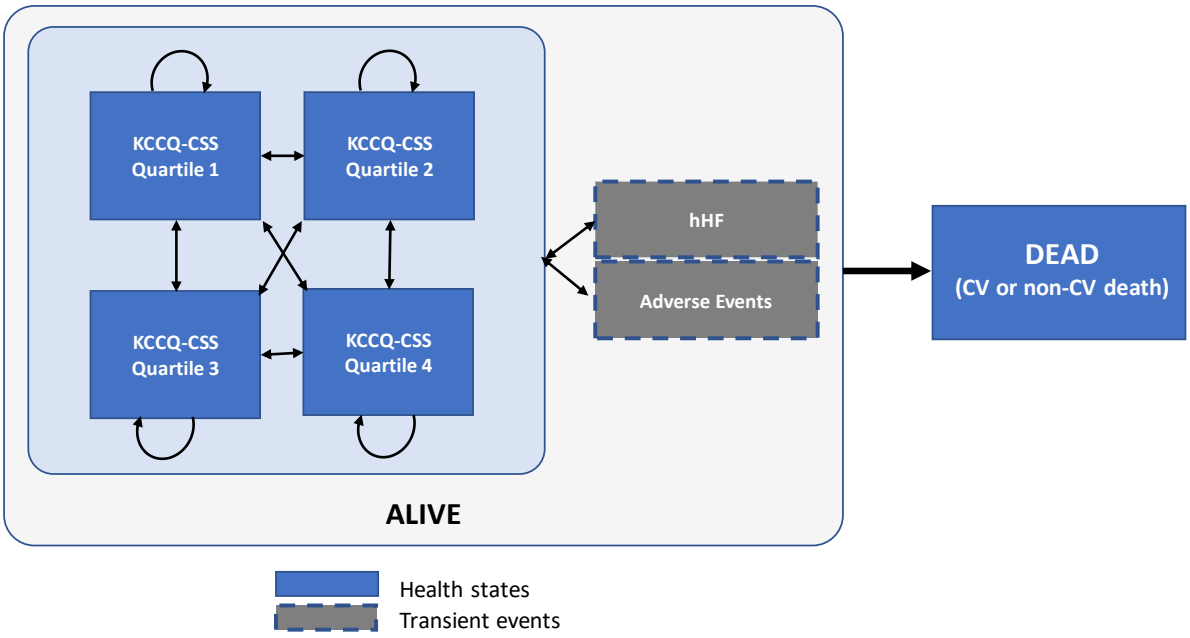
4.2.4 *Modelling approach and model structure*

The company developed a cohort state-transition model with five health states (Figure 10). The four KCCQ-CSS health states represent the different levels of disease severity experienced by patients. KCCQ-CSS quartiles 1 to 4 correspond to KCCQ-CSS scores of 0 to <[redacted], respectively, with higher scores corresponding to a better health status. Patients can transition to a higher or lower KCCQ-CSS quartile; remain in the same state; or die. Patients can have a CV-related death or a non-CV death. In each of the KCCQ-CSS states, patients have a probability of experiencing a hospitalisation for to heart failure (HHF); or a treatment-related adverse event (AE).

The company's model structure allows for the estimation of the relationship between disease progression (measured through movements in the KCCQ-CSS states) on outcomes such as HHF; survival; quality of life; and time on treatment. This was done by introducing different KCCQ-CSS state predictors for each quartile in the HHF; survival; quality of life; and time on treatment risk equations. Therefore, every time a patient moves KCCQ-CSS states in the model, their probability of HHF or death, and their quality of life also changes. This generates indirect benefits associated with empagliflozin (in addition to direct benefits for some of these outcomes), as patients on empagliflozin have a higher probability of moving to the better KCCQ-CSS states. This is further discussed throughout Section 4.2.6; and Section 4.2.8.

Patients can discontinue treatment with empagliflozin at any cycle. After discontinuation, patients in the model were assumed to receive SoC until dead.

Figure 10. Company's model structure



4.2.4.1 EAG critique

The company's justification for the choice of thresholds for the KCCQ-CSS states in the model was that these included an adequate number of patients in each category to permit statistically robust analysis and predicting patient outcomes, when compared to other choices of quartiles. Overall, the EAG is satisfied with the company's choice of KCCQ-CSS states in the model and notes that these are broadly similar to the ones used in TA773.

The company's model did not include renal outcomes, whereas the model used in in TA773 included these. When asked to justify their choice during clarification, the company stated that empagliflozin did not show a statistically significant effect on renal events in EMPEROR-Preserved (hazard ratio of 0.95 with 95% confidence interval of 0.73 to 1.24), whereas in the EMPEROR-Reduced trial this outcome was statistically significant. Given that the inclusion of renal outcomes had a small effect in the cost-effectiveness results in TA773, the EAG considers the company's approach for this STA to be appropriate.

4.2.5 Perspective, time horizon and discounting

A lifetime horizon was adopted in the model and time was discretised into monthly cycles with a half-cycle correction applied. The analysis was carried out from an NHS and Personal Social Services

(PSS) perspective. Costs and health effects are discounted at an annual rate of 3.5%, in line with the NICE Reference Case.

The EAG agrees with the lifetime horizon used, and notes that the model adopts a time horizon of 28 years and that patients' baseline age in the company's model was 72 years.

4.2.6 Treatment effectiveness

Treatment effectiveness was modelled through patients' change in KCCQ-CSS state; the rate of HHF; and mortality. These are discussed in the next sections in detail.

Analyses of overall survival (OS), CV-related mortality and HHF rates (as well as quality of life and time to treatment discontinuation) involved developing regression equations which included a time-varying predictor for patients' KCCQ-CSS state, with or without a treatment group predictor.

4.2.6.1 Transition between KCCQ-CSS states

The company estimated transition probabilities (TPs) between the KCCQ-CSS quartiles from the KCCQ-trial data between the three periods of trial visits (baseline to week 12, 12–32 and 32–52 weeks) by treatment arm and assessed the TPs for variation over time. The company concluded that the probability of patients transitioning between KCCQ-CSS states varied across the three time periods, therefore, decided to have three sets of period-specific TPs in the model in each treatment arm.

Each of the six derived matrices was then converted to monthly TPs. This yielded six sets of monthly TPs representing progression in the three periods used in the analysis (reported in Table 28 in the CS). The transition matrices for the last period (week 32+) were used to predict the changes in KCCQ-CSS scores for the rest of model time horizon. When patients discontinued treatment with empagliflozin in the model, the set of TPs used is that of the SoC patients.

The company used the patient KCCQ-CSS scores for the observed cases including data after treatment discontinuation (OC-AD) population. The company also reported using the last-observation carried-forward (LOCF) imputation method to deal with missing KCCQ-CSS data over the first 52 weeks of observations for each patient.

During clarification, the EAG asked the company to justify why the LOCF method was appropriate to input missing KCCQ-CSS values. The company replied that the mean scores at weeks 12, 32, and 52

(Table 30) were very close between the imputed and non-imputed datasets, and that the distribution of KCCQ-CSS score changes from baseline were also similar between the imputed and non-imputed datasets (Table 8 and Table 9 in company's response to clarification questions). Given the similarity between imputed and non-imputed KCCQ-CSS scores, the company decided to use the LOCF approach in their base case.

Table 30. Comparison of KCCQ-CSS score statistics for imputed and non-imputed data (Randomised set)

Without imputation	Empagliflozin 10mg		Placebo	
Visit	**N**	**Mean (SD)**	**N**	**Mean (SD)**
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]
With imputation	**Empagliflozin 10mg**		**Placebo**	
Visit	**N**	**Mean (SD)**	**N**	**Mean (SD)**
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]

*The higher number of observations at week 12 is due to records from patients with missing scores at baseline. These patients contributed data from week 12 onwards and were kept in the analysis.

Abbreviations: KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; SD, standard deviation.

4.2.6.2 EAG critique

Overall, the EAG notes that it would have been helpful to have more clarity around the collection and analyses of KCCQ-CSS data. In their response to clarification, the company reports that, "*No data beyond week 52 were available* [and that] *EOT* [end of treatment] *change from baseline is presented, which is the last visit (up to week 52) where the patient had an observation.*" the company also reports that, "*EOT data were available but since EOT varies between different patients, the company believes that it should not be utilised in the model as this would lead to biased results*". From the company's response, the EAG concludes that EOT was defined as any point in time before individual

patients concluded their 52 weeks of treatment with treatment. Nonetheless, the clinical study report (CSR) for EMPEROR-Preserved states that, "*when a patient permanently discontinued trial medication or when the required number of events had been reached for the trial, an end-of-treatment (EOT) visit was performed, followed by a follow-up Visit 30 days later. Patients who prematurely discontinued trial medication performed the EOT visit and the follow-up visit, and then continued with scheduled visits (including the second EOT visit) until the trial was stopped"*. The CSR also reports a flow chart with scheduled assessments of efficacy (Table 9.5.1:1), where the end of trial period is defined as 148 weeks.

The EAG remains unclear if and why KCCQ-CSS data beyond 52 weeks in the model were not collected. Regardless, the EAG disagrees with the company's approach to using the LOCF method to deal with missing data. Even though the EAG agrees with the company's assessment that the mean KCCQ-CSS scores at the different time points (and the mean changes from baseline) with and without imputation are broadly similar, the impact of using the observed values without imputation in the model is large, increasing the company's base case ICER from £14,429 to £20,198 per QALY gained.

Therefore, while the mean KCCQ-CSS might not vary much across the two methods, the TP derived with the two methods are substantially different, particularly for the week 32 (month 9) onwards TPs, used throughout the economic model until patients die (Table 31). This is related to the same observation made by the EAG during TA733, that even though the mean and median changes (and absolute) KCCQ-CSS values might be relatively small in the trial (and not achieve the clinically meaningful 5 points), the underlying changes in the clinical scores were much higher on an individual-patient level. An example of this is provided in Table 32, where even though the mean and median change from baseline to week 12 in KCCQ-CSS scores was below [redacted], the quartiles (or deciles) of observations show that , for example, 10% of patients in the empagliflozin arm had an increase in KCCQ-CSS scores of [redacted], while 10% of patients in the same treatment arm had a decrease of [redacted] points, thus "averaging out" mean and median changes.

Importantly, Table 30 shows that the number of observations without imputations is similar across treatment arms in EMPEROR-Preserved suggesting that the data are well balanced, and that missing data is likely due to dropouts or deaths. The appropriateness of the LOCF method also relies on the assumption that the missing observation would produce an identical KCCQ-CSS score from the one captured in the previous data point. Given that the EAG is unclear when EOT occurred for patients

with missing data, and the lack of any data to substantiate the fact that the missing observations all occur in a "plateau" part of the observations, the EAG considers that the more robust approach is to use observed data without imputation.

Table 31. Transition probabilities with and without imputed data (randomised set)

			With imputation			Without imputation		
	From:	To:	Months 1 - 3	Months 4 - 8	Months 9+	Months 1 - 3	Months 4 - 8	Months 9+
Empagliflozin + SoC	1	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	4	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SoC	1	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	4	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SoC, standard of care.

Table 32. Changes in KCCQ-CSS scores without imputation (randomised set)

Without imputation		N*	Mean	SD	p-value	Min	P10	P25	Median	P75	P90	Max
Empagliflozin 10mg	Baseline	[redacted]										
Placebo		[redacted]										
Empagliflozin 10mg	Change from baseline to week 12	[redacted]										
Placebo		[redacted]										

*The total number of patients with available observations on KCCQ-CSS data for each time point (i.e., baseline, week 12,32,52 and EOT). No data beyond week 52 were available. EOT change from baseline is presented, which is the last visit (up to week 52) where the patient had an observation and is the same in the imputed and non-imputed datasets.

Abbreviations: EOT, end of treatment; KCCQ-CSS, Kansas City Cardiomyopathy Questionnaire - clinical summary score; SD, standard deviation.

Finally, in TA773, the EAG asked that the company conducted scenario analyses where the effect of empagliflozin on KCCQ-CSS (sustained by the combination of the proportion of patients in the better KCCQ-CSS states in the empagliflozin arm at month 8 and the low probability of disease progression for both SoC and empagliflozin arms in month 9+) was waned over time in the model. The EAG was worried about the company's underlying modelling assumption that patients still benefit from empagliflozin after they discontinue treatment. Given that the there is a higher percentage of empagliflozin patients in the highest KCCQ-CSS state in the model at month 8, and that the TPs used in month 9+ of the company's model assume that all patients (on treatment and off treatment) have a very small probability of leaving the KCCQ-CSS state they are in at month 8, the benefit associated with empagliflozin was broadly maintained for patients discontinuing treatment after month 8 of the model. The company's assumption that empagliflozin patients experience SoC TPs after discontinuation is only partially conservative and leads to a sustained relative treatment effect for patients in KCCQ-CSS 4 in the model over time.

Due to the company's model structure, this assumption impacts the benefits associated with empagliflozin on HHF and mortality, as these outcomes are dependent on patients' distribution across KCCQ-CSS states.

For this submission, and to pre-empt the concerns raised previously by the EAG, the company included the following scenario analysis in the model:

1. Setting the proportion of patients in the KCCQ-CSS quartiles under the empagliflozin arm equal to those proportions in the placebo arm at 5, 3, 2, and 1 years. The company's scenario analysis assumed that after the proportions were set equal between treatment arms, the TPs respective for each treatment arm would apply.
2. Assuming that the TPs between KCCQ-CSS quartiles for patients on treatment are the same as those for patients off treatment from month 9 onwards in the model.

The EAG considers that the first scenario lacks clinically plausibility as it just creates an artificial drop in the empagliflozin arm proportions, so these are the same as in the SoC arm, but then proceeds to assume the respective TP for each treatment arm. This translates into a scenario where patients experience a sudden loss of treatment effectiveness in the empagliflozin arm but regain it the next cycle of the model.

The second scenario provides a more helpful representation of treatment waning, although it is likely to be overly pessimistic. By assuming that the TPs between KCCQ-CSS quartiles for patients on treatment are the same as those for patients off treatment from month 9 onwards in the model, the proportion of patients in each KCCQ-CSS state is the same in the empagliflozin and SoC arms of the model approximately 4 to 5 years after treatment initiation in the model. This increases the company's base case ICER to £32,482.

4.2.6.3 Hospitalisation for heart failure

The company used count data from EMPEROR-Preserved to model the number of HHFs in the model. There were 541 HHF events observed in the placebo arm and 407 events in the empagliflozin arm of the trial (including repeated hospitalisations).

The monthly rate of HHF events was estimated using a Poisson model with generalised estimating equations with an auto-regressive covariance structure to account for correlations between repeated measures. The company also considered fitting the HHF data with a negative binomial distribution, however, concluded that the fitting procedure failed when using this distribution.

The final Poisson regression used in the model included two predictors: the time-varying KCCQ-CSS states; and treatment received. The regression model was fitted to the ITT population (Table 33)

from EMPEROR-Preserved. The HHF rates were modelled to be constant in each treatment arm, respectively, however the company provided alternative scenarios where the rate of hospitalisations was increased every month. The EAG found a mistake in the coefficients reported in Table 32 of the CS, hence presents the correct estimates in Table 33. The EAG could not find the corresponding p-values of the coefficients, and asks that the company provides these.

Table 33. Poisson regression for HHF, ITT population from EMPEROR-Preserved

Covariate	Coefficient	p-value
Intercept	[redacted]	[redacted]
Treatment effect Empagliflozin 10 mg (reference placebo)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 2 vs Quartile 1)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 3 vs Quartile 1)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 4 vs Quartile 1)	[redacted]	[redacted]

4.2.6.4 EAG critique

As discussed in the EAG review of the company's response to technical engagement (TE) of TA773, a more robust method for estimating HHF in the model would have been to use Kaplan-Meier (KM) data. This would have allowed the rate of HHF to directly vary in the model and in accordance to the underlying observed data (as opposed to the assumption of a constant rate) and would also have allowed first and subsequent hospitalisations to be modelled separately. In this appraisal, the same issues remain relevant as the company used a Poisson regression, assuming a constant risk of hospitalisation in the entire model (regardless of the time-varying element of KCCQ-CSS-linked HHF) and did not differentiate initial and subsequent hospitalisations in the model.

Given that KM data on HHF in EMPEROR-Preserved shows a considerable difference in empagliflozin's effect on first and subsequent hospitalisations, during clarification, the EAG asked that the company used KM data on first and subsequent HHF, respectively, to fit parametric models and extrapolate the rate of HHF events over time in the model. The EAG suggested that if the company's preference was to make HHF dependent on KCCQ-CSS (as was the case in TA773), the same approach undertaken by the company in TA773 of using, "*a counting process setup with start and stop times to create periods defined by the occurrence of each hospitalisation and/or changes in KCCQ-CSS quartiles*" could be used. This approach would allow a patient to have one record per

change in KCCQ-CSS and per HHF with start and stop times of the period defined by the time when these changes occur. The EAG added that regardless of the company's choice of method to estimate HHF, first and subsequent hospitalisations should be modelled separately.

As a response to the EAG's request, the company reported that it conducted an analysis where a Poisson equation with time as a predictor (question B7) was fitted, however, the results generated a negative coefficient (suggesting that HHF risk decreases over time), which the company considered clinically implausible. The company added that the observed pattern was likely attributable to the declining numbers of patients at risk and the disproportionate influence of events near the end of follow-up in the trial. Even though the EAG agrees with the company that a declining rate of HHF over time might not clinically plausible, the EAG notes that a constant rate of hospitalisation is also not clinically plausible.

Furthermore, the company fitted six distributions (exponential, Weibull, lognormal, log-logistic, Gompertz, and generalised Gamma) to time to first HHF KM data, with treatment arm as the only predictor. The company reported that the extrapolated hazard estimates were decreasing or plateauing hazards, which the company considered clinically implausible, and in agreement with the Poisson model that included time as a predictor. The EAG notes that it has not seen any measures of fit (statistical or visual) to the different models used by the company, so it cannot derive any conclusions on the fit of the models, and the respective underlying hazard. Crucially, the company only fitted first HHF with parametric models, and did not undertake the same analysis for subsequent events. The EAG does not understand the company's decision, and notes that it is possible that the trend in the risk of subsequent events could have been different (for example, increasing), which could have added plausibility to the data analysis.

The EAG notes that in the trial [redacted]. The results reported in the EMPEROR-Preserved CSR, not only indicate that the difference across arms in patients with first events was larger ([redacted]) than the difference in the number of patients with second events ([redacted]), but also that it is likely that empagliflozin does not have a benefit in preventing subsequent hospitalisation events. By considering all events from the trial to have been first events in the model the company is therefore likely overestimating the benefit of empagliflozin on reducing HHF events.

The EAG also notes that second HHF eventa occurred "faster" in relation to first HHF events - out of those patients with 2 (or more) events, approximately 50% of patients had already experienced a second event at 3 years after the first event, whereas only about 13% of patients had experienced their first HHF event at 3 years. Furthermore, time to subsequent HHF [redacted] in the empagliflozin than in the placebo arm - at 3 years, [redacted] of patients in the empagliflozin arm had experienced a second HHF, while [redacted] of placebo patients had experienced a second event in the placebo arm.

Therefore, the EAG's view remains that the company's approach does not appropriately capture the hospitalisations in EMPEROR-Preserved, both in the assumption of a constant hospitalisation rate, and in the decision to not separate first and subsequent hospitalisation events. With regards to the latter, the company considered that time to subsequent hospitalisation analysis broke randomisation as baseline was redefined as patients having experienced an initial HHF, thus, the company did not conduct the analysis requested by the EAG. The EAG disagrees with the company's rationale, as second HHF are a conditional event on first HHF but are nonetheless part of the natural disease progression of patients appropriately randomised to the trial.

During clarification the company also provided the number of events observed in EMPEROR-Preserved compared to those predicted in the model (Table 34). The EAG notes that not only the absolute number of HHF events in the model are considerably overestimated in relation to the events observed in the trial for the same period of time, as expected from the company's assumption of a constant rate of HFFs in the model, but the overestimation increases as time progresses in the model. Crucially, even though the observed and predicted differences in events between empagliflozin and placebo are broadly similar at 26 months and at 3 years (indicating less uncertainty in the incremental results of the model), the comparison of incremental estimated vs observed events at the last available time from the trial suggests that the model increasingly overestimates the benefit associated with empagliflozin.

Therefore, despite the model's ability to accurately reproduce the number of incremental HHF observed in the trial up to 3 years, the EAG remains uncertain if HHFs are accurately estimated in the long-term for the trial population, given the company's assumption that HHF is constant and the use of the Poisson model.

Finally, the company conducted a scenario analysis where the rate of HHF increased every month (KCCQ-quartile 1= 0.4%, KCCQ-quartile 2= 0.3%, KCCQ-quartile 3= 0.2%, KCCQ-quartile 4= 0.1%), which resulted in an ICER of £13,86 (lower compared to the base case ICER of £14,429). Nonetheless, the EAG considers that the company's scenario to be flawed and lack face validity as it leads to an even larger overestimation of HHF in the model (Table 35).

The EAG used the company's scenario analysis to run a scenario where the rates of HHF were decreased over time. As expected, this increased the ICER, however did not result in a better fitting model as there was a considerable underestimation of the differential in HHF in the initial part of the model, when compared to the trial data. This reinforces the EAG's view that the "true" risk of HHF varies over time.

The EAG's conclusion remains that given that time to HHF KM data were available from EMPEROR-Preserved, the company could have used these data to model time to first and subsequent HHF, separately. Using the KM HHF data from the trial for first and subsequent events separately would have allowed the company to fit a parametric survival curve to the data and extrapolate into the model's time horizon without having to assume a constant rate of HHF and without having to assume a constant treatment effect with empagliflozin. The EAG notes that using survival curves would have still allowed the company to model HHF by KCCQ-CSS state.

Finally, the EAG notes that in TA773, the company included a scenario in the model where the HHF Poisson model was adjusted by a 0.43 ratio to reflect an overall number of events in the >65 years population. Reducing the overall number of HFF events in both treatment arms of the model was deemed one of the key drivers of the model and increased the final ICER. The EAG expects that decreasing the overall number of HFF in the current model would also lead to an increase in the company's base case ICER.

Table 34. Comparison of total number of HHF observed in the trial and those predicted in the cost-effectiveness model

	EMPEROR-preserved			CE model		
	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference
Total number of HHF (first and subsequent) over 26 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of HHF (first and subsequent) over 3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of HHF (first and subsequent) over 3.58 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 35. Comparison of total number of HHF observed in the trial and those predicted in the cost-effectiveness model in company's scenario analysis

	EMPEROR-preserved			CE model		
	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference
Total number of HHF (first and subsequent) over 26 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of HHF (first and subsequent) over 3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of HHF (first and subsequent) over 3.58 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Finally, during clarification, the EAG raised its concern around the company's decision to exclude age from the risk equations used to estimate HHFs in the model, considering how the likelihood of HHF events is very likely to be impacted by patients' age.

The company replied that the model includes an extended risk equation (as opposed to the reduced base case equation which only considered treatment arm and KCCQ-CSS quartiles as predictors). The company added that, "*other predictors (including age), that underwent a pre-specified variable*

selection process, were included in the extended equations […] *only if found to be statistically significant*". The company also described an iterative process for selecting and removing non-statistically significant variables from the risk equations used in the model.

Overall, the EAG lacks confidence in the company's process and rationale to selecting the final variables included in HFF risk equations:

1. The company reports that age was tested as a covariate in the selection process, however, none of the statistical models reported in the statistical Appendix N of the CS report age as a covariate.
2. The company reported that only statistically significant covariates were included in the final extended equations. However, the company decided to evaluate statistical significance at the 10% level, higher than the more conventionally used 5% significance level (see Table 36).
3. The rationale for using the reduced risk equations is unclear in the CS. The company should have retained all relevant statistically significant variables in the risk equations used in the analysis. Even though the EAG does not know what the final p-values would be for all the covariates included in Table 36 if the non-statistically significant variables were removed from the model, it is likely that more variables other than just treatment and KCCQ-CSS would be appropriate for inclusion in a HFF model.
4. The CS reports that the extended risk equations were a poorer prediction of the observed HHF values in the trial than the reduced risk inclusion. The EAG would have liked to see some investigation into the reason behind this and wonders if the extended models could have provided a better prediction if only predictors considered statistically significant at the 5% level had been included in the model.

The company added that to test the impact of age on HHF, it added the age predictor to the ITT reduced list risk equation and found the latter to be statistically non-significant (reported coefficient estimate of [redacted], and *p*-value of [redacted]). However, the EAG has not seen the outputs of this regression model.

Overall, the EAG is unclear on the appropriateness of using reduced risk equations in the model, and on the plausibility of the parameterised survival models fitted to first HHF KM data. The company did

not investigate including predictors other than treatment in the latter and has only investigated first HHF events.

Table 36. Covariates included in the extended HHF risk equation, estimated with a Poisson model.

Predictors	Estimate	p-value
Intercept	[redacted]	[redacted]
Empagliflozin 10mg (ref: placebo)	[redacted]	[redacted]
Male (ref: female)	[redacted]	[redacted]
CRT (ref: no CRT)	[redacted]	[redacted]
Type II diabetes (ref: no diabetes)	[redacted]	[redacted]
eGFR ≥ 60 (ref: eGFR < 60)	[redacted]	[redacted]
Latin America	[redacted]	[redacted]
North America	[redacted]	[redacted]
Asia	[redacted]	[redacted]
Other region	[redacted]	[redacted]
NT-proBNP (log-scale)	[redacted]	[redacted]
Prior HF (ref: no prior HF)	[redacted]	[redacted]
Time since HF diagnosis (ref: 0-1 years): 1-5 years	[redacted]	[redacted]
Time since HF diagnosis (ref: 0-1 years): 5+ years	[redacted]	[redacted]
NYHA III-IV at baseline (ref: class I-II)	[redacted]	[redacted]
ICD (ref: no ICD)	[redacted]	[redacted]
Prior HF medication (ref: other): ACEI/ARB + BB – IVA	[redacted]	[redacted]
Prior HF medication (ref: other): ARNI + BB - IVA	[redacted]	[redacted]
Updated KCCQ-CSS quartile (ref: Q1 [0, 55.7)): Q2 [55.7, 74)	[redacted]	[redacted]
Updated KCCQ-CSS quartile (ref: Q1 [0, 55.7)): Q3 [74, 88)	[redacted]	[redacted]
Updated KCCQ-CSS quartile (ref: Q1 [0, 55.7)): Q4 [88, 100]	[redacted]	[redacted]
Values in bold represent the variables which would not be considered statistically significant with a p-value of 5%		

4.2.6.5 Treatment discontinuation

In order to estimate time to treatment discontinuation (TTD) in the model, the company fitted parametric survival curves to the TTD KM data from EMPEROR-Preserved. The Weibull, Gompertz, log-logistic, lognormal, and the generalised gamma distributions were fitted to the TTD KM data and assessed for best visual fit; Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) statistics; and clinical plausibility. Similar to HHF outcomes, the company included time-varying KCCQ-CSS and treatment predictors to the selected parametric models.

The company fitted a generalised gamma model as it considered that its closer fit to the tail-end of the KM TTD curve resulted in the more realistic long-term estimates. After treatment discontinuation with empagliflozin, patients were assumed to receive SoC only.

4.2.6.6 EAG critique

During clarification, the EAG raised a concern that the model might reflect an underestimation of time on treatment associated with empagliflozin. Given that clinical expert opinion is that patients (who do not discontinue treatment) stay on empagliflozin for the rest of their lives; the model prediction that patients will stay on treatment for a mean duration of 3.81 years; and a predicted survival for modelled empagliflozin patients is of 8.24 years; the EAG asked at clarification that the company discuss the clinical plausibility of the estimated time on treatment in the model. The CSR for EMPEROR-Preserved reports that 23% discontinued treatment for reasons other than death (which compares to 16% reported in the EMPEROR-Reduced trial).

The company's response at clarification was that even when treatment discontinuation is removed from the model, therefore assuming that patients stay on treatment throughout their lifetime, the deterministic ICER increased only from £14,429 (base case) to £15,126. Therefore, the EAG is satisfied that TTD is not a driver of the economic results.

4.2.6.7 Mortality

The company fitted parametric survival curves to all-cause mortality and to CV-related mortality KM data from EMPEROR-Preserved, separately. The Weibull, Gompertz, log-logistic, lognormal, and the generalised gamma distributions were fitted to the data and assessed for best visual fit; AIC and BIC statistics; and clinical plausibility. Similar to TTD outcomes, the company included time-varying KCCQ-CSS and treatment predictors to the selected parametric models for each survival function.

The economic model uses different probabilities for CV-related and non-CV related deaths (which differed by KCCQ-CSS state and by treatment arm). In order to estimate the probability of patients having a non-CV death the company subtracted the monthly probability of CV-death from the probability of all-cause death in each cycle of the model.

To ensure that the probability of non-CV death was never higher than the corresponding age-and sex-specific probability for the general UK population (taken from the national UK life tables), the maximum probability of non-CV death between EMPEROR-Preserved and the life tables was taken in every cycle of the model. UK life tables were adjusted to exclude CV-related deaths to avoid double counting.

4.2.6.7.1 Overall survival

A total of 422 (14.1%) patients died in the empagliflozin arm of EMPRERORPreserved, while 427 (14.3%) patients died in the placebo arm (Figure 11, HR of 1.00; 95% confidence interval of 0.87 to 1.15).

The company assessed all-cause mortality KM data for the validity of the proportional hazards (PH) assumption and the best fit between independent and jointly fitted curves. The company decided to fit a joint Weibull (PH) model. The company also described an iterative process for selecting and removing non-statistically significant variables from the Weibull risk equations and decided to use the reduced risk equation with only KCCQ-CSS quartiles as predictors of survival (Table 37). The company excluded treatment effect as a predictor of survival from the risk equation because treatment effect was not a statically significant predictor of all-cause mortality.

As with the HHF risk equations, the company included an extended risk equation (as opposed to the reduced base case equation which only considered treatment arm and KCCQ-CSS quartile) in the model, with statistical significance defined at the 10% level.

Figure 11. Observed OS data in EMPEROR-Preserved (reproduced from Figure 19, CS).

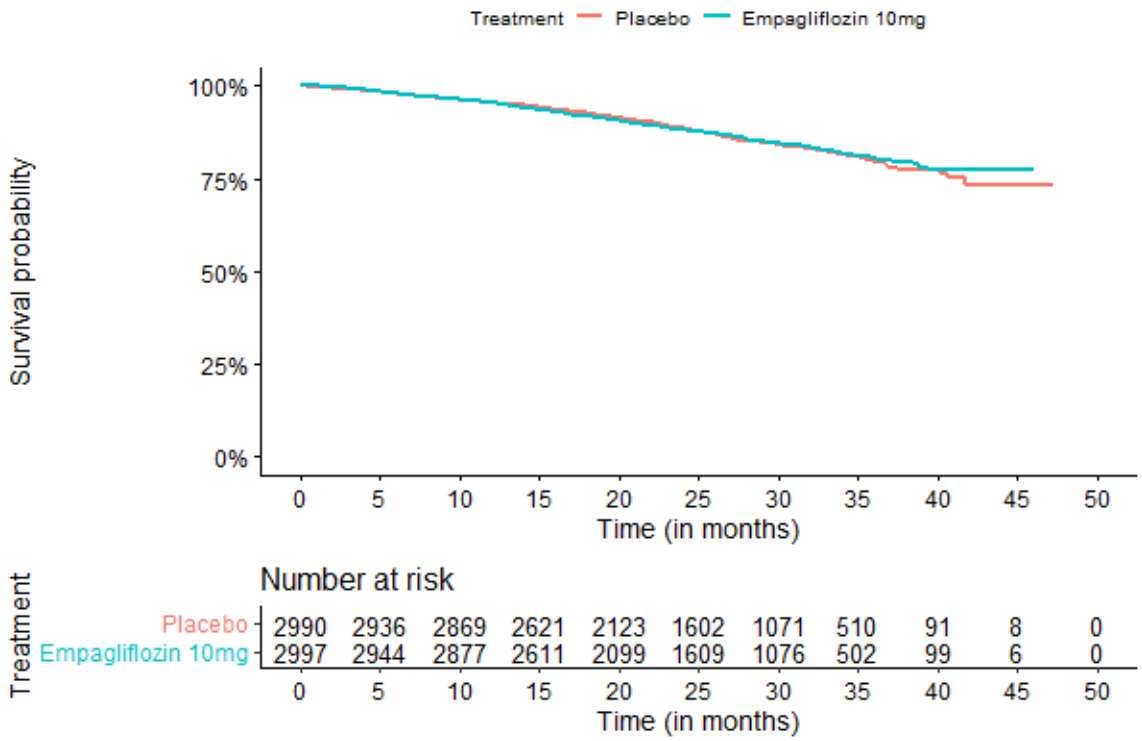


Table 37. Weibull regression for OS, ITT population from EMPEROR-Preserved

Covariate	Coefficient	p-value
Shape	[redacted]	[redacted]
Scale	[redacted]	[redacted]
Treatment effect Empagliflozin 10 mg (reference placebo)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 2 vs Quartile 1)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 3 vs Quartile 1)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 4 vs Quartile 1)	[redacted]	[redacted]

4.2.6.7.2 CV-related death

A total of 219 (7.3%) patients had a CV-related death in the empagliflozin arm of EMPEROR-Preserved, while 244 (8.2%) patients experienced a CV-related death in the placebo arm (Figure 12, HR 0.91, 95% CI: 0.76 to 1.09). Overall, CV-related deaths represented 51% of all deaths in EMPEROR-Preserved.

The company fitted a joint Weibull model to the empagliflozin and SoC arms of EMPEROR-Preserved data. The company justified its approach based on, "*the lack of strong evidence against the*

assumption of proportional hazards; plausibility of cause-specific extrapolations; and face validity with regards to higher survival estimates compared to the all-cause mortality Weibull model."

The company reported the same process as that undertaken for HFF and all-cause mortality, consisting of an iterative process for selecting and removing non-statistically significant variables from the Weibull risk equations. The company also decided to use the reduced risk equation with treatment and KCCQ-CSS quartiles as predictors of survival. The company added that even though the treatment effect for CV mortality was not statistically significant (as observed for all-cause mortality), this was added as a predictor in the CV-model because a numerical difference between the two comparators is clinically plausible for this outcome (Table 38).

Figure 12. Observed CV-related mortality data in EMPEROR-Preserved reproduced from Figure 20, CS).

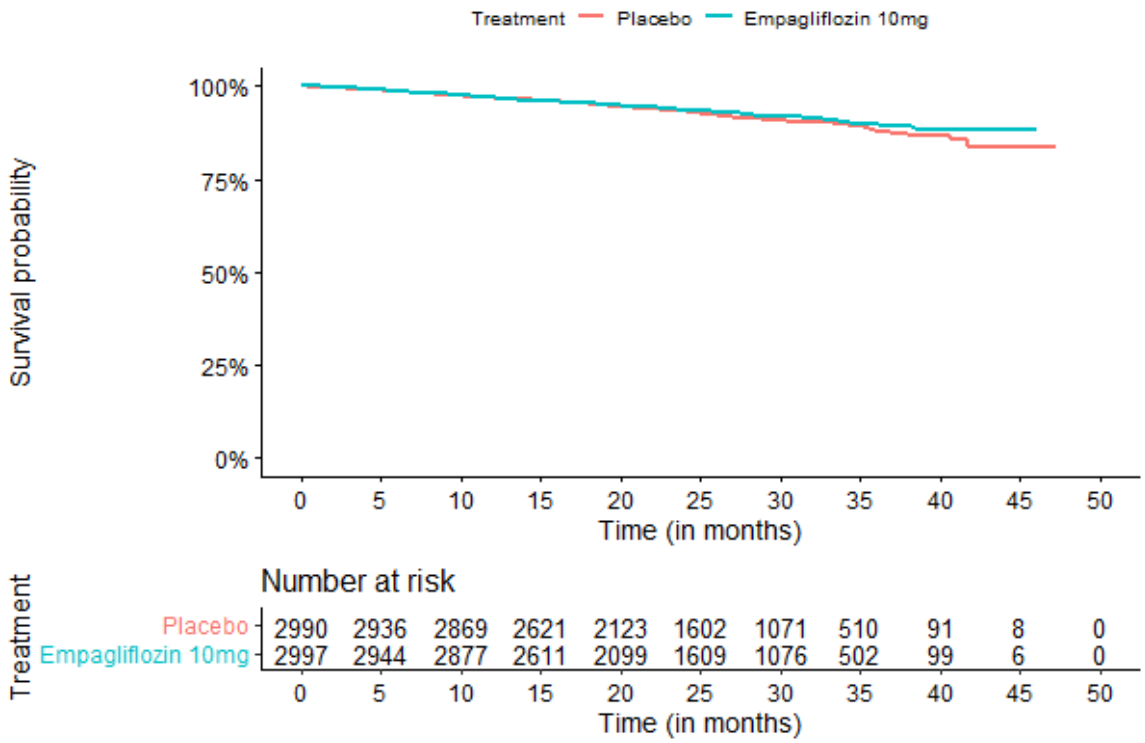


Table 38. Weibull regression for CV-mortality, ITT population from EMPEROR-Preserved

Covariate	Coefficient	p-value
Shape	[redacted]	[redacted]
Scale	[redacted]	[redacted]
Treatment effect Empagliflozin 10 mg (reference placebo)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 2 vs Quartile 1)	[redacted]	[redacted]

Covariate	Coefficient	p-value
KCCQ-CSS: (Quartile 3 vs Quartile 1)	[redacted]	[redacted]
KCCQ-CSS: (Quartile 4 vs Quartile 1)	[redacted]	[redacted]

4.2.6.8 EAG critique

The EAG considers that there is some uncertainty on the company's long-term overall survival (OS) extrapolations. Figure 13 shows the three-best fitting models to OS KM data from EMPEROR-Preserved: the Weibull (company's base case); Gompertz and generalised gamma models.

Based on 10-year KM data from an external literature source found by the EAG (Eriksson *et al.*), the 5-year and 10-year survival observed for patients with HFmrEF/HFpEF was about 60% and 35%, respectively.[17] The population in the Eriksson study was broadly the same age as the population in EMPEROR-Preserved, however and had less severe NYHA stage at baseline, and had a considerably lower proportion of patients with diabetes (20% vs 50%), which the EAG expects to lead to better survival outcomes than patients in EMPEROR-Preserved.

In comparison to Figure 13, the base case Weibull distribution chosen by the company seems to align with the Eriksson outcomes, with over 30% of patients alive at 10 years. Nonetheless, the EAG expects patients in EMPEROR-Preserved to have a higher mortality than patients in Eriksson given the NYHA staging and the presence of diabetes at baseline in the former. The generalised gamma provides more conservative long-term predictions, with 20% of patients being alive at 10 years. The Gompertz curve predicts the most conservative survival estimation at 5 years (about 55%) and is likely to underestimate survival at 10 years (0% of patients are alive).

Crucially, the EAG disagrees with the company's approach of including KKCQ-CSS as predictors of survival in the model given that these generate an indirect survival benefit for empagliflozin in the model. Given that the probability of patients dying is different in every KCCQ-CSS state of the model and that patients in the empagliflozin arm of the model have a higher probability of remaining in the better KQCC-CSS states over time compared with SoC patients, the former also experience a lower probability of death.

Given the KM OS data reported in Figure 11, which shows that survival was similar in the two arms of the EMPEROR-Preserved trial, the EAG considers that including a treatment benefit associated

with empagliflozin is unsubstantiated. Therefore, the EAG requested that the company included a scenario analysis in the model where both the direct and indirect treatment effect of empagliflozin on overall survival was removed from the model. When it is assumed that overall survival is the same across treatment arms but that empagliflozin has an effect on CV-related mortality, the ICER increases from £14,429 to £21,104 per QALY gained.

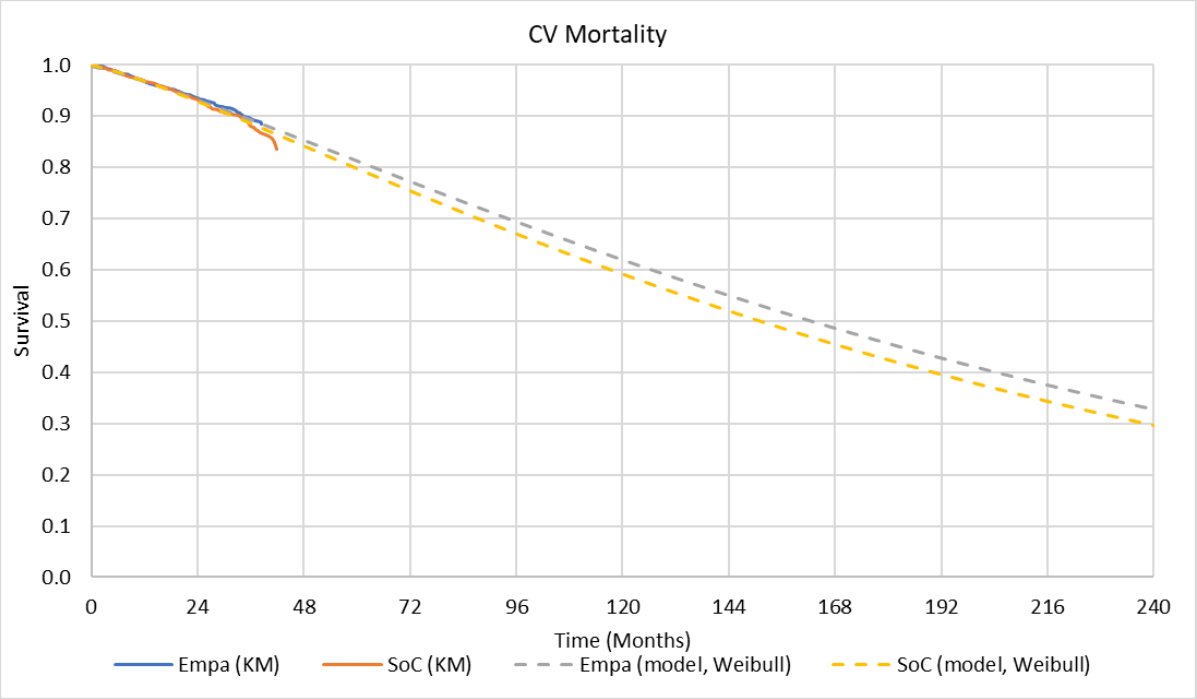
Unfortunately, the company only provided this scenario for the Weibull re-fitted model. Using the generalised gamma in the model, which provides more conservative long-term survival predictions, increases the company's base case ICER to £14,473, whereas using the Gompertz increased the ICER to £17,553. Given that the company did not provide the generalised game and the Gompertz risk equations re-refitted to exclude KCCQ-CSS as a predictor, the EAG cannot present the impact of using these models when there is no survival benefit assumed for empagliflozin. However, the EAG can infer that the £21,104 ICER would be marginally higher if the gamma model was used, and considerably higher if the Gompertz model were used.

Figure 13. Weibull, generalised gamma and Gompertz models fitted jointly to EMPEROR-Preserved OS KM data.

The EAG has some concerns regarding the company's inclusion of a treatment effect in the risk equations to estimate CV mortality in the model. Even though the EAG's clinical experts agreed with the clinical plausibility of empagliflozin being associated with a survival benefit on CV-related death,

the EMPEROR-Preserved trial showed a non-statistically significant effect on this outcome. Furthermore, given the company's decision to include a treatment effect in the CV-mortality risk equations, the decision to do so through a PH model is problematic. Considering the shape of the KM curves observed in Figure 12, it is implausible that a constant treatment effect would be observed throughout the trial (and extrapolated) period (Figure 14), and by assuming PH, the company is overestimating the relative treatment effect associated with empagliflozin.

Figure 14. Weibull models fitted jointly to EMPEROR-Preserved OS KM data



During clarification, the EAG asked that the company included an option in the model to remove the direct and indirect effect (separately) of empagliflozin on CV-related mortality (as requested for all-cause mortality). When it is assumed that overall survival and CV-related mortality are the same in both treatment arms (witch by default means that non-CV deaths are also the same in both treatment arms), the final ICER increases to £26,422 per QALY gained.

Overall, the EAG notes that the OS curves from EMPEROR-Preserved (both for overall and CV survival) do not provide sufficient evidence to substantiate empagliflozin having an impact on patients' survival compared to SoC patients. However, clinical expert opinion consistently reported the plausibility of empagliflozin reducing patients' CV-related mortality. Therefore, the EAG considers that it is crucial that the committee validates the following assumptions:

1. Is empagliflozin likely to reduce CV mortality compared to SoC?
    a. If the answer to the above question is yes, then the EAG notes that there are two possible implications for non-CV deaths in this population:
        i. Either the CV deaths prevented in the empagliflozin arm do not translate into a reduction in overall mortality for these patients - given the age of the HFmrEF/HFpEF population, and the presence of several comorbidities, the proportion of patients who don't die of a CV cause die in a similar time frame of a non-CV cause, therefore suggesting that there will be more non-CV deaths for empagliflozin than SoC patients – in this case, it is crucial that the cost of a CV and non-CV death is incorporated into the economic analysis, to assess the cost-effectiveness of "replacing" an equal number of CV deaths with non-CV deaths for empagliflozin patients. Or;
        ii. The CV deaths prevented in the empagliflozin arm translate into a reduction in overall mortality for these patients vs SoC patients – the proportion of patients who don't die of a CV cause end up dying much later of a non-CV cause, and a similar proportion of empagliflozin and SoC patients die of non-CV causes in the shorter term.
    b. If the answer to question 1 is no, then there is no difference in CV or non-CV deaths for empagliflozin and SoC patients.

In Section 6, the EAG provides scenarios analysis assuming no survival benefit on all-cause mortality associated with empagliflozin, combined with two alternative scenarios where there is a benefit assumed for CV-related death; or there is no impact on the latter. Nonetheless, the EAG did not have time to investigate the impact of including the cost of non-CV deaths in the model and recommends that the company carefully considers including these events in the model, depending on the committee's conclusion on the impact of empagliflozin on survival.

4.2.7 Adverse events

The adverse events included in the model can be found in Table 35 of the CS. These are further discussed in Section 4.

4.2.8 *Health-related quality of life*

EQ-5D-5L data were collected in the EMPEROR-Preserved trial at baseline; 12 weeks; 32 weeks; 52 weeks; 100 weeks; and 148 weeks following randomisation, as well as at treatment discontinuation. Patients were also followed up 30 days following completion of the treatment period. In line with NICE guidance, the company mapped the EQ-5D-5L responses onto the EQ-5D-3L value set using the crosswalk mapping function developed by Hernández Alava *et al*. [18]

Utility scores were analysed using mixed-effects linear regression using all available EQ-5D measurements across all visits. The final model incorporated time-varying predictors such as KCCQ-CSS state and whether a patient had a HHF in the last 0-1 months, 1-2 months, 2-4 months, and 4-12 months.

For HHF events, the reference group had no HHF events to date, and patients were classified back into the reference group once a year had passed from hospitalisation. Therefore, the company assumed that HHF impacted patients' quality of life for 12 months after the event.

The annual disutility associated with HHF in the model was estimated as [redacted] per event. This was calculated by multiplying the coefficients in Table 39 (estimated in the mixed-effects model) for time since HHF by the respective period of time and adding these together [redacted]). The corresponding coefficients in TA773 were the following: [redacted]

The mixed-effects model also included AEs as predictors (urinary tract infections, mycotic genital infections, acute renal failure, hepatic injury, volume depletion, hypotension, hypoglycaemic event, and bone fractures), where patients were returned to the reference group one month after experiencing the event. The coefficients for each AE were applied as monthly disutilities for each event as per Table 40. The company sourced the disutility for genital mycotic infection and hypoglycaemic events from external literature sources.

Unlike the risk equations for HHF and CV-related survival outcomes, treatment was not included as a predictor in the final model as it was found to be not statistically significant. The other predictors of patients' quality of life (in addition to HHF and AEs) included in the final regression are reported in Table 41. The EAG found some discrepancies between the predictors reported in Table 36 of the CS and the predictors included in the model, therefore, in Table 41 the EAG presents the predictors included in the model.

Table 39. Final QoL regression – HHF coefficients used in the model

Covariate	Coefficient	SE	p-value
HHF: <1 month	[redacted]	[redacted]	[redacted]
HHF: 1 to <2 months	[redacted]	[redacted]	[redacted]
HHF: 2 to <4 months	[redacted]	[redacted]	[redacted]
HHF: 4 to <12 months	[redacted]	[redacted]	[redacted]

Table 40. Final QoL regression used in the model - AEs coefficients used in the model

Covariate	Coefficient	SE	p-value	Disutility used in the model
Urinary tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Genital Mycotic Infection	[redacted]	[redacted]	[redacted]	[redacted]
Acute renal failure	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic injury	[redacted]	[redacted]	[redacted]	[redacted]
Volume depletion	[redacted]	[redacted]	[redacted]	[redacted]
Hypotension	[redacted]	[redacted]	[redacted]	[redacted]
Hypoglycaemic event	[redacted]	[redacted]	[redacted]	[redacted]
Bone fracture	[redacted]	[redacted]	[redacted]	[redacted]
*not statistically significant at a 5% significance level				

Table 41. Final QoL regression used in the model (taken form the company's model)

Covariate	Coefficient	SE	p-value
Intercept	[redacted]	[redacted]	[redacted]
Demographics			
Age ≥65	[redacted]	[redacted]	[redacted]
Male (ref: Female)	[redacted]	[redacted]	[redacted]
Region			
Latin America	[redacted]	[redacted]	[redacted]
North America	[redacted]	[redacted]	[redacted]
Asia	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Baseline KCCQ quartile (ref: [88.02, 100])			
KCCQ-CSS: 75 to <90 (Quartile 3)	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 55 to <75 (Quartile 2)	[redacted]	[redacted]	[redacted]
KCCQ-CSS: 0 to 55 (Quartile 1)	[redacted]	[redacted]	[redacted]

Covariate	Coefficient	SE	p-value
Prior atrial fibrillation or flutter	[redacted]	[redacted]	[redacted]
Baseline eGFR higher or equal to 60	[redacted]	[redacted]	[redacted]

The final estimated utility values for each KCCQ-CSS quartiles are presented in Table 42. The utility estimated in the mixed-effects model for quartile 4 (the least severe quartile) was higher ([redacted]) than the utility seen in the UK general population reported by Sullivan *et al.*[19] for the age group of 70-79 years. Therefore, the company used the utility reported in Sullivan (0.7230) for quartile 4 and adjusted the utility values for the 1–3 quartile states based on the relative difference between the estimated utility for quartile 4 and the value reported in Sullivan.

Table 42. Utility values used in the company's base case

KCCQ-CSS health state	Value estimated	Value used in the model
Quartile 1	[redacted]	[redacted]
Quartile 2	[redacted]	[redacted]
Quartile 3	[redacted]	[redacted]
Quartile 4	[redacted]	0.7230

4.2.8.1 EAG critique

During clarification, the company provided the mapped EQ-5D-3L data (from the EQ-5D-5L data from EMPEROR-Preserved) on the mean baseline; change from baseline; and number of observations. The results are reported in Table 43 and show that [redacted]

The company also reported changes in utility values from baseline in Table 19 of the clarification document; however, the EAG considers that the latter lack face validity, thus the EAG estimated

these (see Table 43) but cannot be sure of the statistical significance of these changes. [redacted] Overall, the EAG considers that the QALY gain over the initial 2 years of the model (Table 44) broadly reflects the gains seen in EMPEROR-Preserved up to year 2. However, the model produces a total QALY gain of [redacted], which the EAG is concerned might be an overestimation. The trend in the difference of patients' mean utility observed in EMPEROR-Preserved (Table 43) suggests that the incremental utility gain seen in the empagliflozin arm [redacted] from week 100 to week 148 of the trial, whereas the trend in the model shows an [redacted] of [redacted] in the QALY gains associated with empagliflozin every year of the model.

The majority of the QALY gain in the model comes from the additional time spent by empagliflozin patients in quartile 4 and 3 of the KCCQ-CSS state when compared to SoC patients. Another considerable driver of the QALY gains in the model comes from the reduction in HHF for empagliflozin patients when compared to SoC patients.

Table 43. EQ-5D-3L scores in EMPEROR-Preserved

Time	Empagliflozin 10 mg		Placebo		Difference in mean utility	Change from baseline empagliflozin mean (SD)	Change from baseline SoC mean (SD)	Difference in the change from baseline
	N	Mean	N	Mean				
Baseline	[redacted]							
week 12	[redacted]							
week 32	[redacted]							
week 52	[redacted]							
week 100	[redacted]							
week 148	[redacted]							
[redacted]								

Table 44. QALY gain in the economic model

	Incremental
At week 52 (year 1)	[redacted]
At week 104 (year 2)	[redacted]
At week 156 (year 3)	[redacted]
At week 208 (year 4)	[redacted]
At week 260 (year 5)	[redacted]
At week 312 (year 6)	[redacted]

The EAG considers that the [redacted] disutility associated with HHF in the model is overestimated as it implies that every HHF event would impact patients' quality of life for 1 year after the event, which clinical expert opinion provided to the EAG indicated reflects an overestimation. The experts indicated that the average length of stay in the hospital for HHF for patients with preserved EF is 11 days (which is validated by the 11-day mean stay for HHF in EMPEROR-Preserved). Subsequently, one expert indicated that a reasonable assumption is that 1 day in hospital impacts patients' quality of life for 1 week after discharge. The other clinical expert indicated that 6 months of impact (as a maximum) could also be plausible after discharge. Therefore, the EAG requested that the company conducted two alternative scenario analyses where:

1. It was assumed that HHF events impact patients' QoL for 2.75 months after discharge (corresponding to an impact of 11 weeks after being hospitalised for 11 days).
2. It was assumed that HHF events impact patients' QoL for 6 months after discharge.

When the impact of HHF is assumed to last 2.75 months, the HHF disutility is -[redacted], and when it lasts for 6 months, it is [redacted]. The ICER for the first scenario increases to £17,912 per QALY gained and to £16,511 for the second scenario (compared to the base case ICER of £14,429) . The EAG considers these analyses to provide a more plausible range than assuming every HHF event impacts for 1 year after hospitalisation, therefore presents results using these scenarios in Section 6.

During clarification the EAG also requested that the company included age-related utility decrements throughout the model time horizon using the algorithm published by Ara and Brazier 2010, which the company supplied as a scenario analysis. The EAG considers that the age-related decrements should be use used in the base case results, and therefore reports the results of these analysis in Section 6.

Overall, even though the EAG considers that utility gains are likely to be overestimated in the model, the EAG notes that when the EAG's preferred assumptions with regards to survival; KCCQ-CSS scores; and utilities are used in the model (as discussed in Section 6.2), the QALY becomes more aligned with the trial results.

During clarification the EAG also noted that the final KCCQ-CSS utility values (unadjusted for age) reflected a higher quality of life in EMPEROR-Reduced patients, when compared to EMPEROR-Preserved patients. The EAG noted that the disutility associated with HHF estimated from the EMPEROR-Reduced population was also lower compared with the EMPEROR-Preserved population.

The company commented that this difference is in line with other differences noted in terms of the baseline characteristics of the two populations in the trials. In particular, the mean age in EMPEROR-Preserved is 71 years, and for EMPEROR-Reduced is 67. Furthermore, the company noted that the average presence atrial fibrillation as a comorbidity in EMPEROR-Reduced was 35-37% vs 50-51% in EMPEROR-Preserved, with hypertension in 72% versus 90% of the trial populations, respectively, both of which are associated with a higher risk of adverse events in the EMPEROR-Preserved population. The EAG notes that the baseline utility values in EMPEROR-Reduced were [redacted] in for SoC and empagliflozin, respectively, which compare to [redacted] in EMPEROR-Preserved.

Disutility associated with AEs

Even though the company used disutilty values estimated from its regression analysis that were not statistically significant (see Table 40), and despite some inconsistencies in the sources for the disutilities used, the scenario analysis conducted by the EAG showed that removing the non-statistically significant events reported in Table 40 from the model had a negligible impact on the ICER.

4.2.9 Resource use and costs

4.2.9.1 Treatment and comparator costs

The intervention included in the economic model is empagliflozin formulated as a 10mg tablet taken once a day, in addition to SoC. The list price for empagliflozin is £36.59 for a pack of 28 pills, amounting to a daily cost of £1.31 and a monthly cost of £39.78. Standard of care was modelled as a basket of drugs used in first-line heart failure care, as discussed in Section 4.2.3.

4.2.9.1 Disease management costs

Disease management costs included GP and cardiologist visits, and A&E referrals. Resource use was based on data from the Clinical Practice Research Datalink, as reported by McMurray *et al.* (2018).[20] Unit costs were taken from PSSRU Unit Costs of Health and Social Care[21] and the Schedule of NHS Costs 2019/2020. The company assumed that GP visits were based on patient contact lasting 9.22 minutes and that cardiologist visits were consultant-led, face to face follow-up appointments. The cost of an A&E referral was based on a weighted mean derived from national average unit costs and number of finished consultant episodes (FCEs) for non-admitted emergency medicine. All disease management unit costs were inflated to 2021 by applying the consumer price health inflation factor

from Eurostat,[22] and are reported in Table 50 of the CS. Total cost was £39.62 monthly, for every KCCQ-CSS state.

4.2.9.2 Hospitalisation costs

The acute costs of HHF were based on NHS reference costs for 2019/2020 for non-elective long inpatient stay, computed as the weighted average of reference costs for healthcare resource group (HRG) codes (described in Table 45) and the number of FCEs. The total cost of a HHF event in the model was £3,259.

Table 45. Hospitalisation for heart failure unit costs

Reference cost code and description	Unit cost
Heart Failure or Shock, with CC Score 14+ (EB03A) – Non-Elective (Long Stay)	£4,076
Heart Failure or Shock, with CC Score 11-13 (EB03B) – Non-Elective (Long Stay)	£3,191
Heart Failure or Shock, with CC Score 8-10 (EB03C) – Non-Elective (Long Stay)	£2,634
Heart Failure or Shock, with CC Score 4-7 (EB03D) – Non-Elective (Long Stay)	£2,238
Heart Failure or Shock, with CC Score 0-3 (EB03E) – Non-Elective (Long Stay)	£2,062

4.2.9.3 EAG critique

In TA773, the EAG noted that a better approach to use the appropriate cost year in the analysis would have been to use the 2019/2020 costs sourced from the updated NHS Cost Schedule, instead of using older cost estimates used by the company (2018/2019) and inflating these with the consumer price health inflation factor from Eurostat.[22]

The same issue is applicable for the current submission, given that the updated NHS Cost Schedule 2020/2021 is available. The company's unit costs sourced from the NHS Cost Schedule are therefore outdated by 1 year. The EAG appreciates that the company has used the consumer price health inflation factor from Eurostat to bring prices up to 2021, however, again notes that using the relevant NHS costs for the right year would have been more accurately representative of the costs

incurred in the UK. Furthermore, the EAG notes that inflation index could have been updated to reflect the current cost year (2022), instead of 2021. Overall, the EAG does not expect this to have a considerable impact on the final ICER.

During clarification, the EAG noted that the mean duration of HHF in EMPEROR-Preserved was 11 days; with a median of 8 days; and Q3 of 13 days. The EAG was therefore, concerned that the company's use of the more severe cost code (EB03A), associated with a 53-day long hospitalisation, overestimated HHF costs in comparison to EMPEROR-Preserved. Given the less severe cost code (EB03E) is associated with 13 days in hospital, the EAG asked that the company conducted a scenario analysis in the model where the cost of £2,062 was used to calculate the cost of all HHFs events in the model. The company used the £2,134 (£2,062 inflated to 2021 cost year) to run the scenario analysis and the deterministic ICER increased from £14,429 to £15,215.

Given the length of stay observed in EMPEROR-Preserved, the EAG's preference is to use the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.

4.2.9.4 CV-related mortality

Similar to TA773, the cost of CV death was based on the regression analysis presented in Alva *et al.* which estimated the added inpatient costs for type-2 diabetes mellitus (T2DM) complications, during the UK Prospective Diabetes Study post-trial monitoring period from 1997 to 2007, and used hospitalisation records for patients in England (n=2,791).[23]

The regression analysis reported coefficients for the expected cost impact of T2DM complications, which included fatal myocardial infarction (MI), fatal ischaemic heart disease (IHD), and fatal stroke, as well as age and gender, on inpatient hospitalisation costs. The company used these coefficients to estimate the costs of fatal MI; IHD and stroke. Costs were estimated separately for males and females; aged <65 years and ≥65 years, respectively.

The company then used the percentage of males/females and the percentage aged <65 or ≥65 years from EMPEROR-Preserved and derived a weighted average cost for each event. A simple average was then taken across the cost of the three fatal events to derive the final cost of CV death for the model of £4,295.

4.2.9.5 EAG critique

As in TA733, the EAG disagrees with the use of the chosen estimates from the Alva paper as these relate to the added costs on hospitalisations due to T2DM complications. Therefore, during clarification the EAG asked that the company used the alternative estimates provided in Table 3 of the Alva study, which reported the absolute cost of the events. The company conducted this analysis and the cost of CV death changed from £4,295 to £3,809 (after inflation).

During clarification, the EAG also asked that the company estimated the weighted costs of CV-deaths by the proportion of events leading to CV death observed in EMPEROR-Preserved (Table 11.1.2.4.2:1, page 120 of the CSR). The latter showed that [redacted] of deaths in EMPEROR-Preserved were sudden cardiac deaths and so the company included two options in the model: a conservative approach where the cost of sudden cardiac death was zero, decreasing the overall costs of CV death to £1,452; and a second option where the total HRG costs for cardiac arrest of £1,632 was used, leading to overall costs of CV death of £2,345.

With the EAG's preferred unit costs from Alva, the alternative total cost of CV deaths are: £3,809 (company's base case); £1,452 (with the cost of sudden death assumed to be zero); or £2,345 (with the cost of sudden death assumed to be £1,632).

The EAG's preferred cost for CV death in the model is £1,452 as it represents the most conservative estimate, however the EAG reports the results of using a cost for sudden cardiac arrest in the scenario analysis reported in Section 6.

4.2.9.1 Adverse event costs

The unit costs for outpatient visits were taken from the PSSRU unit costs of health and social care, while inpatient costs were taken from the National Schedule of NHS Costs for 2018/19 and then inflated to 2021 costs using the consumer price health inflation factor from Eurostat.[22]

Similar to the HHF and disease management costs, the EAG's preference would have been for the company to use the 2020/2021 more up to date costs sourced from the NHS Cost Schedule, instead of using older cost estimates and inflating these.

5 Cost effectiveness results

5.1 Company's cost effectiveness results

The results of the company's base case deterministic analysis are presented in Table 46. In the base case analysis, empagliflozin + SoC generates [redacted] incremental QALYs and incremental costs of [redacted] over SoC alone, resulting in an ICER of £14,429 per QALY gained.

Table 46. Company's base case results

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Empagliflozin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14,429
Standard of Care	[redacted]	[redacted]	[redacted]	-	-	-	-

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year;

The company performed probabilistic sensitivity analysis (PSA) to assess the joint parameter uncertainty around the base case results, using 1,000 PSA iterations. Table 47 shows that the company's PSA ICER of £14,564 per QALY gained is similar to the company's deterministic ICER. The probability of empagliflozin being cost effective at the £30,000 threshold is [redacted] (as per Figure 15).

Table 47. Company's mean PSA results

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Empagliflozin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£14, 564
Standard of Care	[redacted]	[redacted]	[redacted]	-	-	-	-

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year;

Figure 15. Cost-effectiveness acceptability curve

Figure 16. Cost-effectiveness plane

5.2 Company's sensitivity analyses

The company's sensitivity analysis can be found in Section B.3.10.3 of the CS, and subgroup analysis can be found in Section B.3.11 of the same document. The company's conclusion is that the most influential parameter in the model is the treatment effect of empagliflozin associated with HHF. When this parameter was set to zero, the ICER increased from £14,429 to £21,339 per QALY gained.

Other highly impactful parameters identified by the company were the disutility for HHF, the treatment effect associated with empagliflozin for CV mortality, inclusion of treatment discontinuation for empagliflozin, and cost of treatment for HHF.

6 Additional economic analysis undertaken by the EAG

6.1 Exploratory and sensitivity analyses undertaken by the EAG

The scenario analyses undertaken by the EAG are explained throughout Section 4 of the report. Results of the exploratory analyses conducted are reported in Table 48. The following analyses were conducted:

1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.
2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality.
3. Assuming that empagliflozin does not have an effect on overall survival (directly or indirectly) or on CV mortality.
4. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.
5. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.
6. Using the age-related decrements from Ara and Brazier 2010.
7. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.
8. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).
9. Using a unit cost for CV death in the model of £2,345 (with the cost of sudden death assumed to be £1,632).

Results in Table 48 show that the key drivers of the economic results are the assumptions made around the impact of empagliflozin on survival. When it is assumed that overall survival is the same across treatment arms but that empagliflozin has an effect on CV-related mortality, the ICER increases from £14,429 to [redacted] per QALY gained. When it is assumed that overall survival and CV-related mortality are the same in both treatment arms (witch by default means that non-CV deaths are also the same in both treatment arms), the final ICER increases to [redacted] per QALY gained. To aid the interpretation of the EAG's scenarios on mortality, the EAG provided the accompanying survival curves and a small description of each scenario in Appendix 8.1.

The second highest driver of the economic results is using the unadjusted TPs between KCCQ-CSS scores, followed by the assumption around the duration of the impact of HHF on patients' quality of

life. Some of the scenarios conducted by the EAG have a small impact when run in isolation, however, have a considerable impact on the ICER when combined (as described in Section 6.2).

Table 48. Results of the EAG's scenario analyses

	Results per patient	Intervention	Comparator	Incremental value
0	**Corrected company base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	£14,429
1	Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
2	Assuming that empagliflozin does not have an effect on overall survival (directly or indirectly) but has an effect on CV mortality.			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
3	Assuming that empagliflozin does not have an effect on overall survival (directly or indirectly) and does not have an effect on CV mortality			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
4	Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
5	Assuming HHF impact patients' QoL for 6 months year after hospitalisation.			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
6	Using the age-related decrements from Ara and Brazier 2010.			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]
7	Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]

8	Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)			[redacted]
9	Using a unit cost for CV death in the model of £2,345 (with the cost of sudden death assumed to be £1,632).			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

6.2 EAG preferred assumptions

The EAG's preferred assumptions are:

9. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.
10. Assuming that empagliflozin does not have an effect on overall survival.
11. Using the age-related decrements from Ara and Brazier 2010.
12. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.
13. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).

In addition to the assumptions listed above, the EAG conducted one set or alternative scenarios:

14. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.
15. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.

Finally, the EAG conducted the alternative combined scenarios described above with the following:

16. Assuming that empagliflozin does not have an effect on CV related deaths (directly or indirectly), or on overall survival.

When empagliflozin is assumed to have an impact on CV mortality (but not on overall survival), the cumulative EAG-preferred assumptions result in a final ICER of [redacted], depending on the duration of the impact of a HHF event on patients' quality of life (Table 49). When empagliflozin is assumed to not have an impact on CV mortality (or overall survival), the cumulative

EAG-preferred assumptions result in a final ICER of [redacted], depending on the duration of the impact of a HHF event on patients' quality of life (Table 49). Given how similar the deterministic base case results are to the probabilistic results, the EAG did not present probabilistic results of the exploratory analysis conducted.

The key driver of the model remains the assumption made around the impact of empagliflozin on mortality. In the combined EAG ICER, the second biggest driver of results is the assumption made for the duration of the impact of an HHF event on patients' quality of life, followed by using the observed TPs between KKCQ-CSS (as opposed to the ones with imputed data).

The EAG caveats again its scenario on mortality assuming that empagliflozin has no impact on overall mortality but reduces patients' CV mortality. This scenario needs further investigation on its implications for patients' non-CV mortality given that it indirectly implies an increase in non-CV deaths for empagliflozin patients, but these have not been considered in the economic model by the company.

Overall, when the EAG's preferred assumptions with regards to survival; KCCQ-CSS scores; and utilities are used in the model, the total HHF events and deaths (both overall and CV-related) remain overestimated (as is the case in the company's base case); however, the incremental events estimated are closer to those observed in EMPEROR-Preserved at 3 years, even if there is an underestimation in the first model cycles (see Appendix 8.2 for the comparison of observed vs estimated outcomes in the EAG's scenarios). The same can be observed for the trend in the utility gains generated in the model. The EAG notes that when it is assumed that empagliflozin does not have an impact on overall survival but does have an impact on CV-related death, the model provides a good prediction of both incremental CV and all-cause deaths in the trial at 3.5 years (with an underestimation at year 1 and year 2). However, when it is assumed that empagliflozin does not have a benefit associated with reducing CV deaths, the model predictions lack face validity when compared to trial outcomes (Table 53, Appendix 8.2).

Finally, as discussed in Section 4.2.6, the company included a scenario analysis in the model which assumed that the TPs between KCCQ-CSS quartiles for patients on treatment are the same as those for patients off treatment from month 9 onwards in the model, which equates to the proportion of patients in each KCCQ-CSS state being the same in the empagliflozin and SoC arms of the model approximately 4 to 5 years after treatment initiation in the model. This scenario is likely to be overly

pessimistic as it assumed that after 4 years EMPAGLIFLOZIÁN does not impact patients' KCCQ-CSS anymore, even if patients are still on treatment. When the EAG ran the company's scenario analysis on the EAG's cumulative ICERs, the impact on the final ICER was considerable, ranging from [redacted], depending on the assumptions used (Table 50).

Table 49. EAG's preferred model assumptions

Preferred assumption	Section in EAG report	Cumulative ICER (£/QALY)
Company base case	Section 5	£14,429
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.	Section 4.2.61 and 4.2.6.2	[redacted]
2. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 3. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality.	Section 4.2.6.7	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010.	Section 4.2.8	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.	Section 4.3.1.2 and Section 4.3.1.3	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).	Section 4.3.1.4 and Section 4.3.1.5.	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden	Section 4.2.8	[redacted]

death assumed to be zero). 6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.		
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.	Section 4.2.8	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation. 8. Assuming that empagliflozin does not have an effect on CV mortality.	Section 4.2.6.7	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation. 8. Assuming that empagliflozin does not have an effect CV mortality.	Section 4.2.6.7	[redacted]

Abbreviations: EAG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 50. EAG's preferred model assumptions when assuming that the TPs between KCCQ-CSS quartiles for patients on treatment are the same as those for patients off treatment from month 9 onwards in the model

Preferred assumption	Section in EAG report	Cumulative ICER (£/QALY)
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.	Section 4.2.8	[redacted]

5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.		
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.	Section 4.2.8	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation. 8. Assuming that empagliflozin does not have an effect on CV mortality.	Section 4.2.6.7	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation. 8. Assuming that empagliflozin does not have an effect CV mortality.	Section 4.2.6.7	[redacted]
Abbreviations: EAG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year		

The EAG also produced a list of issues requiring additional clarification or further analysis from the company. These have been discussed in detail throughout Section 4 and are listed below:

1. The EAG remains unclear if and why KCCQ-CSS data beyond 52 weeks in the model were not collected and asks that the company clarifies this. Crucially, the EAG asks that the company provides the distribution of the missing KCCQ-CSS observations, with a description of when end of treatment occurred for these patients. These data should help clarifying the number of missing observations and when these occurred (and thus for how many timepoints the LOCF imputed values were used).

2. The EAG recommends that the company undertakes a scenario analysis where HHF KM data from EMPEROR-Preserved is used to model time to HHF in the model. Using the KM HHF data would allow the company to fit a parametric survival curve to the data and extrapolate into the model's time horizon without having to assume a constant rate of HHF; a constant treatment effect with empagliflozin; and would allow first and subsequent HHF events to be modelled separately.

3. The final Poisson regression used in the model included two predictors: the time-varying KCCQ-CSS states; and treatment received. The EAG found a mistake in the predictors' coefficients reported in Table 32 of the CS, hence, asks that the company confirms what are the correct values and provides the corresponding p-values associated with the latter.

6.3 Conclusions of the cost effectiveness sections

Two key areas of uncertainty remain in the economic analysis: the effect of empagliflozin on patients' survival and the long-term effect of empagliflozin on patients' change in KCCQ-CSS scores.

The scenario analysis provided by the company help mitigate, at least partially, these areas of uncertainty. However, depending on the assumptions made, the EAG-preferred ICER falls within a wide range of [redacted] per QALY gained. Therefore, the EAG recommends that the committee validates the following assumptions:

1. Is empagliflozin likely to reduce CV mortality compared to SoC?
    a. If the answer to the above question is yes, then the EAG notes that there are two possible implications for non-CV deaths in this population:
        i. Either the CV deaths prevented in the empagliflozin arm do not translate into a reduction in overall mortality for these patients - given the age of the HFmrEF/HFpEF population, and the presence of several comorbidities, the proportion of patients who don't die of a CV cause die in a similar time frame of a non-CV cause, therefore suggesting that there will be more non-CV deaths for empagliflozin than SoC patients – in this case, it is crucial that the cost of a CV and non-CV death is incorporated into the economic

analysis, to assess the cost-effectiveness of "replacing" an equal number of CV deaths with non-CV deaths for empagliflozin patients. Or;

ii. The CV deaths prevented in the empagliflozin arm translate into a reduction in overall mortality for these patients vs SoC patients – the proportion of patients who don't die of a CV cause end up dying much later of a non-CV cause, and a similar proportion of empagliflozin and SoC patients die of non-CV causes in the shorter term.

b. If the answer to question 1 is no, then there is no difference in CV or non-CV deaths for empagliflozin and SoC patients.

2. The duration of the impact of empagliflozin on patients' KCCQ-CSS scores.
3. The duration of the impact of HHF events on patients' quality of life.

The EAG maintains its view that using the KM HHF data from EMPEROR-Preserved (independent from KCCQ-CSS states) would have allowed the company to estimate long-term HHF by relying on observed data and not assuming a constant rate of HHF. Importantly, the use of KM data would have allowed the company to model time to first and subsequent HHF separately and that this could be of importance given the results seen in EMPEROR- Preserved for the ITT population.

7 References

1. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *European Heart Journal* 2021:1-128. doi: 10.1093/eurheartj/ehab368
2. National Institute for Health and Care Excellence. Chronic heart failure in adults: diagnosis and management NICE guideline [NG106] 2018 [Available from: https://www.nice.org.uk/guidance/ng106 accessed 01 Sept 2021.
3. National Institute for Health and Care Excellence (NICE). Empagliflozin for treating chronic heart failure with reduced ejection fraction [TA773]. 2022. https://www.nice.org.uk/guidance/ta773 (accessed April 2022).
4. National Institute for Health and Care Excellence (NICE). Chronic heart failure in adults: diagnosis and management [NG106]. 2018. https://www.nice.org.uk/guidance/ng106 (accessed Sept 2022).
5. Cleland JGF, Bunting KV, Flather MD, et al. Beta-blockers for heart failure with reduced, mid-range, and preserved ejection fraction: an individual patient-level analysis of double-blind randomized trials. *Eur Heart J* 2018;39(1):26-35. doi: 10.1093/eurheartj/ehx564 [published Online First: 2017/10/19]
6. Cleland JGF, Tendera M, Adamus J, et al. The perindopril in elderly people with chronic heart failure (PEP-CHF) study. *European Heart Journal* 2006;27(19):2338-45. doi: 10.1093/eurheartj/ehl250
7. Solomon SD, McMurray JJV, Anand IS, et al. Angiotensin-Neprilysin Inhibition in Heart Failure with Preserved Ejection Fraction. *New England Journal of Medicine* 2019;381(17):1609-20. doi: https://dx.doi.org/10.1056/NEJMoa1908655
8. Yusuf S, Pfeffer MA, Swedberg K, et al. Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: the CHARM-Preserved Trial. *The Lancet* 2003;362(9386):777-81. doi: 10.1016/S0140-6736(03)14285-7
9. NICE. Dapagliflozin for treating chronic heart failure with reduced ejection fraction [TA 679] 2021 [Available from: https://www.nice.org.uk/guidance/ta679.
10. National Institute for Health and Care Excellence (NICE). Final scope for the appraisal of empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]. 2022. https://www.nice.org.uk/guidance/gid-ta10946/documents/final-scope (accessed Sep 2022).
11. ClinicalTrials.gov EMPERIAL - Preserved. NCT03448406. Available at: https://clinicaltrials.gov/ct2/show/NCT03448406 [Accessed: Nov 2022].
12. Ranganathan P, Pramesh C, Aggarwal R. Common pitfalls in statistical analysis: Intention-to-treat versus per-protocol analysis. *Perspect Clin Res* 2016;7(3):144-46. doi: 10.4103/2229-3485.184823
13. Butler J, Anker SD, Filippatos G, et al. Empagliflozin and health-related quality of life outcomes in patients with heart failure with reduced ejection fraction: the EMPEROR-Reduced trial. *Eur Heart J* 2021 doi: 10.1093/eurheartj/ehaa1007 [published Online First: 2021/01/10]
14. Anker SD, Butler J, Filippatos G, et al. Effect of Empagliflozin on Cardiovascular and Renal Outcomes in Patients With Heart Failure by Baseline Diabetes Status: Results From the EMPEROR-Reduced Trial. *Circulation* 2021;143(4):337-49. doi: 10.1161/CIRCULATIONAHA.120.051824 [published Online First: 2020/11/12]

15. Spertus JA, Jones PG, Sandhu AT, et al. Interpreting the Kansas City Cardiomyopathy Questionnaire in Clinical Trials and Clinical Care: JACC State-of-the-Art Review. *Journal of the American College of Cardiology* 2020;76(20):2379-90. doi: https://doi.org/10.1016/j.jacc.2020.09.542
16. Lachin JM. Fallacies of last observation carried forward analyses. *Clinical Trials* 2016;13(2):161-68. doi: 10.1177/1740774515602688
17. Eriksson B. Comorbidities, risk factors and outcomes in patients with heart failure and an ejection fraction of more than or equal to 40% in primary care- and hospital care-based outpatient clinics. *Scand J Prim Health Care* 2018;36:207–15.
18. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15. doi: 10.1016/j.jval.2012.02.008 [published Online First: 2012/08/08]
19. Sullivan PW, Slejko JF, Sculpher MJ, et al. Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making* 2011;31(6):800-4. doi: 10.1177/0272989X11401031 [published Online First: 2011/03/23]
20. McMurray JJV, Trueman D, Hancock E, et al. Cost-effectiveness of sacubitril/valsartan in the treatment of heart failure with reduced ejection fraction. *Heart* 2018;104(12):1006. doi: 10.1136/heartjnl-2016-310661
21. Personal Social Services Research Unit. Unit costs of health and social care 2020 2020 [Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/ accessed March 2021.
22. Eurostat. Harmonised Index of Consumer Prices (HICP): European Commission 2020 [Available from: https://ec.europa.eu/eurostat/web/hicp.
23. Alva ML, Gray A, Mihaylova B, et al. The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84). *Diabet Med* 2015;32(4):459-66. doi: 10.1111/dme.12647 [published Online First: 2014/12/03]

8 Appendices

8.1 Proportion of patients dead in the model

The EAG notes that the impact of empagliflozin on patients' overall survival is small in the company's base. Given that the difference in CV deaths between the empagliflozin arm and the SoC arm is larger than the difference in overall deaths (Figure 9), non-CV deaths (obtained as all cause-deaths minus CV deaths) are, by default, higher in the empagliflozin arm of the model than in the SoC arm.

When empagliflozin is assumed to not have an impact on patients' overall survival, but the benefit of the treatment on CV mortality is maintained, by default, the non-CV mortality in the SoC arm of the model decreases, increasing the ICER (Figure 10).

Figure 17. Company's base case

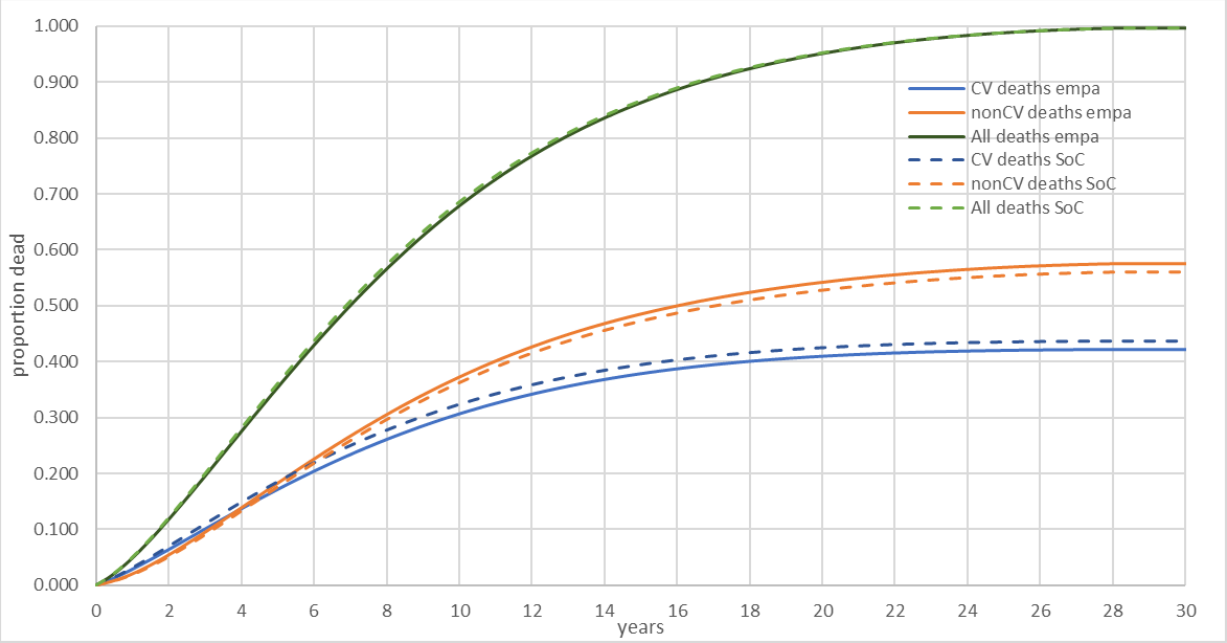


Figure 18. EAG scenario – all cause mortality assumed to be the same across treatment arms and empagliflozin reduced CV mortality

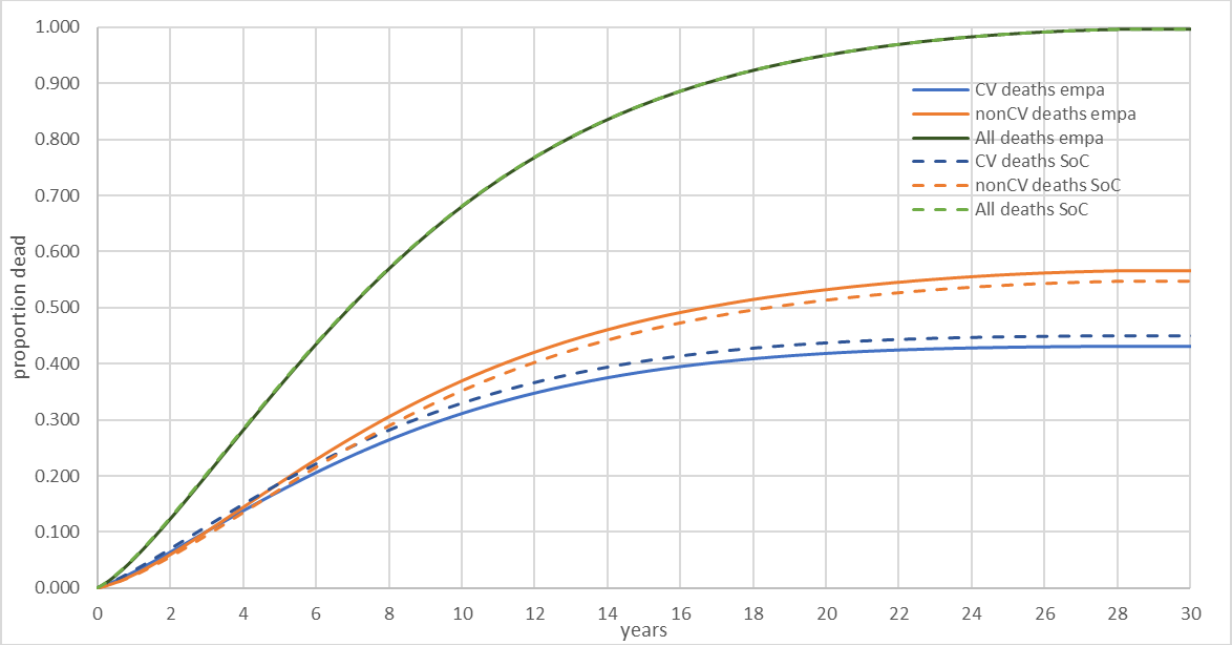
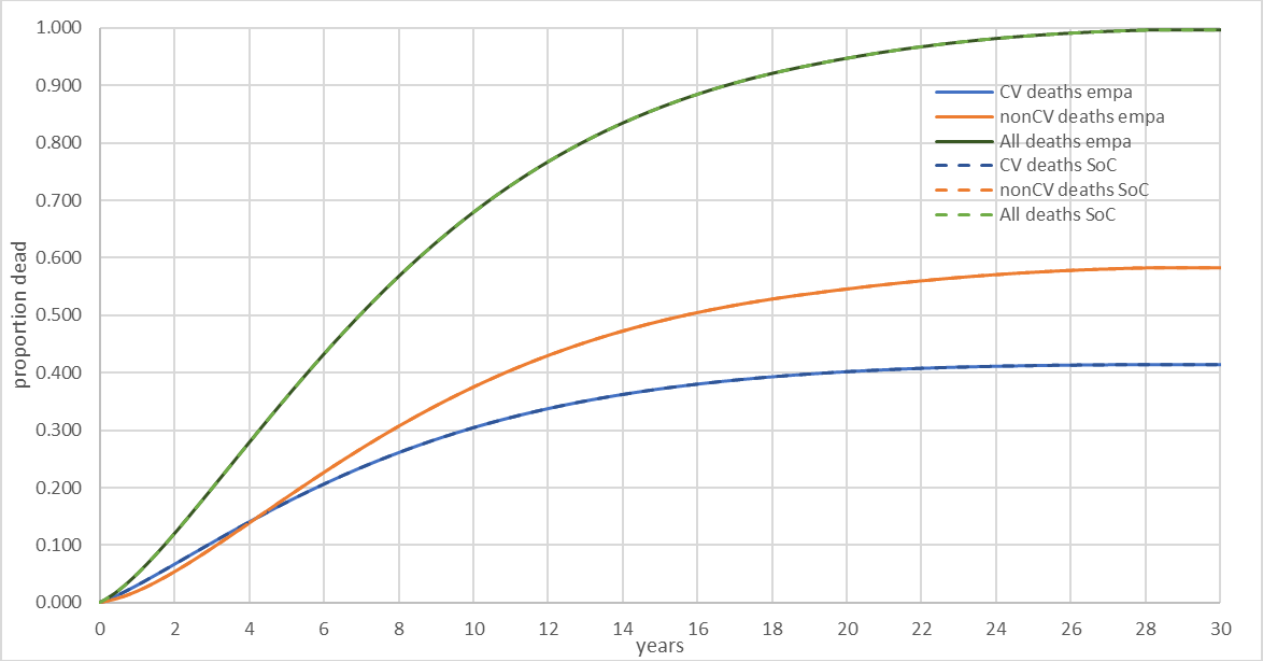


Figure 19. EAG scenario – no overall survival benefit associated with empagliflozin



8.2 Validation of model outcomes in EAG scenarios

Table 51. QALY gain in the economic model

	EMPEROR-Preserved	Base case	EAG preferred assumptions + Assuming that empagliflozin has a benefit on CV mortality + Assuming HHF impact patients' QoL for 11 weeks after HFF	EAG preferred assumptions + Assuming that empagliflozin has a benefit on CV mortality + Assuming HHF impact patients' QoL for 6 months year after HFF	EAG preferred assumptions + Assuming that empagliflozin does not have an effect on mortality + Assuming HHF impact patients' QoL for 11 weeks after HFF	EAG preferred assumptions + Assuming that empagliflozin does not have an effect on mortality + Assuming HHF impact patients' QoL for 6 months year after HFF
At week 52 (year 1)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At week 104 (year 2)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At week 156 (year 3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At week 208 (year 4)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At week 260 (year 5)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
At week 312 (year 6)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]						

Table 52. Comparison of total number of HHF observed in the trial and those predicted in the cost-effectiveness model

	EMPEROR-preserved			CE model – base case			EAG preferred assumptions + Assuming that empagliflozin has a benefit on CV mortality			EAG preferred assumptions + Assuming that empagliflozin does not have an effect on mortality		
	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference
Total number of HHF (first and subsequent) over 26 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of HHF (first and subsequent) over 3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of HHF (first and subsequent) over 3.58 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 53. Comparison of total number of deaths in the trial and those predicted in the cost-effectiveness model

	EMPEROR-preserved			CE model – base case			EAG preferred assumptions + Assuming that empagliflozin has a benefit on CV mortality			EAG preferred assumptions + Assuming that empagliflozin does not have an effect on mortality		
	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (For the equivalent N= 2,997)	Placebo +SoC (For the equivalent N= 2,991)	Difference	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference	Empagliflozin + SoC (N= 2,997)	Placebo +SoC (N=2,991)	Difference
Total number of CV deaths at 26 months	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of CV deaths at 3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of CV deaths over 3.5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total number of overall deaths over 3.5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Single Technology Appraisal

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

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 **5pm on Monday 5 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 'commercial in confidence' in turquoise, all information submitted as 'academic in confidence' in yellow, and all information submitted as 'depersonalised data' in pink.

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Pg 16, text states "The EAG notes that it has not seen any measures of fit (statistical or visual) to the different models used by the company, so it cannot assess the fit of the models, and the respective underlying hazard." The company should provide the additional analysis requested above, by first and subsequent event. The company should also provide measures of fit (statistical and visual) of the parametric used in the analysis and if needed, consider the use of different survival models.	Propose deleting text "visual" fits as this information was provided. Signpost provision of visual fits for HHF rates over time.	Statistical and visual fits of time to first heart failure hospitalisation was provided in Appendix N, Section 6.2 and Clarification Questions, QB10b, Figure 5.	Not a factual inaccuracy. No change required. The EAG maintains its view that it has not seen the different parametric models fitted to the underlying KM HHF data used by the company in this analysis. Figure 5 of the clarification document shows the hazard extrapolations from each parametric model, whereas Section 6.2 in Appendix N links to an Excel spreadsheet where (as far as the EAG could ascertain) only the GEE Poisson models are provided.

Location of incorrect marking	Description of incorrect marking	Amended marking	EAG response
CEM post clarification questions **File name: ID3945 empagliflozin EAG model 25112022CM[ACIC] BI updated 01 12 2022**	Context tab: No AIC marking Characteristics Repository: AIC not underlined Risk Equations – Active AIC not underlined Risk Equations – Storage: AIC/CIC not underlined Clinical inputs AIC not underlined	Underlined and marked in yellow	The EAG thanks the company for providing an updated model with the correct AIC and CIC marking.

	Non-imputed data underlined and highlighted in yellow Utilities Age adjustment to KCCQ-CSS quartile utility values based on UK general population underlined and highlighted in yellow		

(Please add further lines to the table as necessary)

BMJ TAG

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Addendum to the EAG report

January 2023

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135656.

1 Introduction

This document provides the additional results requested by NICE resulting from the Evidence Assessment Group's (EAG's) critique of the company's submission.

2 Additional analysis requested by NICE

The company's base case assumptions regarding the impact of empagliflozin on survival are the following:

A. Empagliflozin has an indirect impact on patients' overall survival (OS) and an indirect impact on CV-related mortality through its improvement in KCCQ-CSS (i.e., patients receiving empagliflozin have better KCCQ-CSS and therefore have a lower probability of dying of CV or non-CV causes).
B. Empagliflozin has a direct impact on patients CV-related mortality (i.e., patients receiving empagliflozin have a lower probability of dying of CV-related causes), independent of the improvement in the KCCQ-CSS.
C. Empagliflozin does not directly impact overall survival (i.e., being on treatment does not directly increase patients' probability of overall survival).

As noted in the EAG report, the EAG disagrees with the assumption made by the company regarding an indirect effect of empagliflozin on OS through KCCQ-CSS. Given the KM OS data reported in Figure 1, which shows that survival was similar in the two arms of the EMPEROR-Preserved trial; and clinical expert opinion, the EAG considers that including a treatment benefit associated with empagliflozin on OS is unsubstantiated.

As also discussed in the EAG report, even though the company did not provide a clinical rationale to reconcile point B and point C, the EAG considers these could be reconciled through the fact that patients with HFmrEF/HFpEF have several comorbidities, thus the proportion of patients who do not die of a CV cause die in a similar time frame of a non-CV cause, therefore, the reduction seen in CV deaths for empagliflozin patients might not translate into an overall survival gain.

During the pre-meeting briefing (PMB), the NICE technical team requested that the following two scenarios were undertaken: 1) empagliflozin affects the direct and indirect probability of patients

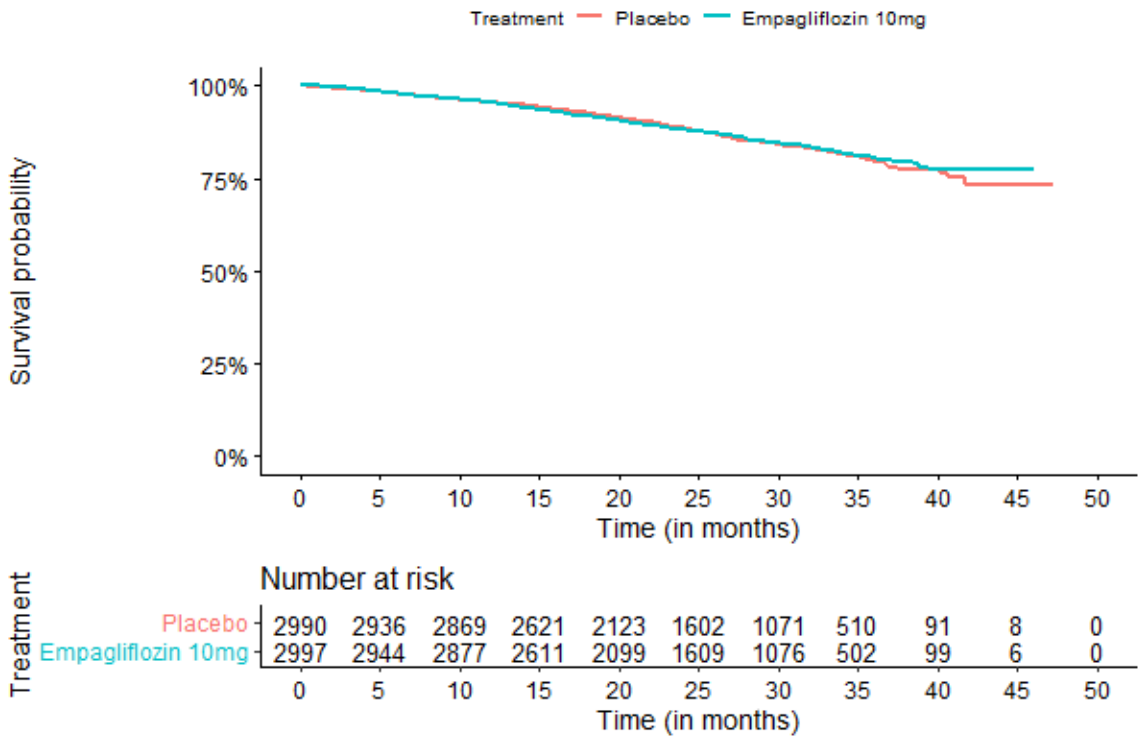
dying of CV or other causes; or 2) empagliflozin has no impact on survival (whether overall of CV-related).

These results can potentially be explained by the OS curves in Figure 1, and the close number of events in both curves (a total of 422 [14.1%] patients died in the empagliflozin arm of EMPREROR-Preserved, while 427 [14.3%] patients died in the placebo arm), which overall, indicate that deaths are very similar and not statistically significantly different across treatment arms (HR of 0.99; 95% confidence interval of 0.865 to 1.131).

Table 1. Weibull regression for OS, ITT population from EMPEROR-Preserved

Covariate	Coefficient	p-value
Shape		
Scale		
Treatment effect Empagliflozin 10 mg (reference placebo)		
KCCQ-CSS: (Quartile 2 vs Quartile 1)		
KCCQ-CSS: (Quartile 3 vs Quartile 1)		
KCCQ-CSS: (Quartile 4 vs Quartile 1)		

Figure 1. Observed OS data in EMPEROR-Preserved (reproduced from Figure 19, CS).



Therefore, the EAG has no means of conducting a scenario analysis where empagliflozin has a direct benefit on OS. Furthermore, a scenario incorporating a direct effect of empagliflozin on OS, using the same approach as the company has for CV mortality, would result in a survival benefit for placebo – the EAG does not consider this clinically plausible. Therefore, the EAG maintained its two original scenario analysis where:

1. Empagliflozin does not have an effect on overall survival (directly or indirectly) or on CV mortality.
2. Empagliflozin does not have an effect on overall survival but has an effect on CV mortality. As acknowledged by the EAG in its original report, the KM curves from EMPEROR-Preserved for CV survival do not provide sufficient evidence to substantiate empagliflozin having an impact on patients' survival compared to SoC patients. However, clinical expert opinion consistently reported the plausibility of empagliflozin reducing patients' CV-related mortality. Therefore, the EAG provided ICERs with its preferred assumptions with and without the benefit of empagliflozin on CV mortality.

The EAG has provided a detailed discussion in its original report regarding the plausibility of scenario 2, which the EAG replicates here to aid the discussion.

It is possible the due to the age of the HFmrEF/HFpEF population (72 years in the EMPEROR-Preserved trial and expected to be close to 80 years in clinical practice) and the presence of several comorbidities in this population (for example, about 50% of patients in the trial had type II diabetes at baseline), the CV deaths potentially avoided by empagliflozin are "replaced" by non-CV deaths within a similar timeframe. Therefore, the EAG recommends that the committee validates the following assumptions:

1. Is empagliflozin likely to reduce CV mortality compared to SoC?
    a. If the answer to the above question is yes, then the EAG speculates that there are two possible implications for non-CV deaths in this population:
        i. Either the CV deaths prevented in the empagliflozin arm do not translate into a reduction in overall mortality. Or;
        ii. The CV deaths prevented in the empagliflozin arm translate into a reduction in overall mortality for these patients vs SoC patients – the proportion of patients who don't die of a CV cause end up dying much later of a non-CV cause, and a similar proportion of empagliflozin and SoC patients die of non-CV causes in the shorter term. The EAG notes that the evidence available (the EMPEROR-Preserved trial data and clinical expert opinion provided to the EAG) does not support this scenario. Therefore, as discussed in the EAG report, the EAG preferred assumption is to assume no impact on OS associated with empagliflozin.
    b. If the answer to question 1 is no, then there is no difference in CV or non-CV deaths for empagliflozin and SoC patients.

2.1 EAG preferred assumptions

As discussed in the EAG report, the EAG's preferred assumptions are:

1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.
2. Assuming that empagliflozin does not have an effect on overall survival (and not changing the company's base case assumption that empagliflozin impacts CV mortality).
3. Using the age-related decrements from Ara and Brazier 2010.

4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.
5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).

In addition to the assumptions listed above, the EAG conducted one set or alternative scenarios:

6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.
7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.

Finally, the EAG conducted the alternative combined scenarios described above with the following:

8. Assuming that empagliflozin has no effect on overall survival and on CV mortality. The EAG reinforces its view that the evidence available from the EMPEROR-Preserved trial data does not support a effect of empagliflozin on survival (CV or non-CV related), and it is only clinical expert opinion that suggests the treatment might impact CV mortality. Therefore, the EAG considers that the most appropriate base case ICER will depend on the committee's view on the latter issue.

When empagliflozin is assumed to have an impact on CV mortality (but not on overall survival), the cumulative EAG-preferred assumptions result in a final ICER of [redacted], depending on the duration of the impact of a HHF event on patients' quality of life (Table 2). When empagliflozin is assumed to not have an impact on CV mortality (or overall survival), the cumulative EAG-preferred assumptions result in a final ICER of [redacted], depending on the duration of the impact of a HHF event on patients' quality of life (Table 2). Given how similar the deterministic base case results are to the probabilistic results, the EAG did not present probabilistic results of the exploratory analysis conducted.

Table 2. EAG's preferred model assumptions

Preferred assumption	Section in EAG report	Cumulative ICER (£/QALY)	Net health benefit £20K threshold	Net health benefit £30K threshold
Company base case	Section 5	£14,429	0.03	0.05
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.	Section 4.2.61 and 4.2.6.2	[redacted]	[redacted]	[redacted]

2. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 3. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality.	Section 4.2.6.7	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010.	Section 4.2.8	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.	Section 4.3.1.2 and Section 4.3.1.3	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).	Section 4.3.1.4 and Section 4.3.1.5.	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 6. Assuming HHF impact patients' QoL for 11 weeks	Section 4.2.8	[redacted]	[redacted]	[redacted]

after hospitalisation.				
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival but has an effect on CV mortality. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.	Section 4.2.8	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 6. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation. 8. Assuming that empagliflozin does not have an effect on CV mortality.	Section 4.2.6.7	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Assuming that empagliflozin does not have an effect on overall survival. 3. Using the age-related decrements from Ara and Brazier 2010. 4. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 5. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation. 8. Assuming that empagliflozin does not have an effect CV mortality.	Section 4.2.6.7	[redacted]	[redacted]	[redacted]

Abbreviations: EAG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

3 Additional figures requested by NICE

Figure 2. Observed all-cause mortality data in EMPEROR-Preserved (zoomed in)

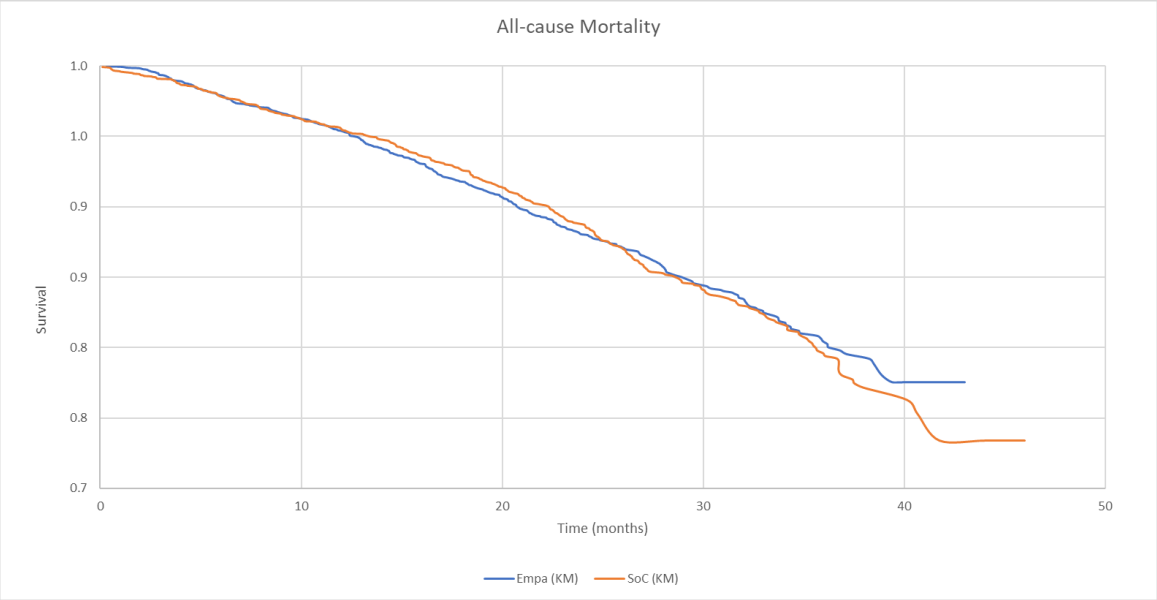


Figure 3. Observed all-cause mortality data in EMPEROR-Preserved (0 -1 scale, taken from CS, Figure 11)

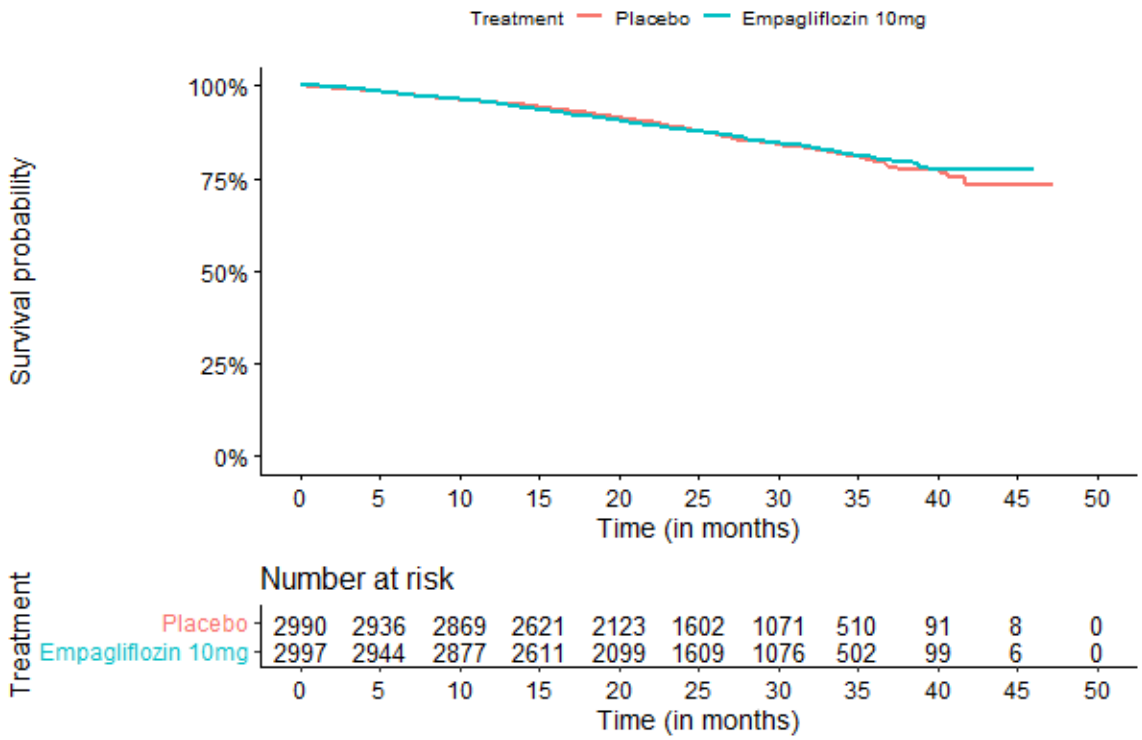


Figure 4. Observed CV-related mortality data in EMPEROR-Preserved (zoomed in)

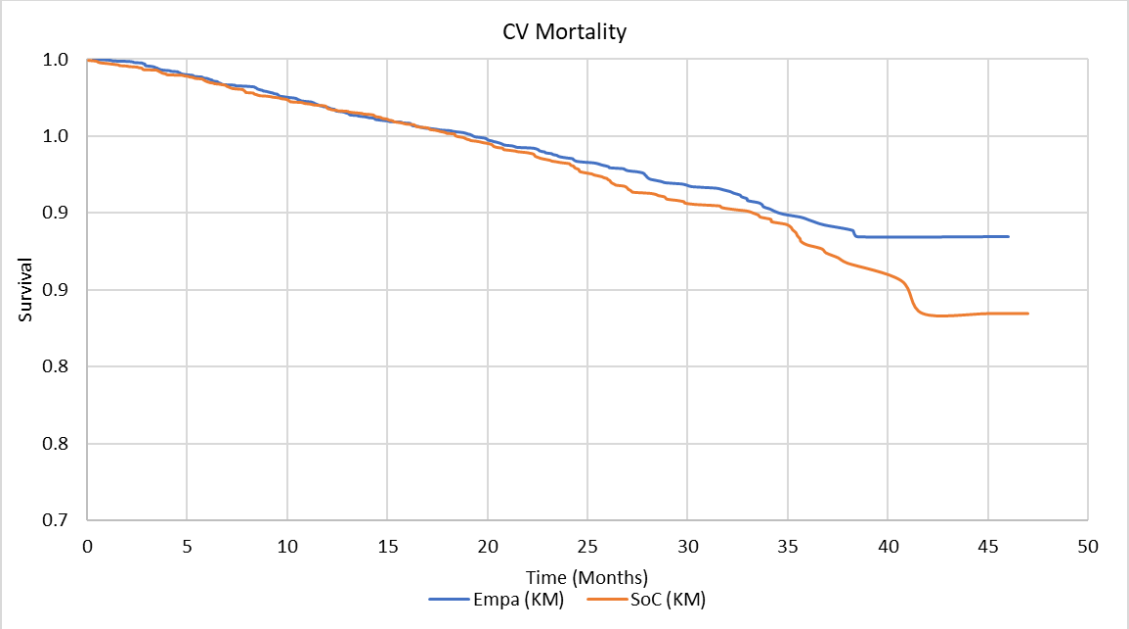
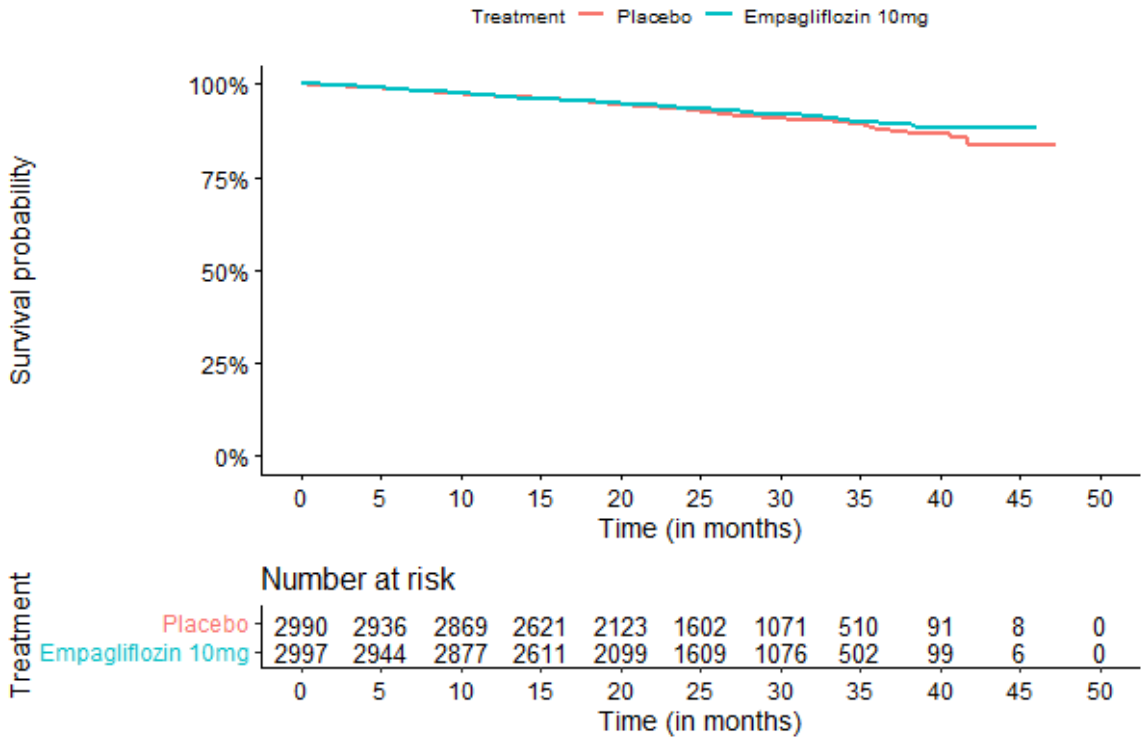


Figure 5. Observed CV-related mortality data in EMPEROR-Preserved (0 -1 scale, taken from CS, Figure 12)


Treatment
Placebo
Empagliflozin 10mg
Survival probability
100%
75%
50%
25%
0%
0
5
10
15
20
25
30
35
40
45
50
Time (in months)
Number at risk
Treatment
Placebo
Empagliflozin 10mg
2990 2936 2869 2621 2123 1602 1071 510 91 8 0
2997 2944 2877 2611 2099 1609 1076 502 99 6 0
0
5
10
15
20
25
30
35
40
45
50
Time (in months)

Scenarios assessing direct and indirect treatment effect of empagliflozin on CV and all-cause deaths:

The table below outlines the ICER outcomes depending on empagliflozin treatment effect assumptions. The scenarios include the EAGs preferred assumptions and assuming that HHF events impact patients' quality of life by 2.75 months

Assumptions			Incremental costs	Incremental LYG	Incremental QALYs	ICER
CV death: Direct and indirect effect	All-cause deaths: Direct and indirect effect	Non-CV death cost included	NA	NA	NA	NA
CV death: Direct and indirect effect	All-cause deaths: Direct and indirect effect	Non-CV death cost not included	NA	NA	NA	NA
CV death: Direct and indirect effect	All-cause deaths: Indirect effect only	Non-CV death cost included	NA	NA	NA	NA
CV death: Direct and indirect effect	All-cause deaths: Indirect effect only	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]
CV death: Direct and indirect effect	All-cause deaths: Indirect effect only	Non-CV death cost included	NA	NA	NA	NA
CV death: Direct and indirect effect	All-cause deaths: No effect	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]
CV death: Indirect effect only	All-cause deaths: No effect	Non-CV death cost included	NA	NA	NA	NA
CV death: Indirect effect only	All-cause deaths: Indirect effect only	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]
CV death: No effect*	All-cause deaths: No effect*	Non-CV death cost included	NA	NA	NA	NA
CV death: No effect*	All-cause deaths: No effect*	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]

NA, not applicable. These scenarios cannot be run due to the structure of the empagliflozin model. *The removal of the indirect treatment effect reflects there is no survival benefit from KCCQ health state occupancy in addition to the removal of the indirect treatment effect.

BMJ TAG

Empagliflozin for treating chronic heart failure with preserved or mildly reduced ejection fraction [ID3945]

Second addendum to the EAG report

January 2023

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135656.

1 Introduction

This document provides the additional results requested by NICE resulting from the Evidence Assessment Group's (EAG's) critique of the company's submission.

2 Additional analysis requested by NICE

The NICE technical team requested that the following scenario was undertaken: empagliflozin affects the direct and indirect probability of patients dying of CV or other causes.

In order to conduct the scenario requested, the treatment coefficient in the Weibull risk equations used to estimate OS in the model has to be used (as in the company's base case the coefficient was set to zero). The EAG notes that the treatment effect coefficient is [redacted] therefore indicating that patients on placebo have a survival [redacted] compared to patients on empagliflozin.

2.1 EAG preferred assumptions

As discussed in the EAG report, the EAG's preferred assumptions are:

1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation.
2. Using the age-related decrements from Ara and Brazier 2010.
3. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events.
4. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero).

In addition to the assumptions listed above, the EAG conducted one set or alternative scenarios:

5. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation.
6. Assuming HHF impact patients' QoL for 6 months year after hospitalisation.

Finally, the EAG conducted the alternative combined scenarios described above with the following: Assuming that empagliflozin affects the direct and indirect probability of patients dying of CV or other causes.

Table 1. EAG's preferred model assumptions

Preferred assumption	Incremental costs	Incremental QALYs	Cumulative ICER	Net health benefit	Net health benefit

			(£/QALY)	£20K threshold	£30K threshold
Company base case	[redacted]	[redacted]	£14,429	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Using the age-related decrements from Ara and Brazier 2010. 3. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 4. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 5. Assuming HHF impact patients' QoL for 11 weeks after hospitalisation. 6. Assuming that empagliflozin affects the direct and indirect probability of patients dying of CV or other causes.	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1. Use the observed KCCQ-CSS TPs data from EMPEROR-Preserved without imputation. 2. Using the age-related decrements from Ara and Brazier 2010. 3. Using the EB03E code to cost HHF in the model, rather than using weighted mean derived from national FCEs for more severe HHFs events. 4. Using a unit cost for CV death in the model of £1,452 (with the cost of sudden death assumed to be zero). 7. Assuming HHF impact patients' QoL for 6 months year after hospitalisation. 8. Assuming that empagliflozin affects the direct and indirect probability of patients dying of CV or other causes.	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: EAG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Assumptions			Incremental costs	Incremental LYG	Incremental QALYs	ICER
CV death: Direct and indirect effect	All cause mortality : Direct and indirect effect	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]
CV death: Direct and indirect effect	All cause mortality: Indirect effect only	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]
CV death: Indirect effect only	All cause mortality : Indirect effect only	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]
CV death: No effect*	All cause mortality : No effect*	Non-CV death cost not included	[redacted]	[redacted]	[redacted]	[redacted]