Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Confidential until published

Canagliflozin in combination therapy for treating type 2 diabetes

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors Vicky Copley
Karen Pickett
Jonathan Shepherd
Emma Loveman
Jeremy Jones

Correspondence to Jonathan Shepherd
Southampton Health Technology Assessments Centre (SHTAC)
University of Southampton
First Floor, Epsilon House
Enterprise Road, Southampton Science Park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

Date completed 17th December 2013

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/63/01.

Declared competing interests of the authors

None

Acknowledgements

We are very grateful to Professor Richard Holt, Professor in Diabetes and Endocrinology at University Hospital Southampton NHS Foundation Trust, for providing expert clinical advice during the production of this report.

We also thank Dr Jill Colquitt, Senior Research Fellow, SHTAC, for acting as internal editor for the ERG report, and Karen Welch, Information Specialist, SHTAC, for appraising the literature search strategies.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Copley V, Pickett K, Shepherd J, Loveman E, Jones J. Canagliflozin in combination therapy for treating type 2 diabetes: A Single Technology Appraisal. Southampton Health Technology Assessments Centre (SHTAC), 2013.

Contributions of authors

V Copley (Senior Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; K Pickett (Research Fellow) critically appraised the clinical effectiveness systematic review and drafted the report; J Shepherd (Principal Research Fellow) critically appraised the clinical effectiveness systematic review, the Network Meta Analysis, drafted the report and project managed the review; E Loveman (Senior Research Fellow) critically appraised the health economic systematic review and the economic evaluation and drafted the report; J Jones (Principal Research Fellow) critically appraised the Network Meta Analysis, the economic evaluation, and drafted the report.

Word count:

49,876

[redacted]

TABLE OF CONTENTS

1 Introduction to ERG Report 18
2 BACKGROUND 18
2.1 Critique of manufacturer's description of underlying health problem 18
2.2 Critique of manufacturer's overview of current service provision 18
2.3 Critique of manufacturer's definition of decision problem 20
3 CLINICAL EFFECTIVENESS 25
3.1 Critique of manufacturer's approach to systematic review 25
3.2 Summary statement of manufacturer's approach 43
3.3 Summary of submitted evidence 45
3.4 Summary 78
4 ECONOMIC EVALUATION 80
4.1 Overview of manufacturer's economic evaluation 80
4.2 Critical appraisal of the manufacturer's submitted economic evaluation 89
4.3 Additional work undertaken by the ERG 130
4.4 Summary of uncertainties and issues 140
5 End of life 141
6 Innovation 141
7 DISCUSSION 142
7.1 Summary of clinical effectiveness issues 142
7.2 Summary of cost effectiveness issues 143
8 REFERENCES 143

LIST OF TABLES

Table 1: Adult T2DM populations included in the manufacturer's decision problem 22
Table 2: The manufacturer and ERG's assessment of whether the identified 11 Phase 3 RCTs were relevant to the systematic literature review (as defined in the scope and decision problem) 31
Table 3: Manufacturer and ERG assessment of trial quality 35
Table 4: ERG appraisal of Network Meta Analysis (NMA) 40
Table 5: Quality assessment (CRD criteria) of MS systematic review 44
Table 6: RCT results for dual therapy, MET background (HbA_{1c} %) 46
Table 7: Proportion of patients achieving HbA_{1c} targets, dual therapy MET background 52 weeks 47
Table 8: NMA results for dual therapy 26 weeks, MET background (HbA_{1c} %) 48
Table 9: NMA results for dual therapy 52 weeks, MET background (HbA_{1c} %) 49
Table 10: RCT results for triple therapy, MET and SU background (HbA_{1c} %) 50
[redacted] 51
Table 12: NMA results for triple therapy 26 weeks, MET and SU background (HbA_{1c} %) 51
Table 13: NMA results for triple therapy 26 weeks, MET and TZD background (HbA_{1c} %) 52
Table 14: NMA results 26 weeks, for insulin background (HbA_{1c} %) 53
Table 15: RCT results for dual therapy MET background (SBP [mmHg]) 54

Table 16: RCT results for dual therapy MET background (DBP [mmHg]) 54
Table 17: NMA results for dual therapy, MET background at 26 weeks (SBP) 55
Table 18: NMA results for dual therapy, MET background at 52 weeks (SBP) 56
Table 19: RCT results for triple therapy, MET and SU background (SBP [mmHg]) 56
Table 20: RCT results for triple therapy, MET and SU background (DBP [mmHg]) 57
Table 21: NMA results for triple therapy 26 weeks, MET and SU background (SBP) ... 57
Table 22: NMA results for insulin background at 26 weeks (SBP) 58
Table 23 - RCT results for dual therapy MET background (weight [kg]) 58
Table 24: NMA results for dual therapy 26 weeks, MET background (weight) 60
Table 25: NMA results for dual therapy at 26 weeks, MET background (BMI) 60
Table 26: NMA results for dual therapy, MET background at 52 weeks (weight) 61
Table 27: NMA results for dual therapy, MET background at 52 weeks (BMI) 62
Table 28: RCT results for triple therapy, MET and SU background at 52 weeks (weight [kg]) 62
Table 29 – NMA results for triple therapy 26 weeks, MET and SU background (weight) 63
Table 30: NMA results for triple therapy 26 weeks, MET and SU background (BMI) 63
Table 31: NMA results for triple therapy at 26 weeks, MET and TZD background (weight) 64
Table 32: NMA results for insulin background at 26 weeks (weight) 65
Table 33: RCT results for dual therapy MET background (HDL-cholesterol) 66
Table 34: RCT results for dual therapy MET background (LDL-cholesterol) 67
Table 35: RCT results for dual therapy MET background (triglycerides) 67
Table 36: RCT results for triple therapy, MET and SU background (HDL-cholesterol).. 68
Table 37: RCT results for triple therapy, MET and SU background (LDL-cholesterol) .. 69
Table 38: RCT results for triple therapy, MET and SU background (triglycerides) 69
Table 39: RCT results for dual therapy MET background (hypoglycaemia) 70
Table 40: NMA results for dual therapy, MET background at 52 weeks (all hypoglycaemic events) 71
Table 41: RCT results for triple therapy, MET+SU background (hypoglycaemia) 71
Table 42: NMA results for triple therapy at 26 weeks, MET and SU background (all hypoglycaemic events) 72
Table 43: NMA results for triple therapy, MET and TZD background (all hypoglycaemic events) 72
Table 44: Pre-defined sub-group analyses by baseline HbA_{1c} %, MET background (HbA_{1c} %) 74
Table 45: Pre-defined sub-group analyses by baseline HbA_{1c} %, MET + SU background (HBA1c %) 75
Table 46: Overall summary of adverse events (broad dataset: safety analysis set)..... 76
Table 47: Summary of specific adverse events (broad dataset: safety analysis set) 77
Table 48: Manufacturer's base case results for canagliflozin in dual therapy combinations (MET background), with correction (see footnote b) 86
Table 49: Manufacturer's base case results for canagliflozin in triple therapy combinations 87
Table 50: Manufacturer's base case results for canagliflozin as 'add on' to insulin therapy 88

Table 51: Critical appraisal checklist of economic evaluation 90
Table 52: NICE reference case requirements 91
Table 53: Model settings (reproduced from MS Table 34 p184) 96
Table 54: Specific comparators selected within each pharmacological class (extract of MS Table 33 p176) 99
Table 55: Economic modelling comparisons conducted (reproduced from MS Table 32 p175). 100
Table 56: Treatment effects, Rates of AEs and 1st year discontinuation due to AEs: Dual therapy, Base cases; Canagliflozin, TZD, GLP-1-a, dapagliflozin (extract of MS Table 185 Appendix 10.9) 102
Table 57: Drift assumptions for biomarkers for all lines of therapy (reproduced from MS Table 36 p186, with CIs added) 105
Table 58: Baseline utilities and sources used in the MS model (MS Table 38) 111
Table 59: Drug acquisition costs applied in the model for all lines of therapy (MS Table 39). 115
Table 60: Direct first year treatment costs associated with complication and co-morbidity (MS Appendix 11) 117
Table 61: Direct complication and co-morbidity annual follow-up costs (MS Appendix 11) 117
Table 62: Direct event costs associated with complication and co-morbidity (MS Appendix 11) 118
Table 63: Summary of adverse event costs (MS Table 40) 119
Table 64: Treatment effect values used in DSA for canagliflozin 100mg vs. SU in dual therapy (extract of MS Table 195 Appendix p322 with additional column) 123
Table 65. Summary of PSA results presented in MS 127
Table 66: Parameter distributions for PSA in ECHO-T2DM (reproduced from MS Table 46 p217) 129
Table 67: Base case incremental cost-effectiveness results for canagliflozin 100mg in dual therapy, ERG analysis compared to MS analysis (1000 patients and 1000 cohorts) 131
Table 68: Base case incremental cost-effectiveness results for canagliflozin 300mg in dual therapy, ERG analysis compared to MS analysis (1000 patients and 1000 cohorts) 131
Table 69: Base case incremental cost-effectiveness results for canagliflozin 100mg in dual therapy, 100,000 patients and no parameter uncertainty (1 cohort) 132
Table 70: Base case incremental cost-effectiveness results for canagliflozin 300mg in dual therapy, 100,000 patients and no parameter uncertainty (1 cohort) 133
Table 71: Fully incremental analysis based upon MS original base case, canagliflozin 100mg in dual therapy. Excludes TZD. 134
Table 72: Fully incremental analysis based upon ERG base case with 100,000 patients and no parameter uncertainty, canagliflozin 100mg in dual therapy. Excludes TZD. . 135
Table 73: Fully incremental analysis based upon MS original base case, canagliflozin 300mg in dual therapy. Excludes TZD. 136
Table 74: Fully incremental analysis based upon ERG base case with 100,000 patients and no parameter uncertainty, canagliflozin 300mg in dual therapy. Excludes TZD. . 137

Table 75: Summary of ERG PSA results for canagliflozin in comparison with dapagliflozin in dual therapy with metformin (1000 patients and 1000 cohorts) 139

LIST OF FIGURES

Figure 1: Schematic structure of the ECHO-T2DM model (reproduced from MS figure 59 p168) 94
Figure 2: Cost-effectiveness of canagliflozin 100mg and other treatments in dual therapy based upon MS original base case. Excludes TZD. 134
Figure 3: Cost-effectiveness of canagliflozin 100mg and other treatments in dual therapy based upon ERG base case with 100,000 patients and no parameter uncertainty. Excludes TZD. 135
Figure 4: Cost-effectiveness of canagliflozin 300mg and other treatments in dual therapy based upon MS original base case. Excludes TZD. 137
Figure 5: Cost-effectiveness of canagliflozin 300mg and other treatments in dual therapy based upon ERG base case with 100,000 patients and no parameter uncertainty. Excludes TZD. 138

LIST OF ABBREVIATIONS

AE	Adverse event
AHA	Anti-Hyperglycaemic Agent
BID	Twice daily
BMI	Body Mass Index
BNF	British National Formulary
CANA	Canagliflozin
CHF	Congestive heart failure
CKD	Chronic kidney disease
CrI	Credible interval
CSR	Clinical study report
CRD	Centre for Reviews and Dissemination
CVD	Cardiovascular disease
DPP-4-i	Dipeptidyl peptidase-4 inhibitor
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ESRD	End stage renal disease
FDA	Food and Drug Administration
FPG	Fasting plasma glucose
GLP-1-a	Glucagon-like peptide-1 agonist
GMI	Genital mycotic infections
HR	Hazard Ratio
HRQoL	Health related quality of life
HTA	Health Technology Assessment
IHD	Ischaemic heart disease
LOCF	Last Observation Carried Forward
LS	Least squares
LYG	Life Years Gained
MACE	Major Adverse Cardiovascular Events
MET	Metformin
MET+SU	Metformin and a sulfonylurea
MET+TZD	Metformin and a thiazolidinedione
MI	Myocardial infarction
mITT	Modified intention to treat
MS	Manufacturer’s submission
MTC	Mixed treatment comparison
NMA	Network meta-analysis
NPH	Neutral Protamine Hagedorn
OAD	Oral antidiabetic drug
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient reported outcomes
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PVD	Peripheral vascular disease
QALY	Quality Adjusted Life Years
RCT	Randomised Controlled Trial
SD	Standard deviation

SE	Standard error
SLR	Systematic Literature Review
SU	Sulfonylurea
SUCRA	Surface Under the Cumulative Ranking
SGLT	Sodium glucose co-transporter
TA	Technology assessment
T2DM	Type 2 diabetes mellitus
TTO	Time Trade Off
TZD	Thiazolidinedione
UKPDS	UK Prospective Diabetes Study
WTP	Willingness-to-pay

SUMMARY

Scope of the manufacturer submission (MS)

The scope of the submission was to assess the clinical effectiveness and cost effectiveness of canagliflozin (100mg and 300mg) as combination therapy in the treatment of type 2 diabetes mellitus (T2DM), compared to other treatments: sulfonylureas (SU), the thiazolidinedione (TZD) drug pioglitazone, dipeptidyl peptidase-4 inhibitors (DPP-4-i), glucagon-like peptide-1 analogues (GLP-1-a), insulin, and the sodium glucose co-transporter 2 (SGLT-2) dapagliflozin.

Summary of submitted clinical effectiveness evidence

The manufacturer's submission to NICE included:

- A systematic literature review comprising 11 randomised controlled trials of canagliflozin
- A systematic literature review of 174 studies (in a total of 208 publications) of the diabetes treatments covered in the scope (including the trials of canagliflozin). This was described in a separate report, and was conducted to support the manufacturer's network meta-analysis.
- A network meta-analysis of 64 trials to provide direct and indirect evidence of canagliflozin compared to the other diabetes treatments in the scope.

Of the 11 included canagliflozin trials:

- three assessed dual therapy
    - canagliflozin 300mg and 100mg versus sulfonylurea (glimepiride) on a background of metformin (DIA3009)
    - canagliflozin 300mg and 100mg versus placebo or DPP-4 inhibitor (sitagliptin) on a background of metformin (DIA3006)
    - canagliflozin 300mg and 100mg versus placebo (DIA3008 sulfonylurea sub-study)

- three assessed triple therapy
    - Canagliflozin 300 mg versus DPP-4 inhibitor sitagliptin 100 mg on a background of metformin and sulfonylurea (DIA3015)
    - canagliflozin 300 mg and 100 mg versus placebo on a background of metformin and sulfonylurea (DIA3002)[1;21;21;21;2]
    - canagliflozin 300mg and 100mg versus placebo/ DPP-4 inhibitor (sitagliptin) on a background of metformin and thiazolidinedione (DIA3012)

- one assessed add-on to insulin therapy
    - canagliflozin 300mg and 100mg versus placebo on a background of insulin and standard of care (DIA3008 insulin sub-study)

Four trials evaluated canagliflozin in various patient populations (older patients, with moderate renal impairment, Asian patients, patients at high risk for cardiovascular disease (CVD). These all appear to be placebo controlled trials and are not considered in great detail in the submission.

The ERG considered that only the trials with active comparator arms were relevant to the NICE scope and has focused on these in this report: two dual therapy trials (DIA3009 and DIA3006) and one triple therapy trial (DIA3015). All three were non-inferiority trials and were judged to be of to be of generally good methodological quality.

A number of outcome measures were included in the decision problem, matching the NICE scope, including HbA_{1c} control, hypoglycaemia, cardiovascular risk factors (e.g. blood pressure, serum lipids), weight change, diabetes complications (e.g. cardiovascular), mortality, and adverse effects of treatment.

Canagliflozin 100mg was shown to be non-inferior to glimepiride at week 52 (-0.01%, (95% CI -0.109; 0.085)). Canagliflozin 300mg in combination with metformin was shown to be statistically superior to glimepiride in combination with metformin in HbA_{1c} at week 52 (difference of mean change, minus glimepiride: -0.12%, (95% CI -0.217;-0.023)), based on modified intention to treat analysis (mITT). However, per-protocol analysis showed it to be only non-inferior and not superior to glimepiride. Canagliflozin 100mg in combination with metformin was non-inferior to sitagliptin and metformin in the difference of mean change in HbA_{1c} at week 52, 0.00% (95% CI -0.12; 0.12). Likewise, canagliflozin 300mg in combination with metformin was non-inferior to sitagliptin and metformin in the difference of mean change in HbA_{1c} at week 52, -0.15% (95% CI -0.27; -0.03)). Canagliflozin 300mg also demonstrated statistical superiority to sitagliptin in reducing HbA_{1c} (with the upper limit of the 95% CI less than 0.0%), however it was not found to be superior in per-protocol analysis.

Canagliflozin 300mg in combination with metformin and sulfonylurea was non-inferior to sitagliptin in combination with metformin and sulfonylurea in the difference of mean change in

HbA_{1c} at week 52 (-0.37% (95 CI -.05; -0.25). Canagliflozin 300mg also demonstrated statistical superiority in both mITT and per-protocol analysis.

The two dual therapy trials found that, at week 52, both doses of canagliflozin (100mg and 300mg) in combination with metformin were associated with statistically significant greater decreases in patients' systolic and diastolic blood pressure from baseline levels than the comparators (glimepiride and sitagliptin, both combined with metformin). The triple therapy trial found that canagliflozin (300mg) combined with metformin and a sulfonylurea resulted in the statistically significant greater improvement in both systolic and diastolic blood pressure at 52 weeks than the DPP-4-i sitagliptin combined with metformin and a sulfonylurea.

In both the dual therapy trials canagliflozin 300mg and 100mg were associated with statistically significant greater weight loss at 52 weeks than sulfonylurea glimepiride or the DPP-4-i sitagliptin. In the triple therapy trial, at 52 weeks canagliflozin (300mg) combined with metformin and a sulfonylurea resulted in a statistically significant greater improvement in BMI than the DPP-4-i sitagliptin (100mg) combined with metformin and a sulfonylurea.

The manufacturer's pooled analysis of all active- and placebo- controlled canagliflozin trials showed a similar rate of any adverse events between canagliflozin and non-canagliflozin (76.8% for all canagliflozin and 75.8% for all non canagliflozin). There was a slightly higher percentage of patients discontinuing due to adverse events for canagliflozin versus comparator (6.4% versus 5.0%), and a 10% higher rate of patients experiencing adverse events related to the study drug for all canagliflozin. The percentage of patients with serious adverse events and the percentage of deaths was similar between canagliflozin and non-canagliflozin.

The MS includes a network meta-analysis (NMA) to compare canagliflozin 100mg and 300mg to other active treatments included in the decision problem. Separate networks of evidence were constructed by background therapy: metformin (26 weeks and 52 weeks); metformin and sulfonylurea (26 weeks); insulin (26 weeks). A number of outcomes were included, though it is only HbA_{1c} and BMI from the NMA that appear to be consistently used as input data in the economic model. The probability of canagliflozin ranking as best treatment varied across the networks and by outcome measure. For the metformin background at 52 weeks canagliflozin 300mg was the third highest ranking treatment in terms of reducing HbA_{1c} (probability of 71%) behind the SU glibenclamide (73%) and the GLP-1 agonist liraglutide 1.2mg once daily (89%).

Canagliflozin 300mg also ranked highly in the triple therapy and insulin background networks on this outcome.

Summary of submitted cost effectiveness evidence

The manufacturer's submission to NICE includes:

i) A review of published economic evaluations of treatments for T2DM.
ii) A report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of canagliflozin in dual therapy with metformin is compared with dapagliflozin, DPP-4-i, GLP-1-a, SU and TZD. The cost effectiveness of canagliflozin in triple therapy is compared with DPP-4-i, GLP-1-a and insulin. The cost effectiveness of canagliflozin as an add-on to insulin is compared with DPP-4-i, GLP-1-a, and dapagliflozin.

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations in T2DM of canagliflozin, agents in the class of SGLT-2 inhibitors (dapagliflozin) and intended comparators (DPP-4-i, GLP-1-a, SU). The MS does not make any general conclusions about the systematic review.

The cost effectiveness analysis (CEA) uses a stochastic micro-simulation model (the ECHO-T2DM model) to estimate the cost-effectiveness of canagliflozin compared with SU, TZD, DPP-4-i, GLP-1-a, dapagliflozin and insulin in adults with T2DM. Cohorts of individual hypothetical patients are created and simulated over time. Parameter uncertainty is captured by many cohorts of patients. The model adopts a 40 year time horizon and has an annual cycle length.

The model has three parallel sets of micro-vascular complications and five types of macro-vascular complications represented as Markov states. There is a health state where the patient is free from complications, and then micro-vascular health states relating to chronic kidney disease (includes 7 stages); neuropathy (includes 5 conditions); and retinopathy (includes 4 conditions). Patients can also experience macro-vascular complications and peripheral vascular disease.

Twenty one scenario analyses were performed. Deterministic sensitivity analyses were undertaken on 35 parameters or sets of parameters. Probabilistic sensitivity analysis is a default feature of the ECHO-T2DM model and was conducted on all simulations.

For dual therapy combinations pairwise incremental costs per QALYs gained were calculated compared with TZD, SU, dapagliflozin, DPP-4-i and GLP-1-a. Canagliflozin 100mg was dominated by TZD. Other ICERs were £1,537 for SU, £8,674 for dapagliflozin, £97 for DPP-4-i and £50,005 for GLP-1-a where canagliflozin was less expensive and less effective than GLP-1-a. In the base case canagliflozin 300mg was dominated (MS states dominates in error) by TZD and led to ICERs of £4,899 compared with SU, £18,349 compared with DPP-4-i, £27,419 compared with dapagliflozin, and £76,214 compared with GLP-1-a where canagliflozin was less expensive and less effective.

For triple therapy combinations with metformin and SU, canagliflozin 100mg dominates DPP-4-i and GLP-1-a. The ICER compared with insulin LA was £263. Canagliflozin 300mg dominates GLP-1-a and has ICERs of £13,287 and £607 for DPP-4-i and insulin LA respectively.

For triple therapy combinations with metformin and TZD canagliflozin 100mg led to an ICER of £1,095 compared with DPP-4-i and canagliflozin 300mg led to an ICER of £21,430 compared with DPP-4-i.

Commentary on the robustness of submitted evidence

Strengths

The assessment of clinical effectiveness is based on a systematic review generally conducted to a reasonable standard and reported in adequate detail, though with some uncertainties (see below).

The manufacturer has conducted an extensive clinical trials programme to evaluate canagliflozin in both dual and triple therapy, employing a range of relevant outcome measures. The methodological quality of the trials is generally good. However, not all of the trials included active comparators, and therefore many of the comparisons in the scope are not covered by direct trial evidence.

The economic model captures all important aspects of the disease pathway and extrapolates intermediate outcomes to final outcomes in a robust and consistent manner, drawing upon

standard sources from the literature. The model has been very well validated against external data.

Weaknesses and Areas of uncertainty

The manufacturer's decision problem does not fully meet the scope in terms of interventions and comparators included. Principally, the submission does not cover people with T2DM inadequately controlled on monotherapy with a sulfonylurea. It is stated that relatively few patients would receive sulfonylurea monotherapy (most would receive metformin), and those that did would not be suitable for canagliflozin due to renal impairment. Expert clinical advice to the ERG suggests that not all people taking a sulfonylurea monotherapy would do so because metformin was contraindicated on this basis. Omission of this intervention could exclude a significant minority of people with diabetes who may be treated in the NHS. (Note that the only trial identified by the manufacturer evaluating a canagliflozin on a sulfonylurea background would not meet the scope in the view of the ERG as there was no active comparator.) Furthermore, there is no direct evidence comparing canagliflozin with other anti-diabetic drugs as add-on to insulin presented in the submission.

The NMA has been conducted to a reasonable standard, and is reported in adequate detail. However, caution is advised due to the low volume of direct evidence for some of the links in the networks. It should be noted that not all of the results of the NMA are used in the economic model (HbA_{1c} and BMI are the two NMA outcomes that are used most consistently in the model).

The manufacturer's quality assessment of trials included in the NMA noted some uncertainties in methodological quality, presumably due to poor reporting in the primary publications. Furthermore, the ERG notes that the quality assessment judgements appear not to have been fully checked by a second reviewer or done independently. There is potential for the results of the NMA to be subject to bias (though trials which were not double-blinded were excluded as a sensitivity analysis).

The economic model is complicated, consisting of around 20,000 lines of R code (in 29 separate files) with an Excel front- and back-end which contain additional lines of visual basic for applications (VBA) code. There are around 800 user-specifiable inputs. It has not been possible in the time available to fully check all inputs or model wiring.

The model uses patient-level simulation. This is common in models of T2DM due to the potentially complex nature of the disease pathway. However base case results are reported from simulations which use only 1000 patients and which may not therefore robustly capture the ICERs. Base case results also incorporate parameter uncertainty.

The manufacturer does not supply a fully incremental base case analysis and consequently the place of canagliflozin in the cost-effectiveness hierarchy is not clear.

The base case results reported in the MS generally show few incremental QALYs accruing to canagliflozin over a lifetime horizon of 40 years, at small incremental cost. Canagliflozin generally appears cost-effective in these results. However because of its small incremental benefit even slight variation to parameter assumptions can lead to a large change in the ICER.

Parameter distributions used in probabilistic sensitivity analysis (PSA) are not well documented. The extent to which PSA explores uncertainty in the decision problem is not clear.

Summary of additional work undertaken by the ERG

The ERG conducted additional analyses to:

- examine the variation in final ICER associated with re-running the base cases
- examine the variation in final ICER arising with no parameter uncertainty and 100,000 patients
- perform fully incremental analyses for dual therapy
- identify the probability that canagliflozin is cost effective in comparison with dapagliflozin
- examine the effect of modifications to the assumed clinical efficacy of glimepiride and exenatide

Fully incremental analysis conducted by the ERG for dual therapy in combination with metformin found that, with no parameter uncertainty and 100,000 patients, canagliflozin 100 mg is associated with an ICER of £84,800 compared to the next best option, DPP-4-i. The incremental differences for this comparison are £245 in total costs and 0.003 total QALYs over the lifetime horizon of the model. Canagliflozin 300 mg is associated with an ICER of £17,903

compared to the next best option, DPP-4-i, with incremental total costs of £585 and incremental total QALYs of 0.033 over the lifetime horizon of the model.

Canagliflozin 300 mg has a lower probability of being cost-effective than dapagliflozin in dual therapy with metformin at a WTP threshold of £20,000 per QALY (46.7% probability). It has a 50% probability of being cost-effective compared to dapagliflozin at the £30,000 WTP threshold. Canagliflozin 100 mg has a 52.5% probability of being cost-effective compared to dapagliflozin at both the £20,000 and £30,000 WTP thresholds.

1 Introduction to ERG Report

This report is a critique of the manufacturer's submission (MS) to NICE from Janssen-Cilag Ltd on the clinical effectiveness and cost effectiveness of canagliflozin for treating type 2 diabetes mellitus (T2DM). It identifies the strengths and weaknesses of the MS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the MS was requested from the manufacturer by the ERG via NICE on 7th November 2013. A response from the manufacturer via NICE was received by the ERG on 22nd November 2013 and this can be seen in the NICE evaluation report for this appraisal.

The submission comprises the following key documents:

- The main MS document (310 pages)
- Appendices document (386 pages)
- Systematic clinical effectiveness literature review and network meta-analysis (NMA) report[3] (433 pages)
- Systematic review of Health Related Quality of Life (HRQoL) in diabetes patients report (174 pages)
- Swedish Institute for Health Economics systematic review of diabetes economic models and cost effectiveness[4] (116 pages)
- Economic model Technical Report[5] (152 pages)
- Economic model User Guide[6] (99 pages)

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The MS provides a clear and accurate overview of T2DM.

2.2 Critique of manufacturer's overview of current service provision

The MS details the current management approach to T2DM in clinical practice, including the use of individualised treatment goals for patients, and the current clinical pathway for the use of

blood-glucose-lowering therapy in T2DM, drawing on NICE clinical guidelines for diabetes[7-11] and NICE Pathways in Diabetes.[12] The description is accurate and reflects NICE guidance, except that it does not detail that:

- acarbose may used as a monotherapy instead of metformin or a sulfonylurea in patients unable to use other oral glucose-lowering medications[13] (but this is a minor point as this agent is not specified to be of interest in the NICE scope for any comparisons of canagliflozin with or against other therapies and it is not mentioned in the NICE scope).
- of the thiazolidinediones, the marketing authorisation for rosiglitazone has been suspended[7] and, although pioglitazone is still recommended by NICE and used in practice, from March 2009 it has had black triangle status[7] and in July 2011the MHRA raised concerns about the increased risk of bladder cancer in patients taking this drug.[7;14;15] However, the manufacturer does note in the economic analysis section (MS p. 233) that there are increasing concerns about adverse events associated with pioglitazone and that its use in clinical practice is decreasing.

The MS states that the recommendation in CG 87[7] that patients can receive triple therapy with metformin, sulfonylurea and insulin if dual therapy with metformin and sulfonylurea fails is not favoured in clinical practice, because many patients do not want to start injectable treatment. The clinical expert consulted by the ERG concurred that many patients would prefer to use an oral agent first and that a minority of triple therapy patients use insulin.

The manufacturer details on MS p. 35 the proportion and number of patients who are treated on different mono-, dual and triple therapy and add-on therapy to insulin regimens, using data from an unpublished report (CEGEDIM, 2013).[16] The data presented in the MS suggest that most people with diabetes will be treated with metformin as part of their therapy. Metformin is the standard first-line therapy recommended by NICE[7] and clinical expert advice to the ERG is that most patients will receive metformin. In the systematic literature review, the manufacturer has focused on regimens involving metformin and patients treated on a background of metformin. However, canagliflozin's licenced indication does not limit its use to regimens involving metformin and states that it may be used as an add-on therapy to other anti-hyperglycaemic agents. The manufacturer has not provided data on the proportion or number of patients receiving other therapeutic combinations to which canagliflozin potentially could be added (e.g. it could be used as part of dual therapy with a sulfonylurea or triple therapy with agents other than metformin). It is therefore unclear what proportion of patients may receive triple or dual

therapy regimens other than those including metformin and who could also potentially receive canagliflozin as part of their treatment in practice[17] (and who are thus not represented in the systematic literature review; see section 3.1.2 of this report). The clinical expert consulted by the ERG estimated that around 5-10% of patients experience persistent problems with metformin that lead to discontinuation. He stated that although this is proportionally a minority of patients, given the high prevalence of diabetes this means that a large number of patients do not receive metformin.

2.3 Critique of manufacturer's definition of decision problem

Population

Table 1 summarises the patient populations included in the manufacturer's decision problem for the assessment of dual, triple and add-on therapy to insulin and whether they match those specified in the final scope issued by NICE. In line with the final scope, the manufacturer has included adults with T2DM that has been inadequately controlled with monotherapy (for the assessment of dual therapy), dual therapy (for the assessment of triple therapy) and insulin alone or with other agents (for the assessment of add-on therapy to insulin). However, they have not included patient populations treated on all the background therapies specified in the scope (as we outline below).

For the assessment of dual therapy, the manufacturer has included only patients treated on a background of monotherapy with metformin and has excluded patients treated on a background of monotherapy with a sulfonylurea. They state two reasons for this: 1) the sulfonylurea sub-study of the DIA3008 trial[18] showed a high rate of hypoglycaemic events when sulfonylurea was used in combination with canagliflozin (though the ERG considers that this study does not meet the decision problem as there was no active comparator), and 2) they argue that a high proportion of patients treated with a sulfonylurea would not be eligible for treatment with canagliflozin as many patients receiving a sulfonylurea would have been treated with it rather than metformin as they have renal impairment (specifically those with a CrCl < 45 mL/min), which is also a contraindication for canagliflozin. Clinical expert advice to the ERG suggests that the manufacturer's rationale is not completely justified, as not all patients taking a sulfonylurea monotherapy would do so because metformin was contraindicated due to renal impairment. For example, some patients experience gastrointestinal problems on metformin and may therefore be prescribed a sulfonylurea instead. The clinical expert consulted by the ERG also stated that hypoglycaemia is an adverse event commonly associated with sulfonylureas and the high rate

of hypoglycaemia observed for dual therapy with sulfonylurea and canagliflozin is to be expected and is not a reasonable justification for excluding this patient group (though as stated above, the DIA3008 SU sub-study[18] does not meet the decision problem in the view of the ERG). While most patients treated with dual therapy will receive an add-on to metformin (which is the standard first-line therapy recommended by NICE[7]), the ERG suggests that the exclusion of patients treated on a background of a sulfonylurea is not reasonable, and that given the high prevalence of diabetes in the UK, this excludes a significant number of diabetes patients who may be treated in the NHS and who fall within the licenced indication outlined in the SmPC (i.e. canagliflozin can be used as an add-on therapy to sulfonylureas). We discuss the implications of this omission for the clinical and cost-effectiveness analysis in section 4.4 of this report.

For the assessment of triple therapy, the manufacturer considers patients inadequately controlled on a background of a) metformin and a sulfonylurea, and b) metformin and a thiazolidinedione. Other patient groups specified in the scope that are not included are: patients whose diabetes was inadequately controlled on dual therapy with metformin in combination with a DPP-4-inhibitor or a GLP-1 analogue, or with a sulfonylurea in combination with a thiazolidinedione, a DPP-4-inhibitor or a GLP-1 analogue.

[redacted] Patients on GLP-1-a and metformin would probably move to insulin, and a SGLT-2 inhibitor would only be considered as a third line agent with a DPP-4-i and metformin for a minority of patients in whom sulfonylurea is contraindicated. The manufacturer has not provided reasons for not including patients treated with a sulfonylurea in combination with a thiazolidinedione, a DPP-4-inhibitor or a GLP-1 analogue.

In summary, the population described in the decision problem covers most of the people with T2DM inadequately controlled on mono- and dual therapy in the NHS, who would have been treated on drug combinations including metformin (which is the standard first line therapy recommended by NICE).[7] However, the smaller proportion of patients treated with other agents, and who could potentially be eligible for treatment with canagliflozin according to its licenced indication, is not represented in the decision problem. In particular, the ERG considers that the omission of patients treated on a background of monotherapy with a sulfonylurea was not justified.

Table 1: Adult T2DM populations included in the manufacturer's decision problem

Treatment regimen	Population included in the decision problem	Matches NICE final scope?
Dual therapy	Adults with T2DM inadequately controlled on monotherapy with metformin	Yes, but does not include adults with T2DM inadequately controlled on monotherapy with a sulfonylurea, which was also stated to be a population of interest in the scope
Triple therapy	Adults with T2DM inadequately controlled on dual therapy with either : • metformin and a sulfonylurea • metformin and a thiazolidinedione	Yes, but does not include all the triple populations which were stated to be of interest in scope. Specifically, it omits adults with T2DM inadequately controlled on dual therapy with: • metformin in combination with a DPP-4-inhibitor or a GLP-1 analogue • a sulfonylurea in combination with a thiazolidinedione, a DPP-4-inhibitor or a GLP-1 analogue
Add-on therapy to insulin	Adults with T2DM inadequately controlled on monotherapy with insulin and standard of care	Yes. Scope states that population of interest is adults with T2DM inadequately controlled on insulin monotherapy or on insulin with up to two oral agents. Although, note that the manufacturer has not specified how many agents are used as part of "standard of care".

Intervention

The marketing authorisation for canagliflozin was granted in November 2013. In line with the final scope and the licenced indication, the intervention described in the decision problem is canagliflozin at a dose of 100mg or 300mg once daily in combination with oral anti-diabetic agents and/or insulin in patients whose diabetes was inadequately controlled on other agents/insulin. As detailed in the MS, the SmPC states that canagliflozin is initially administered at a dose of 100mg once daily, and this can then be increased to 300mg once daily if better glycaemic control is needed, patients have tolerated canagliflozin and the higher dose is not contraindicated.[17]

On MS p. 14 of the decision problem section of the MS, the manufacturer states that in clinical practice, it is expected that canagliflozin will be used in combination with metformin for dual therapy, in combination with metformin and a sulfonylurea for triple therapy and in combination with insulin plus standard of care. The MS focuses on these intervention combinations, plus triple therapy with canagliflozin, metformin and a thiazolidinedione, in the decision problem (MS p. 47) and the systematic literature review. These uses of canagliflozin only partly reflect the

licenced indication and NICE's final scope – which states that the aim of this appraisal is to examine the effectiveness of canagliflozin within its licensed indication. The SmPC does not limit use of canagliflozin to only dual and triple therapy regimens including metformin; it states that canagliflozin can be used as an add-on therapy to other anti-hyperglycaemic agents, including insulin. Therefore the use of canagliflozin specified in the decision problem is narrower than defined in the licenced indication and may not reflect all the potential uses of canagliflozin in clinical practice. Furthermore, the ERG notes that the combination of canagliflozin, metformin and thiazolidinedione may be of limited relevance to clinical practice. Clinical expert advice to the ERG is that the use of thiazolidinediones is declining due to concerns about weight gain, heart failure and osteoporosis.

Comparators

For the assessment of dual therapy with a combination of canagliflozin and metformin, the comparators stated in the manufacturer's decision problem are all those specified in the final scope: sulfonylureas, thiazolidinediones (although the scope more specifically states 'pioglitazone'), DPP-4-inhibitors, GLP-1 receptor agonists and dapagliflozin, all in combination with metformin. These comparators are appropriate for the NHS and reflect the drug combinations recommended by NICE in CG 87,[7] TA 203,[10] TA 248,[9] and TA 288.[8]

For the assessment of triple therapy with a combination of canagliflozin, metformin and a sulfonylurea, the comparators stated in the manufacturer's decision problem are: DPP-4-inhibitors and GLP-1 analogues, both in combination with metformin and a sulfonylurea. The manufacturer has therefore included two of the four comparators stated to be of interest in the final scope. These comparators are appropriate for the NHS and reflect the drug combinations recommended by NICE in CG 87,[7] TA 248,[9] and TA 203.[10] The comparators specified in the scope that have not been considered by the manufacturer are: thiazolidinediones (although, again, the scope more specifically states 'pioglitazone') and insulin, both combined with metformin and a sulfonylurea. The manufacturer states that these comparators have not been included as "insufficient clinical data were found to populate the NMA [network meta-analysis]" (MS p. 41). For the assessment of the triple therapy combination of canagliflozin, metformin and thiazolidinediones, the comparator is DPP-4-inhibitors combined with metformin and thiazolidinediones. Clinical expert advice to the ERG is that this combination is not generally used in clinical practice, and, as noted above, the use of thiazolidinediones is declining. The ERG therefore suggests that this comparator is of less relevance to clinical practice.

Comparators stated to be of interest in NICE's scope and that have not been included are: GLP-1-a and insulin, both combined with metformin and thiazolidinediones. The manufacturer has not included these as again they state that there was insufficient evidence to populate the NMA.

For the assessment of add-on therapy to insulin (canagliflozin in combination with insulin with or without metformin), the comparators stated in the manufacturer's submission are: DPP-4-i, GLP-1-a and dapagliflozin, all with or without metformin. The manufacturer states that although NICE does not recommend use of DPP-4-i or GLP-1-a in combination with insulin, these combinations are used in clinical practice. Clinical expert advice to the ERG concurs that these agents are used in clinical practice with insulin.

On the whole, the comparators outlined in the decision problem are appropriate for the NHS and are those recommended in NICE guidance. The most notable potential comparator not included (due to insufficient evidence) is insulin, metformin and sulfonylurea, which is the main triple therapy regimen recommended by NICE in CG 87,[7] but as noted above, clinical expert advice to the ERG indicated that this regimen is not commonly used in practice.

Outcomes

The manufacturer has included all the outcomes specified in the final scope in the decision problem:

- HbA_{1c} control
- Frequency and severity of hypoglycaemia
- Change in cardiovascular risk factors (including blood pressure and serum lipids)
- Weight change
- Complications of diabetes (for example, cardiovascular, renal and eye)
- Mortality
- Adverse effects of treatment (including genitourinary tract infection and urinary tract infection)
- HRQoL

The outcomes are appropriate and are clinically meaningful. The manufacturer has not omitted any clinically important outcomes.

Economic analysis
The economic analysis proposed in the decision problem matches the final scope and is appropriate for the NHS. The manufacturer has used a 40 year time horizon, which is sufficient for capturing differences in outcomes and costs for intervention in diabetes.

Other relevant factors
In line with the final scope, the decision problem states that subgroup analyses based on patients' baseline HbA_{1c} will be included in the submission (the manufacturer has used a HbA_{1c} level of ≥ 9% as this is the pre-defined category used in some of the trials, which the ERG agrees is reasonable). The manufacturer states that they have not included subgroup analyses by patients' baseline BMI, as suggested in the final scope, because the trial data showed that this was not associated with change in HbA_{1c} and other outcomes. The ERG again agrees that this is reasonable following clarification provided by the manufacturer.

The MS states that there are no known equity or equality issues and the ERG agrees with this.

3 CLINICAL EFFECTIVENESS

3.1 Critique of manufacturer's approach to systematic review

3.1.1 Description of manufacturer's search strategy

Summary
Overall, the literature searches in the MS are systematic and of a reasonable standard. They do contain some minor inconsistencies, though these are not deemed sufficient for amendment or re-run. There is good coverage of grey literature, especially for conferences and some hand searching was undertaken. Only one clinical trial registry, clinicaltrials.gov, was documented as having been searched. The ERG elected to run searches on some additional clinical trial registries and websites (see below).

Clinical effectiveness searches
The database selection matched the minimum NICE criteria (MS section 6.1 and MS Appendix 2). The MS reports the database host platforms used and the years searched. Searches were undertaken in February 2013 and updated in August 2013 (NB. The accompanying clinical systematic literature review (SLR) report by Amaris[3] does not mention that searches were

updated in August 2013, yet the same number of references are reported as having been screened in both the SLR report and the MS. The ERG would have expected there to be an updated number of screened references in the MS). A good range of relevant conference proceedings have been searched (eight in total).

The clinical effectiveness literature search (which was conducted jointly for the clinical SLR and the NMA) is clearly split into three logical sets to represent population/disease, intervention and trial design. There is a reasonable balance of database index descriptor terms (e.g. MeSH) and free text terms in the strategy. Syntax and Boolean operators are used correctly. A precision-optimised RCT search filter (similar to the one in the Cochrane Handbook) has been utilised and the search was restricted to English language. Search strategy terms have been successfully translated from Medline to Embase. In the Cochrane Central Register of Controlled Trials (Central) it is correct that no trials search filter has been applied as it is a trials database.

There is a minor discrepancy in the reporting of the choice of database host, as in MS appendix 2 section 10.2.3 Table 1 page 36 states “Ovid” whereas the text in MS section 6.1.1 Table 8 page 51, records “Medline and Embase were searched using the Embase.com interface”.

There were no details in the search strategy tables in MS appendix 2 of the number of hits per line, nor total recall, making it hard to quickly check the return of hits. The searching of in-house company databases is specified by the MS as not applicable, with no further details given. However, MS table 10 provides citations from canagliflozin clinical study reports ‘from the Janssen internal database’.

Only one on-going trials registry has been searched (ClinicalTrials.gov) on March 11th 2013 (MS appendix 10.2.1). It would have been advisable for this search to have been updated in August 2013 to synchronise with the update searches of the bibliographic databases. A more extensive search for on-going trials could have been undertaken using other trial registries. Hand searching of NICE, SMC, AWSMG was undertaken in Feb and updated in August 2013. The ERG conducted searches of the following on-going trials registries: UK Medicines Information (UKMi); National Horizon Scanning; European Medicines Agency (EMA); Food and Drugs Administration (FDA); Johnson & Johnson; UK Clinical Research Network Study Portfolio (UKCRN); controlled-trials.com multi register search; World Health Organisation International

Clinical Trials Registry Platform (WHO ICTRP); Australian and New Zealand trial Registry. The results of this search are discussed below.

Cost effectiveness searches

MS section 7.1, outlines the approach to the search for cost-effectiveness studies. The search strategies are provided in a separate report by the Swedish Institute for Health Economics (IHE).[4] They were undertaken in November 2012, with no mention of update searches. A variety of appropriate databases have been searched. There is a typo in text referring to the search strategy description - “sulfonlyurea” instead of sulfonylurea, however it appears correct in the search table in Appendix 1 of the IHE report.[4] A succinct cost filter has been used. The Pubmed search appears to have used only free text terms in any field rather than using a combination of free text and index terms. In Embase, index terms have been used to represent T2DM, however only free text terms have been used to represent cost effectiveness, although the terms would appear to be appropriate and correctly linked.

Cost of resource use searches

Cost of resource use searches were conducted in November 2012 and appear in a separate report by the manufacturer.[19] A wide range of co-morbidities have been included in the search and there is a balance of index terms and free text terms. The Medline search is reported in the text to have been conducted in Ovid, however the ERG notes that the syntax used suggests a different host platform was used.

Quality of life searches

The searches were conducted on consultancy by York Health Economics Consortium (YHEC) and contain an appropriate quality of life search filter to identify specific utility values such as EQ5D or SF-36. in diabetes or complications of the condition. There is a good balance of index and free text terms. Lines 5-16 are repeated perhaps unnecessarily again in lines 22-32, making sets 16 and 33 identical, although perhaps the rationale was for greater overtness for the linking of T2DM and quality of life and all types of diabetes and complications and quality of life. The anatomy of the search strategy is clearly explained. A good range of grey literature and clinical trials registries were also efficiently searched.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection

The inclusion and exclusion criteria are clearly stated on MS p. 53. The same exclusion and inclusion criteria were used to select studies for both the systematic literature review and the NMA. Therefore, the criterion regarding the type of intervention to be included ("at least one treatment of interest" [MS p. 53]) was wider than interventions/comparators specified in the decision problem and included all those stated to be of interest in the NICE scope, so that indirect comparisons of canagliflozin and the comparators listed in the scope could be made.

The manufacturer restricted inclusion to studies that focused on patients whose T2DM was inadequately controlled. This is wider than the populations described in the decision problem, which more specifically focused on patients inadequately controlled on mono-, dual and insulin therapy on specified combinations of agents as their background therapy. Studies of children were excluded, which is in line with the decision problem and licence, which is for use of canagliflozin in adults only.

In terms of outcomes, the manufacturer restricted inclusion to studies that had measured HbA_{1c}, fasting plasma glucose (FPG), blood pressure, body weight/BMI and the incidence of hypoglycaemia. These outcomes reflect those specified in the decision problem, except for FPG which was not stated to be of interest in the decision problem. The ERG considers that FPG provides little clinically meaningful information as an outcome, as it is an acute parameter that easily fluctuates and only reflects an individual's glycaemic control at one moment in time.[20] However, results for FPG are reported in the SLR publication accompanying the MS[3] (reference 72 in the MS), and, in line with the decision problem, are not reported in the MS. The manufacturer omitted other outcomes stated in the decision problem and scope from their inclusion criteria, specifically: change in cardiovascular risk factors other than blood pressure, complications of diabetes, mortality, adverse effects of treatment and HRQoL, though they do report results of these for the canagliflozin trials included the MS. This means that the review may not have identified all relevant studies that measured these outcomes.

The manufacturer also restricted study selection to only RCTs and studies of 12 weeks or more duration, and excluded studies not in English, which the ERG agrees is reasonable. No limits were placed on inclusion regarding the methodological quality or risk of bias of the RCTs.

Setting was not used as an inclusion criterion and this reflects the decision problem and the NICE scope.

Overall, the ERG considers that the inclusion criteria were appropriate, but in parts were wider than the decision problem. The restriction of the outcomes criteria to only some of those specified in the decision problem and scope means that potentially relevant studies may have been missed for the systematic literature review and the NMA.

Additional inclusion criteria were applied to screen studies for the NMA (see MS p. 112). Trials including special populations (renal impairment, CVD or risk of CVD) (for the analysis of metformin, MET+SU and MET+TZD background), or patients on mixed background therapies or therapies not recommended in the UK (specifically insulin degludec, liraglutide 1.8mg and lixisenatide), and were excluded. Special populations were excluded to reduce heterogeneity (with the exception of the DIA3008 insulin sub-study which included patients at increased risk of CVD, on the basis that this was the only canagliflozin trial available to be included in the insulin background NMA). Trials reporting outcomes at 26 or 52 weeks ± 4 weeks (for the base case) or at 26 or 52 weeks ± 10 weeks (for the sensitivity analyses) were included. For the triple therapy networks trials that did not examine outcomes at 26 +/- 10 weeks were excluded. Please see Table 4 for the ERG's appraisal of the appropriateness of these criteria.

A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart is available in Figure 1 in the accompanying SLR report[3] (p. 26). This shows the number of references included and excluded at each stage of the review, as well as reasons for exclusion at the initial and full text screening stages. However, it is unclear how the 208 citations that were identified for data extraction and use in the systematic literature review in the accompanying SLR report (MS p. 54 and SLR[3] p. 26) were further refined to the 11 canagliflozin studies detailed on MS p. 57 that were eventually considered for inclusion in the MS. The manufacturer has not detailed the process or criteria used for further study selection. The ERG requested clarification from the manufacturer about this. In their response, the manufacturer stated that Janssen had provided 10 studies (totalling to the 11 studies identified when including the two sub-studies of one RCT) to Amaris (who conducted the SLR on behalf of the manufacturer), and that these had not been identified through the original searches, as the studies were unpublished at the time the searches were conducted. These are shown in the PRISMA diagram in the cell for "additional papers identified through other sources". They stated that the

studies were screened using the same process as for all the other papers. The ERG considers the process for selecting these studies to be acceptable.

Furthermore, a total of 648 full texts were screened for eligibility, and the total number of references 'discarded as did not meet the inclusion criteria' sums to 434 by the ERG's calculation. This should leave 214 publications for inclusion in the systematic literature review detailed in the accompanying report,[3] rather than 208 publications as specified. The ERG asked the manufacturer to clarify this discrepancy, and they responded stating that there were an additional 6 studies (5 of which were conference abstracts) that did not meet the criteria for the clinical systematic review but which were included in the sensitivity analysis of the NMA (citations provided).

The manufacturer has not discussed any potential biases in their study selection.

3.1.3 Identified studies

The MS identified 11 Phase 3 RCTs (reported in 17 publications in addition to a number of conference abstracts) as relevant to the MS (MS p. 54), including two sub-studies of one trial (DIA3008) (MS p.55). The manufacturer did not identify any relevant non-RCTs, as they restricted inclusion to RCTs only. The manufacturer supplied the ERG with electronic copies of the CSRs and all the linked publications for 10 of the 11 identified RCTs. They did not supply a copy of the CSR for the DIA3014 trial (the "Asian study"), which was the only publication identified for this trial. All of the RCTs were sponsored by Janssen Research & Development, LLC. Of the 11 RCTs, the manufacturer excluded five from further consideration and gave reasons for this on MS p. 57 and p. 58 (though note, these trials were included in various pooled sets of analyses for adverse events reported in the MS). This resulted in six RCTs finally being included in the review (MS p. 59) (DIA3009,[21] DIA3006,[22] DIA3015,[23] DIA3002, DIA3012, and the DIA3008 insulin sub-study). The six studies identified for inclusion by the manufacturer meet the inclusion criteria, but three are not relevant to the decision problem or final scope, as we detail below.

The ERG checked the relevance of the originally identified 11 RCTs against the final scope and decision problem and Table 2 shows our assessment of whether or not these trials are relevant. The ERG agreed with the five exclusions the manufacturer had made, but also identified that

three of the included RCTs (DIA3002, DIA3012, and the DIA3008 insulin sub-study) were not relevant, for the reasons detailed in Table 2 – namely the absence of active comparators.

Therefore, the following trials identified by the manufacturer were assessed by the ERG as relevant to the review, and we have focused our review of the evidence submitted by the manufacturer to these:

- DIA3009[21] - CANagliflozin Treatment And Trial Analysis-Sulfonylurea SGLT2 Add-on to Metformin Versus Glimepiride (CANTATA-SU)
- DIA3006[22] - CANagliflozin Treatment and Trial Analysis - DPP-4 Inhibitor Comparator (CANTATA-D)
- DIA3015[23] - CANagliflozin Treatment And Trial Analysis - DPP-4 Inhibitor Second Comparator (CANTATA-D2)

Two of these three trials examined dual therapy with canagliflozin (at a dose of 100mg or 300mg) among patients treated on a background of metformin, with DIA3009[21] comparing it to sulfonylurea (glimepiride, titrated up to a maximum dose of 6mg or 8mg depending on local regulations – see trial CSR p. 36) and DIA3006[22] comparing it to sitagliptin (a DPP-4-i) (100mg). The other trial (DIA3015[23]) examined triple therapy with canagliflozin (300mg) compared to sitagliptin (100mg) among patients treated on a background of metformin and a sulfonylurea. There were no relevant add-on to insulin therapy studies.

Table 2: The manufacturer and ERG's assessment of whether the identified 11 Phase 3 RCTs were relevant to the systematic literature review (as defined in the scope and decision problem)

Trial	Intervention	Patient population – treatment background	Manufacturer's assessment of trial relevance	ERG's assessment of trial relevance for this report
DIA3009	Dual therapy - canagliflozin 300mg and 100mg versus SU (glimepiride)	MET	Include	Include
DIA3006	Dual therapy - canagliflozin 300mg and 100mg versus placebo or versus DPP-4-i (sitagliptin)	MET	Include	Include
DIA3008 SU sub-study	Dual therapy - canagliflozin	MET + SU	Exclude	Exclude - not relevant to the decision problem as no active

	300mg and 100mg versus placebo			comparator
DIA3015	Triple therapy - canagliflozin 300 mg versus sitagliptin 100 mg	MET + SU	Include	Include
DIA3002	Triple therapy - canagliflozin 300 mg and 100 mg versus placebo	MET + SU	Include	Exclude - not relevant to the decision problem as no active comparator
DIA3012	Triple therapy - canagliflozin 300mg and 100mg versus placebo/sitagliptin	MET + TZD	Include	Exclude – purpose of the study was to compare the efficacy and safety of two different doses of canagliflozin with placebo. Patients received placebo up to 26 weeks and then switched to sitagliptin for the next 26 weeks to maintain the double-blind design. Trial publication and MS do not report efficacy data for this arm up to 52 weeks.
DIA3008 insulin sub-study	Add-on to insulin therapy - canagliflozin 300mg and 100mg versus placebo	Insulin and standard of care	Included	Exclude – trial compares two different doses of canagliflozin and placebo in a sub-population of patients receiving insulin ≥30 IU/day and each group was stratified at randomisation for patients receiving different background therapy combinations of insulin and other agents or insulin alone. [redacted] Around a third of patients randomised to the placebo group had no background therapy apart from insulin. Thus the study does not match the comparator specified in the scope, “ One or more oral anti-diabetic agents (in combination with insulin)”.
Special population studies				
DIA3010 “Older subjects”	Placebo or canagliflozin 100mg or 300mg	Any approved oral or injectable treatment	Exclude	Exclude

DIA3004 "Renal impairment"	Placebo or canagliflozin (100mg or 300mg)	Any approved oral or injectable treatment	Exclude	Exclude
DIA3014 "Asian study"	Unclear	All lines of therapy	Exclude	Exclude – a copy of the trial publication was not provided to ERG, but the manufacturer stated that patients in this trial had been in receipt of "all lines of therapy" (MS p. 58) as their background treatment so it is likely that this study is not relevant to the decision problem or the scope
DIA3008 "CV safety study"	Placebo, canagliflozin 100 mg, or canagliflozin 300 mg	Patients with increased cardiovascular risk	Exclude	Exclude

The MS provides an overview of the trial designs and interventions used in DIA3009[21], DIA3006[22] and DIA3015[23] trials in MS p. 60, and p. 61, Table 6 in Appendix 3, and Tables 7 and 8 in Appendix 4, MS Figure 5 (p. 60) and Figures 1 to 6 in Section 10.4.2 of Appendix 4. The patient inclusion and exclusion criteria are provided in Table 11 (MS p. 63 and 64) and Table 12 (MS p. 65). Consort flow charts showing the number of patients eligible, randomised, and who completed each stage of the trials or dropped out are provided in Figures 7 to 12 in Section 10.4.5 in Appendix 4 (p. 69 to 71). Baseline characteristics are shown in Tables 9 and 11 in section 10.4.3. The manufacturer has provided a list of the outcomes reported in the review and clinical justification for their inclusion in the review on MS p. 67 to 68. Table 13 in section 10.4.4 of Appendix 4 (p. 66 to 68) details the primary and secondary outcomes measured in each study and includes information about the reliability, validity and current use of each measure in current clinical practice. Information about the statistical analyses in the trials is provided in Table 14 (MS p. 69 to 72), including power and sample size calculations, description of the mITT analysis, and information about the subgroup analyses are available on MS p. 73. In summary, the ERG considers that the MS provides an adequate level of detail about the canagliflozin trial characteristics.

The MS states that baseline patient characteristics were similar across treatment arms within the DIA3009,[21] DIA3006[22] and DIA3015[23] trials (MS p. 67) and the ERG agrees with this conclusion. The MS also states that patient baseline characteristics were similar between the dual therapy studies (DIA3009[21] and DIA3006[22]), except that the duration of diabetes was slightly higher in DIA3006 than in DIA3009 (MS p. 67), and the ERG agrees with this. The ERG

notes that, as might be expected, proportionally more patients in the triple therapy study (DIA3015) than in the dual therapy studies had an HbA_{1c} of > 9% and had experienced diabetes-related complications. These patients also had a longer duration of disease.

The MS states that there is currently only one ongoing study of canagliflozin for T2DM: the CANVAS trial (DIA3008) of patients with a diagnosis of T2DM who have an elevated cardiovascular risk. This a safety study carried out in 4,330 patients which will report in 2017. The ERG notes that this trial is not relevant to the decision problem or scope, as it compares canagliflozin plus standard of care and placebo plus standard of care.

The MS appears to have identified all relevant head-to-head trials of canagliflozin. Other than this, the ERG identified one further ongoing RCT which is possibly relevant from searches of the UK Clinical Research Network (UKCRN) database (UKCRN study ID 7616). The study is a four-arm Phase 3 RCT comparing the efficacy, safety and tolerability of canagliflozin with placebo and sitagliptin in the treatment of patients with T2DM that is inadequately controlled on metformin and sulfonylurea. The ERG were unable to identify any further information about this study from internet searches and were unable to establish contact with the lead investigator, but note that it is funded by the manufacturer (Johnson & Johnson Pharmaceutical Research and Development).

Finally, although the ERG deemed the three placebo-controlled trials DIA3002, DIA3012, and DIA3008 (insulin sub-study) as relevant to this report, it is noted that these trials are included in the NMA with the placebo arms used to connect treatments together indirectly. Furthermore, the economic model requires treatment effectiveness versus placebo to be specified. Whilst the ERG has not focused on these three trials in this report for the reasons stated above, given their inclusion in the NMA and the use of placebo comparison in the economic model, there is a brief discussion of their methodological quality in section 3.1.4.

Superseded – see erratum

3.1.4 Description and critique of the approach to validity assessment

The MS quality assessed all the included trials, and they used the NICE criteria (MS section 6.4). The canagliflozin clinical trials appear to be of good methodological quality and the ERG generally agrees with the manufacturer's assessment of study quality (Table 3).

Table 3: Manufacturer and ERG assessment of trial quality

		DIA3009	DIA3006	DIA3015
1. Was randomisation carried out appropriately?	MS:	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
Comment: Computer generated randomisation				
2. Was concealment of treatment allocation adequate?	MS:	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
Comment: An interactive web/voice response system was used				
3. Were groups similar at outset in terms of prognostic factors?	MS:	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
Comment:				
4. Were care providers, participants and outcome assessors blind to treatment allocation?	MS:	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
Comment: In all three trials blinding was maintained throughout the core double blind period until database lock whereby the study was unblinded to permit completion of CSR for regulatory filing (though Centre, patients and local sponsor personnel remained blinded). Unblinding could also occur if rescue mediation was required.				
5. Were there any unexpected imbalances in drop-outs between groups?	MS:	No	No	No
	ERG:	No	No	No
Comment: There was a slightly higher rate of discontinuations in the active comparator arms of the trials (e.g. sitagliptin, and glimepiride)				
6. Is there any evidence that authors measured more outcomes than reported?	MS:	No	No	No
	ERG:	No	No	No
Comment: Generally all outcomes specified in MS Appendix 4 (Tables 7 and 8) are reported, with following exceptions: time to initiation of rescue therapy (DIA3009 and DIA3006); time to discontinuation; coefficient of durability (DIA3015); and time to develop an increase in HbA_{1c} ≥0.3% from week 26 to week 52 (in patients with decrease from baseline in HbA_{1c} ≥ 0.4%)(DIA3015). Limited data are reported for exploratory outcomes such as HRQOL and PROs (all three trials).				
7. Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	MS:	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
Comment: Primary efficacy analysis was performed in the modified intent-to-treat (mITT) population (randomised participants who received ≥1 dose of study drug) using a last observation carried forward (LOCF) approach. The number of randomised patients not dosed (and therefore excluded from the mITT) was minimal (DIA3006 n=0; DIA3009 n=2; DIA3015 n=1). The MS does not discuss the appropriateness of using the LOCF approach in the evaluation of diabetes treatment (e.g. the assumption that change in HbA_{1c} would be maintained following treatment withdrawal). A per-protocol analysis was performed as supportive analyses. Only mITT results are reported in the MS (some per-protocol results are reported in the trial journal publications).				

The manufacturer judged the quality of three placebo controlled RCTs included in the MS (DIA3002, DIA3012 and the DIA3008 insulin sub-study) to be similar to the three active comparator trials (MS Table 16). As discussed above, the ERG has not focused on the placebo controlled trials in this report, but does note their inclusion in the NMA and the use of placebo

comparison in the economic model. Whilst the ERG has not undertaken an independent quality assessment of these three trials it is noted that they are similar in design to the active comparator trials, having been conducted as part of the canagliflozin clinical trial programme. It would therefore be reasonable to assume they are of similar good quality to the other trials in the programme.

3.1.5 Description and critique of manufacturer's outcome selection

The outcomes addressed in the manufacturer's decision problem include all of those specified in the scope, and are generally reported in the manufacturer's systematic review of the literature (though see below).

Generally the outcomes measured in the trials are reported in the MS, however see the specific exceptions noted in Table 3 (Manufacturer and ERG assessment of trial quality). Body Mass Index (BMI) is not reported in the MS but is included in the NMA.

MS Appendix 10.4.1 (Table 7) reports that a variety of exploratory HRQoL measures and Patient Reported Outcomes (PRO) were assessed using questionnaires, in the trials. These included established generic instruments such as the EQ-5D, SF-36, and Current Health Satisfaction Questionnaire (CHES-Q) and disease-specific instruments: Impact of Weight on Quality of Life Lite (IWQOL-Lite); Multi-dimensional Diabetes Questionnaire (MDQ); Diabetes Utility Index (DUI); Diabetes Intention, Attitude, and Behavior Questionnaire (DIAB-Q).

The HRQoL and PRO outcomes are not reported in the main submission document (or in the trial journal publications), however MS Appendix 14 reports various sets of analyses of these instruments, some of which are pooled across trials, and some which are reported by trial (e.g. mainly trial DIA3009). Regression analyses are also reported to identify predictors of HRQoL. A description of the HRQoL and PRO instruments with supporting references is also provided. Some of the instruments are reported to have been validated, with some having been developed by the manufacturer.

Superseded – see erratum

3.1.6 Description and critique of the manufacturer's approach to trial statistics

The MS reports the results for the relevant outcomes in the DIA3009,[21] DIA3006[22] and DIA3015[23] trials in MS section 6.5.3.1 (p. 79 to 108) and Section 10.6 of Appendix 6 (specifically, change in HbA_{1c}, blood pressure, body weight and lipid levels). No interim data for the DIA3009,[21] DIA3006[22] and DIA3015[23] trials are presented. Trial results for weight change are only reported in kg in the SLR and not in terms of BMI, which was also reported in the trial CSRs. However, BMI results are reported in the NMA. Limited results for the HRQoL and PRO outcomes are reported in MS Appendix 14.

The MS states that analyses of the change in HbA_{1c}, blood pressure, body weight and lipid levels endpoints were performed using ANCOVA. The MS reports that non-inferiority and superiority analyses were conducted for the HbA_{1c} outcome in each of the three trials. The trials used a pre-specified non-inferiority margin of 0.3% for the upper limit of the two-sided 95% confidence interval for each comparison. If non-inferiority was shown superiority was concluded when the upper limit of the 95% CI was less than 0.0%. The ERG notes that the non-inferiority analyses were adequately powered in each of the trials. Units of measurement, effect sizes, measures of variance and the number of patients included in each analysis are provided. The results presented are based on the mITT population (defined as all randomised patients who took at least one dose of the study drug), which used a LOCF approach to impute missing data. As mentioned in section 3.1.4 of this report, the MS does not discuss the appropriateness of the LOCF approach in the evaluation of diabetes treatment (e.g. the assumption that change in HbA_{1c} would be maintained following treatment withdrawal). Per-protocol analyses were also reported in the original trial publications, with the per-protocol populations defined as all patients who completed follow-up, did not need glycaemic rescue therapy and who had no protocol violations. The manufacturer has only presented the mITT analyses in the MS and has not provided any commentary on any differences or similarities between the per-protocol and mITT analyses. In our presentation of the trial results in this report (see section 3.3), the ERG has commented on whether or not the mITT and per-protocol population analyses were consistent.

For all three trials, the manufacturer has presented a post-hoc analysis of the proportion of patients achieving an HbA_{1c} of less than 7.5%, the current NHS target in the Quality Outcomes Framework (QOF),[24] rather than the pre-defined endpoints in the trials of patients achieving less than 7% or less than 6.5% at week 52.

[redacted]

The manufacturer presents the results for the sub-group analyses by baseline HbA_{1c} in DIA3006[22] and DIA3015[23] on MS p. 107, using the following categories: <8%, 8% to <9% and ≥9%. Subgroup analyses by baseline HbA_{1c} were not conducted in DIA3009.[25] The subgroup analysis results by baseline BMI are not presented, with the MS just stating that "there were no significant difference [sic] in terms of change in mean HbA_{1c} over time by BMI groups" (MS p. 107 and p. 108).

In summary, the manufacturer's approach to trial statistics is on the whole appropriate, but the ERG considers that the MS should have discussed the appropriateness of the LOCF approach to imputing missing data in the mITT analyses and made apparent any differences and similarities between the mITT and per-protocol analyses for the HbA_{1c} outcome. Furthermore, the ERG suggests that the post-hoc analyses of the proportion of patients achieving an HbA_{1c} of less than 7.5% should be interpreted with caution.

3.1.7 Description and critique of the manufacturer's approach to the evidence synthesis

A narrative synthesis is provided of the canagliflozin trials, with data presented in tables and figures with some textual commentary (MS Section 6.5).

The main form of synthesis in the MS is the NMA. Separate networks were constructed for: metformin background at 26 weeks, metformin background at 52 weeks, metformin and sulfonylurea background at 26 weeks, metformin and pioglitazone background at 26 weeks, and insulin (with or without oral antidiabetic drugs (OAD)) background at 26 weeks.

Meta-analyses of direct evidence (i.e. two or more trials of a comparison) were conducted, when possible, for each of the networks of the NMA. Direct meta-analysis was conducted using a frequentist model. The inverse variance method was used, or in cases of zero events the Mantel-Haenszel method was used (MS section 6.7.5.3). Heterogeneity was suspected if Cochran's Q test was significant (at <10%), or I^2 was higher than 50%. These are standard methods for assessing the likelihood of heterogeneity in meta-analysis.[26] The results of the

direct comparisons are provided in the accompanying SLR report,[3] but do not appear to be in the main MS.

Outcomes of relevance to the economic model (i.e. change in HbA_{1c}, weight, BMI and SBP, and incidence of hypoglycaemia) are reported in the main submission (see MS p. 111) and further outcomes such as FPG are reported in the accompanying SLR report[3]. However, as noted later in this report (section 4.2.4) the economic model appears to use data from the NMA for HbA_{1c} and BMI, but for other outcomes input data is taken from the NMA or from specific clinical trials. Consequently, of all the outcomes included in the NMA the results for HbA_{1c} and BMI have most influence on the cost-effectiveness results in the submission.

Table 4 reports the ERG's critical appraisal of the manufacturer's NMA. In general the NMA is judged to be of reasonable quality according to the criteria, with a few caveats as noted in the table. The key caveats are:

- For the triple therapy networks, studies which did not measure outcomes at 26 +/- 10 weeks were excluded from the NMA after full text screening. This raises the possibility that some of those excluded could have been used to construct 52 week networks for triple therapy. This would have meant less reliance on 26 week effectiveness data in the economic model, and therefore the assumption of maintenance of treatment effects between 26 and 52 weeks (see point 4 in Table 4).
- Critical appraisal of included studies appears to have been conducted by one reviewer and only a 20% sample checked by a second. From a brief examination of the manufacturer's tabulated critical appraisal judgments it appears that the methodological quality is reasonable, however, for a number of studies quality was judged to be unclear (see point 7 in Table 4).
- Some of the networks include few studies, particularly for triple therapy. The volume of direct evidence for some of the loops in the networks was low (approximately 2 trials for each loop, typically), with several treatments connected directly by only one trial (see point 8 in Table 4).
- Statistically significant heterogeneity was noted in a number of the direct comparisons across the networks, and was suggested to be caused by differences in variables such as HbA_{1c} and weight/BMI at baseline between studies (see point 9 in Table 4).

Superseded – see erratum

- It is not stated whether the inconsistency tests reported were adjusted to take account of multi-arm studies or multiple testing in networks with multiple loops. Very little information is given on the methods adopted for these tests, which do not appear to follow the methods outlined by the Decision Support Unit (DSU).

Table 4: ERG appraisal of Network Meta Analysis (NMA)

Appraisal criteria	Criteria met
A. CONCEPTUAL BASIS	
1. Is a justification given for conducting an MTC?	No. An explicit rationale is not given. However, given the number of comparators included in the scope it would be expected that canagliflozin would not yet have been directly compared to all of them, hence the justification for indirect comparison.
B. SYSTEMATIC PROCESSES	
2. Is a comprehensive and transparent search strategy reported?	Yes. The manufacturer conducted one overall literature search (MS section 6.1, and MS Appendix 10.2) to inform their clinical SLR of canagliflozin trials (MS sections 6.2 to 6.5) and their NMA (MS section 6.7). The ERG considers the search to be generally comprehensive (see ERG report section 3.11).
3. Are inclusion / exclusion criteria adequately reported?	Yes. MS Table 9 specifies the inclusion criteria for the clinical systematic literature review, and additional criteria for the NMA are specified in MS section 6.7.2.1. The additional criteria were proposed to ensure comparable pooling across trials (e.g. in terms of background therapy, trial length, patient co-morbidities). The ERG considers that the additional criteria are generally appropriate. The criteria exclude treatments not recommended in the UK by NICE or the Scottish Medicines Consortium.
4. Is the number of included /excluded studies from the MTC reported, with reasons for exclusions?	Yes, Figure 1 of the accompanying SLR[3] shows a PRISMA diagram, reporting the number of studies included (n=64 in the base case, plus a further in 21 in sensitivity analysis) and excluded (n=123) from the NMA and the reasons for exclusion. The manufacturer supplied a list of references excluded from the NMA at the request of the ERG. The ERG notes from the PRISMA flow chart that a total of 51 references were discarded from the NMA for not meeting the inclusion criteria. Specifically, studies with a time assessment not within the ranges of 26 weeks (+/- 10 weeks) or 52 weeks (+/- 10 weeks - only for the metformin background) were excluded. It is possible that some of these discarded studies could have been used to construct 52 week networks for triple therapy (i.e. MET and SU, and MET and TZD) and insulin therapy background, though the ERG has not conducted a thorough check of the list of excluded references provided due to time constraints.
5. Is a visual representation of the data networks provided?	Yes. MS Figures 34 to 38 illustrate the evidence networks for the five separate networks. The accompanying SLR report[3] provides separate network diagrams for each outcome included in each network.
6. Are the data from included studies extracted and tabulated?	Yes. Data were extracted by one reviewer. All data included in the meta-analysis were checked by a second analyst. Tables 48 to 104 in MS Appendix 10.7 provide data on study and patient characteristics, and data inputs into the NMA. The ERG consider that the level of detail provided is generally adequate.
7. Is the quality of the included studies assessed?	Yes. The manufacturer used the NICE recommended critical appraisal criteria as also used for the canagliflozin trials in MS section 6.4. The results of the appraisal are in MS Appendix 10.7.8 (Tables 105 to Table 108). No commentary or overall summary of study methodological quality is provided. It should be noted that critical appraisal appears to have been conducted by one reviewer and only a 20% sample checked by a second. The ERG considers this a limitation of the manufacturer's systematic review (see section 3.2). The ERG has not performed an independent critical appraisal of the studies included in the NMA due to the

Superseded - see Draft

	large number of studies. From a brief examination of the manufacturer's tabulated critical appraisal judgments it appears that the methodological quality is reasonable, however, for a number of studies quality was uncertain due to deficiencies in reporting. Studies that the manufacturer considered to be lower quality (defined as not double blinded) were excluded in sensitivity analysis.
C. STATISTICAL ANALYSIS	
8. Are the statistical procedures adequately described and executed?	Yes. The MS reports details of the type of model (Bayesian hierarchical), inconsistency assessment, adjustments for multi-arm and cross over trials, procedures for selection of random and fixed effect models, assumptions and procedures for binary and continuous outcome data, model simulation procedures (e.g. number of iterations); methods of handling missing data; and other assumptions. Nodes of the NMA were treatment and dose-specific with the exception of insulin, and metformin and SU dual therapy. A class effect was assumed for insulin in line with NICE guidelines for management of T2DM. One of the main assumptions is that background therapies are important treatment effect modifiers, hence the decision to model separate networks by background therapy (i.e. MET; MET+SU; MET+TZD; insulin). This appears to be a reasonable assumption given that, for example, incidence of hypoglycaemia is higher with certain drugs than others (e.g. SU, insulin). To reiterate, however, the manufacturer has not created networks for all background therapies in the scope (e.g. SU monotherapy), as discussed earlier in the ERG report. Furthermore, not all comparisons in the decision problem were made due to insufficient evidence on certain outcomes to create a network. This mainly applied to the MET+TZD background where there were few available clinical trials. The ERG noted that the some of the networks are quite small, particularly for triple therapy. Also, the volume of direct evidence for some of the loops in the networks was low (approximately 2 trials for each loop, typically), with several treatments connected directly by only one trial. The results of the NMA in places are therefore reliant on methodological quality of individual studies, and as stated above, there are some uncertainties regarding methodological quality.
9. Is there a sufficient discussion of heterogeneity?	Yes. The MS states various assumptions to increase comparability of studies in the NMA (i.e. separate networks by background therapy; inclusion of trials with similar outcome assessment periods; exclusion of trials considered as a potential source of heterogeneity in sensitivity analysis). Fixed effect (FE) and random effects (RE) models were developed and the one associated with the lowest Deviance Information Criterion (DIC) was selected (MS section 6.7) (see also below). For pairwise direct comparisons statistical heterogeneity was assessed through Cochran's Q test and the I^2 statistic. MS Appendix 10.7.11 discusses cases of statistical heterogeneity amongst the direct comparisons. The accompanying SLR report[3] provides further discussion of suspected heterogeneity for direct pairwise comparisons and provides forest plots for comparisons associated with a suspected heterogeneity (based on I^2>50%) to visually assess the heterogeneity between results from individual trials and pooled analyses (though only for HbA_{1c} in the metformin background network at 26 weeks). Statistically significant heterogeneity was noted in a number of the direct comparisons across the networks, and was suggested to be caused by differences in variables such as HbA_{1c} and weight/BMI at baseline between studies.
10. Is the type of model used (i.e. fixed or random effects) reported and	Yes. Use of a fixed or random effects model was governed by the DIC (a measure of the goodness of fit of models), and the model with the lowest DIC was selected. In the majority of cases a fixed effect model was used in the base case

justified?	analysis. When a random effects model was used in the base case the fixed effect model was reported as sensitivity analysis (though not vice versa) (MS Page 120, and MS Table 21). However, the ERG is not convinced by the decision adopted in the MS to select between random and fixed effects models based on the DIC. While this rule may be considered to be generally appropriate (see DSU TSD 2,[27] page 14) it is not clear that the evidence networks presented in the MS are adequate to support this, with the majority of links in each network being based on a single trial.
11. Was sensitivity analysis conducted?	Yes. MS section 6.7.5.12 reports a number of parameters varied in the sensitivity analysis, including addition of trials with longer outcome assessment timepoints, and omission of poorer quality trials and trials with a potential source of heterogeneity or inconsistency.
12. Is any of the programming code used in the statistical programme provided (for potential verification?)	Yes. The ERG have checked the WinBUGS code for the NMA reported in the MS, Appendix 10.7.10 of the MS. These conform with example code provided in DSU Technical Support Documents (TSD 2[27]) and include appropriate adjustments for multi-arm studies. Appropriate models have been adopted for fixed effect and random effect models both of continuous outcomes (HbA_{1c}, weight, BMI, SBP) and of binomial data. The ERG checked a sample of the input data used in the NMA (tabulated in Appendix 10.7.7 of the MS) and re-ran a sample of the analyses producing results consistent with those presented in the MS.
D. PRESENTATION AND INTERPRETATION OF THE EVIDENCE	
13. Is there a tabulation/ illustration of results for each intervention and for each outcome?	Yes. MS section 6.7.6 provides forest plots for treatments by outcome measure, in terms of mean differences or OR. Note that the forest plots do not include all of the individual treatments within each of the drug classes that were included in the NMA. However, MS Appendix 10.7.14 tabulates results for all of the treatments included in the NMA, in terms of mean differences, SD, credible interval, and the probability that canagliflozin was more effective for outcomes in relation to other comparators using the Surface Under the Cumulative Ranking (SUCRA). The accompanying SLR report[3] also reports the probabilities of ranking treatments (e.g. the probability that canagliflozin is ranked 1st, 2nd etc). The main submission document only reports outcomes used in the cost-effectiveness model (i.e. change from baseline in HbA_{1c}, weight, BMI and SBP and incidence of all hypoglycaemic events). However, some additional outcomes are reported in the MS appendix and in MS reference 72 (proportion of patients reaching HbA_{1c} <7%; FPG).
14. Is there a narrative commentary on the results?	Yes. Results are summarised in bullet form in the MS section 6.7.6 with accompanying forest plots. MS appendix 10.7.14 and MS reference 72 both provide tabular results, but MS reference 72 provides more narrative commentary of these results.
15. Does the discussion of the results reflect the data presented?	Yes
16. Have the authors commented on how their results compare with other published studies (e.g. MTCs), and offer any explanation for discrepancies?	No.
17. Have the authors discussed whether or not there are any differences in effects between the direct and indirect	Yes. The manufacturer conducted inconsistency assessments for each loop of each network by performing an indirect comparison and comparing it to the direct evidence from the NMA (citing a DSU Technical Support Document 4[28] on methods for handling inconsistency) (see MS section 6.7.5.8). Inconsistent loops were identified in the MET (52 weeks) and MET+SU networks. Trials considered

evidence?	to cause inconsistency were excluded in sensitivity analysis (see MS appendix 10.7.12 and 10.7.13) While DSU Technical Support Document 4[28] does indeed indicate that comparisons of direct and indirect evidence (the latter derived using the method described by Bucher et al[29]) may be used for assessing consistency they raise concerns over their use in networks containing multi-arm trials (such as DIA3006 and DIA3009) and the multiple testing required in networks with multiple loops. Neither the MS nor the accompanying SLR report[3] state whether the inconsistency tests reported were adjusted to take account of multi-arm studies or multiple testing in networks with multiple loops. Both reports provide very little information on the methods adopted for these tests, which do not appear to follow the methods outlined in the DSU document.[28]

3.2 Summary statement of manufacturer's approach

The quality of the manufacturer's systematic review is reasonable with some exceptions (see below), based on the CRD criteria (see Table 5). The MS provides a high level of detail on the procedures followed and the evidence included, though much of this detail can only be found in the appendices and also in a separate SLR report.[3]

There is a slight discrepancy in the description of the screening process for titles and abstracts between the MS and the separate SLR report[3] The MS states that titles and abstracts were screened by a first reviewer and a second reviewer quality checked the assessment. However, the SLR report[3] mentions that titles and abstracts 'were reviewed by two independent reviewers' (page 17), rather than them being checked (which would imply that the person checking would have seen the first reviewer's screening decision).

The MS does not explicitly state the process for screening full publications for inclusion in the systematic review. However, the SLR report[3] reports that screening decisions were quality checked by a second reviewer, with discrepancies resolved through discussion between the two reviewers (page 17).

Data extraction (and, it is implied, quality assessment) was conducted by one reviewer and a second reviewer performed a quality check of 20% of the publications included in the systematic literature review. The CRD recommend that as a minimum, one researcher should extract the data with a second researcher independently checking the data extraction forms for accuracy and detail. The ERG considers that the data extraction of all of the publications should have been checked, rather than just 20% of them. The MS reports that all data included in the meta-

Superseded – see erratum

analysis were checked by a second analyst. This is appropriate given the possibility of human error in data entry and consequent impact on the meta-analysis results, particularly given the prominence of the NMA in the MS. The CRD also recommend that quality assessment is performed independently by two reviewers. The ERG notes that this does not appear to have been done in the MS and given that quality assessment can involve subjective interpretations this is a significant limitation of the manufacturer's systematic review.

The submitted evidence only partially reflects the decision problem in the MS for the reasons already outlined (see ERG sections 2.3 and 3.1.2). Overall there is a low chance of systematic error in the systematic review of the MS based on the methods employed, with the caveat about the assessment of quality assessment of studies.

Table 5: Quality assessment (CRD criteria) of MS systematic review

CRD Quality Item	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	Yes
2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified	Yes
3. Is the validity of included studies adequately assessed?	Partial. The MS uses the NICE recommended criteria. The quality of the canagliflozin RCTs is generally good and the ERG agrees with the manufacturer's assessment, though with a few caveats (see Table 3). There is minimal discussion of the results of the quality assessment in the MS. The trials included in the NMA are assessed using the same quality assessment criteria (MS Appendix 10.7.8, Tables 105 to Table 108). However, there is no discussion of the results of the quality assessment of the trials included in the NMA. (although a sensitivity analysis was conducted in which trials which were not double-blinded were excluded (see SLR report 24[3])). The ERG notes that there are some uncertainties regarding quality of some of the studies in the NMA due to poor reporting in the primary studies (see Table 4) Independent quality assessment by two reviewers does not appear to have been performed on studies included in the manufacturer's systematic review and this is a significant limitation of the systematic review.
4. Is sufficient detail of the individual studies presented?	Yes, however, much of the detail can only be found in the appendices.
5. Are the primary studies summarised appropriately?	Yes, though see Table 4 for the ERG appraisal of the NMA for discussion of limitations in the analysis.

3.3 Summary of submitted evidence

Results are presented in the following sub-sections for the three canagliflozin trials the ERG consider relevant to the decision problem (DIA3006, DIA3009 and DIA3015), at week 26 or 52 as appropriate and for the mITT (NB. DIA3009 measured outcomes up to 104 weeks, however limited data are presented for this time point, and none of the study hypotheses included this time point, therefore we have not reported results here). The results are reported by each outcome measure in turn (outcomes that are included in the manufacturer's economic model have been prioritised here). Data have been reproduced here from the MS as well as from the trial journal publications where necessary. The results of the manufacturer's NMA are also presented, by outcome measure. The main manufacturer's submission document provides NMA forest plots showing mean differences between canagliflozin and selected comparators from each NMA model (MS section 6.7.6). However the following sub-sections report tabulated data showing median differences and rankings for *all treatments* included in each NMA (taken from MS appendix 10.7, and MS reference 72, which only present model median differences, rather than study mean differences, for comparisons of canagliflozin against active comparators) as the ERG considered these to be more informative than the findings reported in the MS for only selected comparators. The tabulated NMA results are ranked with a SUCRA of 100% indicating that the treatment always ranks first, and a SUCRA of 0% indicating that a treatment always ranks last.

3.3.1 Summary of results for HbA_{1c}

<u>Dual therapy</u>

In DIA3009, canagliflozin 300mg in combination with metformin was shown to be statistically superior to glimepiride in combination with metformin in reducing HbA_{1c} from baseline to week 52 (difference of mean change, minus glimepiride: -0.12%, (95% CI -0.217;-0.023)) (Table 6). Canagliflozin 100mg was shown to be non-inferior to glimepiride at week 52 (-0.01%, (95% CI -0.109;0.085)). The ERG notes that the per protocol analyses presented in the journal publication show similar results for canagliflozin 100mg, with it found to be non-inferior to glimepiride at week 52 (difference of LS means: 0.05% (95% CI -0.05;0.14)).[21] However, in the per protocol analysis, canagliflozin 300mg was only found to be non-inferior, and not superior, to glimepiride at week 52 (difference of LS means: -0.06% (95% CI -0.15;0.03)).

In DIA3006, canagliflozin 300mg,100mg and sitagliptin, each in combination with metformin, were statistically superior to placebo and metformin in the difference of mean change in HbA_{1c} at week 26 (-0.77% (95% CI -0.914;-0.636); -0.62% (95% CI -0.758;-0.481); -0.66% (95% CI -0.795;-0.516, respectively)) (Table 6). At week 52 both canagliflozin 100mg and 300mg were non-inferior to sitagliptin in the difference of mean change in HbA_{1c} (0.00% (95% CI -0.12; 0.12) and -0.15% (95% CI -0.27; -0.03)) respectively. Canagliflozin 300mg also demonstrated statistical superiority to sitagliptin in reducing HbA_{1c} (with the upper limit of the 95% CI less than 0.0%). The per-protocol analyses are not reported in the journal publication for DIA3006. The ERG notes that the trial CSR shows that in the per-protocol analyses both canagliflozin 100mg and 300mg were non-inferior to sitagliptin at week 52 (difference of LS means: 0.12% (95% CI 0.005;0.242) and -0.03% (95% CI -0.145;0.088), respectively). In this analysis, however, canagliflozin 300mg was not superior to sitagliptin. Overall, the per-protocol analyses for both dual therapy trials showed less favourable results for canagliflozin than the mITT analyses.

Table 6: RCT results for dual therapy, MET background (HbA_{1c} %)

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 LOCF Mean (SD)	7.00 (0.769)	6.90 (0.806)		7.04 (0.984)
Change from baseline LS Mean (SE)	-0.82 (0.039)	-0.93 (0.039)		-0.81 (0.039)
Diff. of LS Means (SE) (minus glimepiride at week 52) (95% CI)	-0.01 (0.050) (-0.109;0.085)	-0.12 (0.050) (-0.217;-0.023)		
DIA3006	**(n = 368)**	**(n = 367)**	**(n = 181)**	**DPP-4 i (sitagliptin) (n = 366)**
Value at Week 26 LOCF Mean (SD)	7.13 (0.861)	6.98 (0.819)	7.76 (1.216)	7.08 (0.970)
Change from baseline LS Mean (SE)	-0.79 (0.044)	-0.94 (0.044)	-0.17 (0.060)	-0.82 (0.044)
Diff. of LS Means (SE) (minus placebo at week 26) P-Value 95%CI	-0.62 (0.071) <0.001 (-0.758;-0.481)	-0.77 (0.071) <0.001 (-0.914;-0.636)		-0.66 (0.071) N/A[a] (-0.795;-0.516)
Change from baseline LS Mean (± SE change) week 52[b]	-0.73 (0.05)	-0.88 (0.05)		-0.73 (0.05)
Diff. of LS Means (minus	0.00 (-0.12; 0.12)	-0.15		

sitagliptin at week 52)[b] (95% CI)		(-0.27; -0.03)		

mITT population

[a] Statistical comparison between canagliflozin and sitagliptin was not performed at week 26, though in the text of the MS it is stated p<0.001 versus placebo for all active treatment arms (page 79)

[b] Data from trial publication[22]

[redacted]

[redacted]

[redacted]

Table 7: Proportion of patients achieving HbA_{1c} targets, dual therapy MET background 52 weeks

[redacted]				

[redacted]

Table 8 reports the HbA_{1c} % results of the random effects NMA for the metformin background at 26 weeks, with treatments ranked according to the SUCRA. This network comprised 28 trials, of which there were six direct comparisons (i.e. comparisons where at least two studies compared the same interventions, varying from 2 to 6 studies). Statistical heterogeneity (according to the p value of the Cochran test) was present for five of the six comparisons, which the manufacturer suggests could be due to different levels of HbA_{1c} at baseline across studies.[3] (NB. The ERG notes that there appears to be a seventh direct comparison depicted in the NMA illustration in the accompanying SLR report,[3] of exenatide 10µg versus placebo (2 trials), but no details of heterogeneity are given). The percentage that the treatment always ranks first was highest for the GLP-1-a drugs exenatide 2mg and liraglutide (93% and 79% respectively). Canagliflozin 100mg had a probability of lowering HbA_{1c} of 4% and 11% compared to exenatide 2mg and liraglutide respectively. The corresponding probabilities for canagliflozin 300mg were 8% and 22%. The percentage that the treatment always ranks first for canagliflozin 300mg and 100mg was 61% and 44% respectively

Table 8: NMA results for dual therapy 26 weeks, MET background (HbA_{1c} %)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	Crl	Prob	Median	sd	Crl	Prob	
Exenatide 2mg QWK	0.63	0.36	[-0.08 ; 1.35]	4%	0.51	0.36	[-0.20 ; 1.22]	8%	93%
Liraglutide 1.2mg QD	0.32	0.27	[-0.21 ; 0.85]	11%	0.19	0.27	[-0.34 ; 0.72]	22%	79%
Pioglitazone 45mg QD	0.33	0.37	[-0.39 ; 1.06]	17%	0.21	0.37	[-0.52 ; 0.93]	28%	75%
Glimepiride	0.14	0.20	[-0.25 ; 0.54]	22%	0.02	0.20	[-0.38 ; 0.42]	46%	64%
Glibenclamide	0.18	0.41	[-0.62 ; 0.99]	32%	0.06	0.41	[-0.74 ; 0.87]	44%	63%
Pioglitazone 30mg QD	0.13	0.25	[-0.37 ; 0.65]	29%	0.01	0.25	[-0.49 ; 0.52]	49%	61%

Cana 300mg QD	0.13	0.20	[-0.26 ; 0.51]	25%					61%
Sitagliptin 100mg QD	0.03	0.20	[-0.36 ; 0.43]	43%	-0.09	0.20	[-0.49 ; 0.30]	69%	48%
Linagliptin 5mg QD	-0.01	0.34	[-0.69 ; 0.67]	51%	-0.13	0.34	[-0.82 ; 0.55]	66%	45%
Cana 100mg QD					-0.13	0.20	[-0.51 ; 0.26]	75%	44%
Exenatide 10µg BID	-0.02	0.28	[-0.57 ; 0.54]	53%	-0.15	0.28	[-0.69 ; 0.42]	70%	43%
Saxagliptin 5mg QD	-0.04	0.28	[-0.60 ; 0.52]	56%	-0.17	0.28	[-0.73 ; 0.40]	73%	41%
Vildagliptin 100mg QD	-0.05	0.23	[-0.50 ; 0.42]	58%	-0.17	0.23	[-0.62 ; 0.30]	78%	38%
Dapagliflozin 10mg QD	-0.24	0.28	[-0.80 ; 0.32]	81%	-0.36	0.28	[-0.92 ; 0.20]	91%	21%
Exenatide 5µg BID	-0.29	0.35	[-0.97 ; 0.41]	80%	-0.41	0.35	[-1.10 ; 0.28]	89%	21%
Placebo	-0.65	0.20	[-1.04 ; -0.26]	100 %	-0.77	0.20	[-1.17 ; -0.38]	100 %	1%

Table 9 reports the HbA_{1c} % results of the random effects NMA for the metformin background at 52 weeks. This network comprised 17 trials, of which only one was a direct comparison (i.e. comparison where at least two studies compared the same intervention). This was canagliflozin 300mg versus 100mg (two trials), of which there was no statistical heterogeneity present, though the manufacturer notes that the mean duration of diabetes at baseline across the studies was quite heterogeneous.[3] The percentage that the treatment always ranks first was highest for the GLP-1-a liraglutide (89%), followed by the SU glibenclamide (73%) and canagliflozin 300mg (71%). The percentage for canagliflozin 100mg was 51%. Canagliflozin 100mg was associated with the probability of lowering HbA_{1c} of at least 54% compared to the DPP-4-i drugs sitagliptin vildagliptin and saxagliptin, and compared to the SUs gliclazide and glipizide. It was also associated with a probability of lowering HbA_{1c} of 74% compared to the SGLT2 drug dapagliflozin. The corresponding probability for canagliflozin 300mg compared to these drugs was at least 81% (except for the probability of gliclazide).

Table 9: NMA results for dual therapy 52 weeks, MET background (HbA_{1c} %)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	Crl	Prob	Median	sd	Crl	Prob	
Liraglutide 1.2mg QD	0.40	0.36	[-0.33 ; 1.11]	10%	0.27	0.36	[-0.46 ; 0.98]	17%	89%
Glibenclamide	0.19	0.34	[-0.55 ; 0.78]	26%	0.06	0.34	[-0.69 ; 0.65]	41%	73%
Cana 300mg QD	0.13	0.19	[-0.25 ; 0.52]	18%					71%
Pioglitazone 30mg QD	0.11	0.32	[-0.44 ; 0.84]	32%	-0.02	0.32	[-0.57 ; 0.72]	53%	68%
Exenatide 10µg BID	0.02	0.30	[-0.65 ; 0.55]	47%	-0.11	0.30	[-0.78 ; 0.42]	69%	52%
Glimepiride	0.00	0.22	[-0.45 ; 0.46]	50%	-0.13	0.22	[-0.58 ; 0.33]	78%	51%
Cana 100mg QD					-0.13	0.19	[-0.52 ; 0.25]	82%	51%

Sitagliptin 100mg QD	-0.01	0.22	[-0.48 ; 0.44]	54%	-0.14	0.23	[-0.61 ; 0.31]	81%	49%
Gliclazide	-0.03	0.39	[-0.79 ; 0.78]	54%	-0.16	0.39	[-0.92 ; 0.66]	71%	48%
Glipizide	-0.08	0.31	[-0.72 ; 0.53]	65%	-0.20	0.31	[-0.85 ; 0.40]	82%	41%
Vildagliptin 100mg QD	-0.07	0.27	[-0.60 ; 0.51]	64%	-0.20	0.27	[-0.73 ; 0.39]	82%	40%
Saxagliptin 5mg QD	-0.13	0.41	[-0.99 ; 0.67]	69%	-0.26	0.41	[-1.11 ; 0.54]	82%	35%
Dapagliflozin 10mg QD	-0.14	0.32	[-0.81 ; 0.47]	74%	-0.27	0.32	[-0.94 ; 0.34]	87%	31%
Placebo	-0.62	0.26	[-1.17 ; -0.12]	99%	-0.75	0.26	[-1.31 ; -0.25]	99%	2%

The ERG has not reported NMA results for the proportion of patients reaching HbA_{1c} <7%
For these results please refer to MS Appendix 10.7.14.

Triple therapy

The DIA3015 trial found that canagliflozin 300mg in combination with metformin and sulfonylurea was non-inferior to sitagliptin in combination with metformin and sulfonylurea in the difference of mean change in HbA_{1c} at week 52 (-0.37% (95 CI -.05; -0.25) (

Table 10). Canagliflozin 300mg also demonstrated statistical superiority (upper limit of the 95% CI less than 0.0%). The ERG notes that per protocol analysis results reported in the trial paper were consistent with the mITT analyses (difference of LS means: -0.21% (95% CIs -0.34 to -0.08)).

Table 10: RCT results for triple therapy, MET and SU background (HbA_{1c} %)

	Canagliflozin 300mg	**Active comparator**
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 LOCF - Mean (SD), %	7.12 (0.875)	7.50 (1.083)
Change from baseline - LS Mean (SE)	-1.03 (0.048)	-0.66 (0.049)
Diff. of LS Means (SE) (minus sitagliptin) 95% CI	-0.37 (0.064) (-0.50; -0.25)	

mITT analysis

[redacted]

[redacted]

[redacted]

Table 12 reports the HbA_{1c} % results of the random effects NMA for the metformin and sulfonylurea background at 26 weeks. This network comprised nine trials, including one direct comparison (i.e. comparison where at least two studies compared the same intervention). This was long acting insulin versus biphasic insulin (two trials), of which there was no statistical heterogeneity present, though the manufacturer notes that both of these trials had a lower mean BMI at baseline than other trials in the network.[3] The percentage that the treatment always ranks first was highest for biphasic insulin (90%) followed by long-acting insulin (70%). The percentage that the treatment always ranks first for canagliflozin 300mg and 100mg was 65% and 45% respectively. Canagliflozin 100mg had a probability of lowering HbA_{1c} of 12% and 24% compared to biphasic insulin and long-acting insulin, respectively. The corresponding probabilities for canagliflozin 300mg were 18% and 40%.

Table 12: NMA results for triple therapy 26 weeks, MET and SU background (HbA_{1c} %)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Biphasic insulin	0.70	0.99	[-1.13 ; 2.62]	12%	0.46	0.92	[-1.22 ; 2.25]	18%	90%
Long-acting insulin	0.35	0.94	[-1.43 ; 2.16]	24%	0.11	0.87	[-1.52 ; 1.77]	40%	70%
Cana 300mg QD	0.24	0.70	[-1.09 ; 1.58]	26%					65%
Sitagliptin 100mg QD	0.07	0.82	[-1.48 ; 1.64]	43%	-0.17	0.60	[-1.30 ; 0.97]	72%	51%
Exenatide 10µg BID	0.03	0.94	[-1.78 ; 1.79]	47%	-0.21	0.87	[-1.87 ; 1.42]	69%	47%
Cana 100mg QD					-0.24	0.70	[-1.58 ; 1.09]	74%	45%
Exenatide 5µg BID	-0.08	1.00	[-1.98 ; 1.82]	56%	-0.32	0.93	[-2.10 ; 1.44]	75%	40%
Linagliptin 5mg QD	-0.13	1.00	[-2.04 ; 1.79]	61%	-0.37	0.95	[-2.15 ; 1.43]	78%	37%
Placebo	-0.75	0.70	[-2.08 ; 0.59]	93%	-1.00	0.60	[-2.12 ; 0.16]	97%	5%

Table 13 reports the HbA_{1c} % results of the fixed effect NMA for the metformin and thiazolidinediones (pioglitazone) background at 26 weeks. This was the smallest network with only two RCTs included and no direct comparisons, so caution is advised in the interpretation of the results. The percentage that the treatment always ranks first was highest for canagliflozin 300mg (87%) followed by sitagliptin (67%) and canagliflozin 100mg (46%).

Table 13: NMA results for triple therapy 26 weeks, MET and TZD background (HbA_{1c} %)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Canagliflozin 300mg	0.14	0.10	[-0.05 ; 0.33]	8%					87%
Sitagliptin 100mg	0.07	0.13	[-0.19 ; 0.33]	29%	-0.07	0.13	[-0.33 ; 0.19]	69%	67%
Canagliflozin 100mg					-0.14	0.10	[-0.33 ; 0.05]	92%	46%
Placebo	-0.63	0.10	[-0.82 ; -0.44]	100%	-0.77	0.10	[-0.96 ; -0.58]	100%	0%

Insulin background

As stated earlier (see section 3.1.3), there were no RCTs of canagliflozin included in the MS with insulin as a background therapy that the ERG considered were relevant to the scope. This sub-section therefore just comprises the NMA results (fixed effect) for insulin background therapy at 26 weeks (Table 14). This network comprised 14 trials, of which two were direct comparisons. This was vildagliptin versus placebo (two trials), and pioglitazone versus metformin (2 trials). Statistically significant heterogeneity was found for vildagliptin versus placebo. The manufacturer notes that baseline HbA_{1c} and duration of diabetes was heterogeneous across trials.[3] The manufacturer also states that of the 14 trials, 9 trials included patients taking insulin with oral antidiabetic drugs, and 5 included patients taking insulin monotherapy. [3] Further, the manufacturer points out that the one canagliflozin trial included in this network (DIA3008 insulin sub-study) included patients with a history of, or who were at increased risk of, cardiovascular disease (CVD). A sensitivity analysis was not possible for this background and the manufacturer notes this as a limitation of the analysis. Caution is advised given the clinical and statistical heterogeneity present in this network.

Table 14: NMA results 26 weeks, for insulin background (HbA_{1c} %)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Pioglitazone 45mg QD	0.34	0.18	[-0.01; 0.69]	3%	0.20	0.18	[-0.15; 0.54]	14%	91%
Cana 300mg QD	0.14	0.04	[0.06; 0.22]	0%					78%
Glibenclamide	0.53	0.75	[-0.92; 2.00]	24%	0.39	0.75	[-1.06; 1.86]	30%	76%
Exenatide 10µg BID	0.09	0.13	[-0.17; 0.35]	25%	-0.05	0.13	[-0.31; 0.21]	64%	65%
Pioglitazone 30mg QD	0.04	0.14	[-0.23; 0.30]	40%	-0.10	0.14	[-0.37; 0.17]	78%	56%
Cana 100mg QD					-0.14	0.04	[-0.22; -0.06]	100%	49%
Sitagliptin 100mg QD	-0.01	0.08	[-0.17; 0.16]	55%	-0.15	0.08	[-0.31; 0.02]	96%	47%
Dapagliflozin 10mg QD	-0.01	0.08	[-0.17 ; 0.15]	55%	-0.15	0.08	[-0.31; 0.01]	96%	47%
Metformin	-0.02	0.18	[-0.38; 0.33]	55%	-0.16	0.18	[-0.52; 0.19]	82%	45%
Vildagliptin 100mg QD	-0.11	0.11	[-0.32; 0.10]	84%	-0.25	0.11	[-0.46; -0.04]	99%	29%
Saxagliptin 5mg QD	-0.20	0.10	[-0.40; 0.00]	98%	-0.34	0.10	[-0.54; -0.14]	100%	17%
Placebo	-0.61	0.04	[-0.69; -0.53]	100%	-0.75	0.04	[-0.83; -0.67]	100%	1%

The percentage that the treatment always ranks first was highest for pioglitazone 45mg (91%) followed by canagliflozin 300mg (78%). The percentage that the treatment always ranks first for canagliflozin 100mg was 49%. Canagliflozin 100mg had a probability of lowering HbA_{1c} of 55% and higher compared to all three DPP-4-i drugs included in the network (sitagliptin, vildagliptin, saxagliptin), dapagliflozin, metformin and placebo. For canagliflozin 300mg this probability was 96% (except for metformin).

3.3.2 Summary of results for blood pressure

Dual therapy

In DIA3009, at week 52, both doses of canagliflozin (100mg and 300mg) in combination with metformin were associated with statistically significant greater decreases in patients' systolic and diastolic blood pressure from baseline levels than the SU glimepiride combined with metformin (Table 15 and Table 16). The DIA3006 trial similarly found that at week 52, both doses of canagliflozin in combination with metformin resulted in statistically significant greater reductions in patients' systolic and diastolic blood pressure from baseline levels than the DPP-4-i sitagliptin 100mg combined with metformin.

Table 15: RCT results for dual therapy MET background (SBP [mmHg])

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 LOCF Mean (SD)	126.1 (12.6)	124.8 (12.8)		129.4 (13.9)
Change from baseline LS Mean (SE)	-3.27 (0.6)	-4.56 (0.6)		0.20 (0.6)
Diff. of LS Means (SE) (minus glimepiride) (95% CI)	-3.48(0.7) (-4.9;-2.1)	-4.76 (0.7) (-6.2;-3.4)		
DIA3006	**(n = 368)**	**(n = 367)**	**(n = 183)**	**DPP-4 i (sitagliptin) (n = 366)**
Value at Week 26 LOCF Mean (SD)	124.5 (13.908)	123.7 (12.691)	129.9 (13.453)	126.5 (12.714)
Change from baseline LS Mean (SE)	-3.84 (0.602)	-5.06 (0.605)	1.52 (0.829)	-1.83 (0.611)
Diff. of LS Means (SE) (minus placebo) P-Value (95% CI)	-5.36 (0.979) <0.001 (-7.280;-3.439)	-6.58 (0.981) <0.001 (-8.504;-4.653)		-3.34 (0.984) N/A (-5.273;-1.413)
Change from baseline LS Mean (SE change) week 52[b]	−3.5 (0.6)	−4.7 (0.6)		−0.7 (0.6)
Diff. of LS Means (minus sitagliptin at week 52)[b] P-Value (95% CI)	−2.9 <0.001 (−4.5, –1.3)	−4.0 <0.001 (−5.6, –2.4)		

[b] Data from journal publication[22]

Table 16: RCT results for dual therapy MET background (DBP [mmHg])

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 LOCF Mean (SD)	76.85 (8.5)	76.35 (8.1)		78.66 (8.9)
Change from baseline LS Mean (SE)	-1.84 (0.4)	-2.54 (0.4)		-0.14 (0.4)
Diff. of LS Means (SE) (minus glimepiride) (95% CI)	-1.70 (0.5) (-2.7;-0.8)	-2.39 (0.5) (-3.4;-1.5)		
DIA3006	**(n = 368)**	**(n = 367)**	**(n = 183)**	**DPP-4 i (sitagliptin) (n = 366)**

Value at Week 26 LOCF Mean (SD)	75.72 (9.261)	75.93 (7.864)	78.21 (8.038)	76.68 (8.062)
Change from baseline LS Mean (SE)	-2.19 (0.391)	-2.09 (0.394)	0.28 (0.539)	-1.11 (0.398)
Diff. of LS Means (SE) (minus placebo) (95% CI)	-2.47 (0.637) (-3.716;-1.217)	-2.37 (0.638) (-3.618;-1.113)		-1.39 (0.640) (-2.641;-0.131)
Change from baseline LS Mean (SE change) week 52[b]	−1.8 (0.4)	−1.8 (0.4)		−0.3 (0.4)
Diff. of LS Means (minus sitagliptin at week 52)[b] (95% CI)	−1.4 (−2.4, –0.5)	−1.5 (−2.5, –0.5)		

mITT analysis
[b] Data from trial publication[22]

Table 17 reports the SBP results of the fixed effects NMA for the metformin background at 26 weeks. This network comprised 11 trials, of which there were four direct comparisons. These were canagliflozin 300mg versus 100mg (two trials), dapagliflozin versus placebo (two trials), pioglitazone versus glimepiride (three trials), sitagliptin versus placebo (two trials). There was no statistically significant heterogeneity for these direct comparisons.[3] The GLP-1 analogue exenatide had the highest probability of ranking first as the best treatment, with a percentage of 92%. This was followed by canagliflozin 300mg (84%) and canagliflozin 100mg (67%). Canagliflozin 300mg and 100mg had probabilities of lowering SBP of 12% and 26%, respectively, compared to exenatide. Canagliflozin 300mg had a probability of lowering SBP of 94% compared to canagliflozin 100mg. NMA results for DBP were not reported in either the MS or the accompanying SLR report.[3]

Table 17: NMA results for dual therapy, MET background at 26 weeks (SBP)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Exenatide 2mg QWK	1.83	1.54	[-1.19 ; 4.83]	12%	0.99	1.54	[-2.03 ; 4.01]	26%	92%
Cana 300mg QD	0.84	0.55	[-0.23 ; 1.92]	6%					84%
Cana 100mg QD					-0.84	0.55	[-1.92 ; 0.23]	94%	67%
Pioglitazone 45mg QD	-0.17	1.53	[-3.18 ; 2.82]	55%	-1.02	1.53	[-4.03 ; 1.99]	75%	64%
Dapagliflozin 10mg QD	-0.73	1.40	[-3.46 ; 2.02]	70%	-1.58	1.40	[-4.32 ; 1.18]	87%	56%
Pioglitazone 30mg QD	-0.78	2.48	[-5.68 ; 4.08]	62%	-1.61	2.48	[-6.51 ; 3.22]	74%	55%
Sitagliptin 100mg QD	-1.98	0.72	[-3.39 ; -0.57]	100%	-2.83	0.72	[-4.24 ; -1.41]	100%	35%
Liraglutide 1.2mg QD	-2.10	0.99	[-4.05 ; -0.15]	98%	-2.94	0.99	[-4.88 ; -1.00]	100%	34%
Glimepiride	-4.28	0.63	[-5.50 ; -3.04]	100%	-5.11	0.62	[-6.35 ; -3.89]	100%	7%

Placebo	-4.43	0.79	[-5.99 ; -2.89]	100%	-5.27	0.79	[-6.84 ; -3.73]	100%	6%

Table 18 reports the SBP results of the fixed effects NMA for the metformin background at 52 weeks. This network comprised four trials, with one direct comparison (2 trials comparing canagliflozin 300mg with 100mg, with no statistically significant heterogeneity). Canagliflozin 300mg had the highest probability of ranking as the best treatment, with a percentage of 87%, followed by pioglitazone (86%) and canagliflozin 100mg (65%). Canagliflozin 300mg had a probability of lowering SBP of at least 98% compared to all active comparators included in the analysis, except for pioglitazone. NMA results for DBP were not reported in either the MS or the accompanying SLR report.[3]

Table 18: NMA results for dual therapy, MET background at 52 weeks (SBP)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Cana 300mg QD	1.22	0.59	[0.05 ; 2.36]	2%					87%
Pioglitazone 30mg QD	2.05	2.69	[-3.22 ; 7.37]	22%	0.83	2.69	[-4.44 ; 6.15]	38%	86%
Cana 100mg QD					-1.22	0.59	[-2.36 ; -0.05]	98%	65%
Sitagliptin 100mg QD	-2.84	0.82	[-4.44 ; -1.22]	100%	-4.04	0.82	[-5.64 ; -2.44]	100%	30%
Liraglutide 1.2mg QD	-3.50	1.55	[-6.55 ; -0.43]	99%	-4.71	1.55	[-7.73 ; -1.66]	100%	17%
Glimepiride	-3.52	0.76	[-5.02 ; -2.05]	100%	-4.73	0.76	[-6.22 ; -3.26]	100%	15%

Triple therapy

In the triple therapy trial DIA3015, canagliflozin (300mg) combined with metformin and a sulfonylurea resulted in the statistically significant greater improvement in both SBP and DBP at 52 weeks than the DPP-4-i sitagliptin (100mg) combined with metformin and a sulfonylurea (Table 19 and Table 20).

Table 19: RCT results for triple therapy, MET and SU background (SBP [mmHg])

	Canagliflozin 300mg	**Active comparator**
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 LOCF - Mean (SD)	125.6 (13.641)	130.8 (15.103)
Change from baseline - LS Mean (SE)	-5.06 (0.656)	0.85 (0.666)
Diff. of LS Means (SE) (minus sitagliptin) P-Value 95% CI	-5.91 (0.883) <0.001 (-7.642;-4.175)	

mITT analysis

Table 20: RCT results for triple therapy, MET and SU background (DBP [mmHg])

	Canagliflozin 300mg	Active comparator
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 LOCF - Mean (SD)	75.92 (8.817)	78.32 (9.116)
Change from baseline - LS Mean (SE)	-3.04 (0.407)	-0.30 (0.413)
Diff. of LS Means (SE) (minus sitagliptin) P-Value 95% CI	-2.73 (0.547) NR (-3.809; -1.659)	

mITT analysis

Table 21 reports the SBP results of the fixed effects NMA for the metformin and sulfonylurea background. Three trials were included in this network, with no direct comparisons. Canagliflozin 100mg, followed by canagliflozin 300mg, had the highest probabilities of ranking first, with percentages of 89% and 78%, respectively. Both doses of canagliflozin had a probability of lowering SBP of at least 88% compared to placebo, long-acting insulin and the DPP-4-i sitagliptin. Canagliflozin 300mg had a probability of lowering SBP of 33% compared to canagliflozin 100mg. NMA results for DBP were not reported in either the MS or the accompanying SLR report.[3]

Table 21: NMA results for triple therapy 26 weeks, MET and SU background (SBP)

	Cana 100mg				Cana 300 mg				SUCRA
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Cana 100mg QD					0.61	1.38	[-2.07 ; 3.34]	33%	89%
Cana 300mg QD	-0.61	1.38	[-3.34 ; 2.07]	67%					78%
Placebo	-2.26	1.39	[-5.01 ; 0.47]	95%	-1.66	1.39	[-4.35 ; 1.10]	88%	50%
Long-acting insulin	-4.17	2.47	[-9.06 ; 0.69]	95%	-3.56	2.47	[-8.43 ; 1.25]	93%	26%
Sitagliptin 100mg QD	-5.76	1.66	[-9.02 ; -2.53]	100%	-5.16	0.91	[-6.94 ; -3.38]	100%	7%

An NMA was not conducted to analyse change in SBP for patients treated with triple therapy with canagliflozin on a background of MET and TZD, because only one trial provided evidence for this network.

<u>Insulin background</u>

Table 22 reports the SBP of the fixed effects NMA for the insulin background at 26 weeks. Four trials were included in this network, with no direct comparisons. As noted above, the one canagliflozin trial included in this network (DIA3008 insulin sub-study) included patients with a history of, or who were at increased risk of, CVD, and therefore is likely to be heterogeneous. Canagliflozin 300mg ranked first of the included treatments with a percentage of 94%, followed by exenatide (70%) and canagliflozin 100mg (63%). Canagliflozin 300mg had a probability of lowering SBP of at least 96% compared to placebo, canagliflozin 100mg and all the other active comparators included in the analysis, except for the GLP-1 analogue exenatide. NMA results for diastolic blood pressure were not reported in either the MS or the accompanying SLR report.[3]

Table 22: NMA results for insulin background at 26 weeks (SBP)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Cana 300mg QD	1.71	0.72	[0.31 ; 3.12]	1%					94%
Exenatide 10µg BID	0.38	1.97	[-3.46 ; 4.23]	42%	-1.33	1.97	[-5.19 ; 2.52]	75%	70%
Cana 100mg QD					-1.71	0.72	[-3.12 ; -0.31]	99%	63%
Dapagliflozin 10mg QD	-0.92	1.54	[-3.93 ; 2.08]	72%	-2.62	1.54	[-5.64 ; 0.40]	96%	51%
Saxagliptin 5mg QD	-3.99	1.54	[-7.01 ; -0.96]	99%	-5.70	1.54	[-8.73 ; -2.67]	100%	12%
Placebo	-4.02	0.71	[-5.42 ; -2.63]	100%	-5.73	0.72	[-7.13 ; -4.34]	100%	10%

3.3.3 Summary of results for body weight

<u>Dual therapy</u>

Table 23 shows that in both dual therapy trials (DIA3009 and DIA3006), canagliflozin 300mg and 100mg were associated with statistically significant greater weight loss in kg at 52 weeks than the SU glimepiride or the DPP-4-i sitagliptin. Results for change in BMI from baseline were only reported in the NMAs (see below).

Table 23 - RCT results for dual therapy MET background (weight [kg])

Trial no. (acronym)	**Canagliflozin 100mg**	**Canagliflozin 300mg**	**Placebo**	**Active comparator**
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 LOCF Mean (SD)	83.37 (19.756)	82.78 (19.092)		87.54 (19.914)
Change from baseline LS Mean (SE)	-4.2 (0.2)	-4.7 (0.2)		1.0 (0.2)

Diff. of LS Means (SE) (minus glimepiride) P-value (95% CI)	-5.2 (0.3) <0.001 (-5.7;-4.7)	-5.7 (0.3) <0.001 (-6.2;-5.1)		
DIA3006	**(n = 368)**	**(n = 367)**	**(n = 181)**	**DPP-4 i (sitagliptin) (n = 366)**
Value at Week 26 LOCF Mean (SD)	85.61 (21.700)	82.15 (20.168)	85.85 (22.411)	86.70 (20.669)
Change from baseline LS Mean (SE)	-3.7 (0.2)	-4.2 (0.2)	-1.2 (0.3)	-1.2 (0.2)
Diff. of LS Means (SE) (minus placebo) P-Value (95% CI)	-2.5 (0.3) <0.001 (-3.1;-1.9)	-2.9 (0.3) <0.001 (-3.5;-2.3)		-0.0 (0.3) (-0.6;0.6)
	(n = 368)	**(n = 367)**		**(n = 366)**
Change from baseline LS Mean (SE change) week 52[b]	−3.3 (0.2)	−3.7 (0.2)		−1.2 (0.2)
Diff. of LS Means (minus sitagliptin at week 52)[b] P-Value (95% CI)	−2.4 <0.001 (−3.0, –1.8)	−2.9 <0.001 (−3.4, –2.3)		

mITT
[b] Data from trial publication[22]

Table 24 reports the body weight results of the fixed effects NMA for the metformin background at 26 weeks. This network comprised 21 trials, of which there were six direct comparisons: canagliflozin 300mg versus 100mg (two trials); dapagliflozin versus placebo (two trials); pioglitazone versus glimepiride (two trials); sitagliptin versus placebo (four trials); vildagliptin versus placebo (two trials); exenatide versus placebo (two trials). Statistically significant heterogeneity was found for the vildagliptin versus placebo comparison, where the two trials differed in baseline BMI.[3] Canagliflozin 300mg had a probability of being the best treatment of 94% and ranked first of all the treatments included. The GLP-1 analogue exenatide ranked second with a probability of 92% and dapagliflozin ranked third with a probability of 84%. Canagliflozin 300mg had a probability of reducing weight of at least 98% compared to placebo and all the active comparators, except compared to exenatide and the SGLT-2 dapagliflozin. It had a probability of 100% of reducing weight compared to canagliflozin 100mg.

Table 24: NMA results for dual therapy 26 weeks, MET background (weight)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	Crl	Prob	Median	sd	Crl	Prob	
Cana 300mg QD	0.52	0.15	[0.22 ; 0.82]	0%					94%
Exenatide 10µg BID	0.59	0.58	[-0.55 ; 1.71]	15%	0.07	0.58	[-1.06 ; 1.20]	45%	92%
Dapagliflozin 10mg QD	0.21	0.33	[-0.44 ; 0.84]	26%	-0.31	0.33	[-0.96 ; 0.33]	83%	84%
Cana 100mg QD					-0.52	0.15	[-0.82 ; -0.22]	100%	78%
Liraglutide 1.2mg QD	-0.41	0.26	[-0.93 ; 0.10]	94%	-0.93	0.26	[-1.44 ; -0.42]	100%	64%
Exenatide 5µg BID	-0.55	0.54	[-1.60 ; 0.50]	85%	-1.07	0.54	[-2.12 ; -0.01]	98%	63%
Exenatide 2mg QWK	-0.57	0.49	[-1.53 ; 0.39]	88%	-1.09	0.49	[-2.04 ; -0.13]	99%	62%
Placebo	-1.83	0.20	[-2.23 ; -1.44]	100%	-2.35	0.20	[-2.74 ; -1.95]	100%	45%
Sitagliptin 100mg QD	-2.07	0.18	[-2.42 ; -1.73]	100%	-2.59	0.18	[-2.94 ; -2.24]	100%	38%
Vildagliptin 100mg QD	-2.64	0.42	[-3.48 ; -1.81]	100%	-3.16	0.42	[-3.99 ; -2.33]	100%	31%
Glimepiride	-4.02	0.16	[-4.34 ; -3.70]	100%	-4.54	0.16	[-4.86 ; -4.23]	100%	18%
Pioglitazone 30mg QD	-4.24	0.50	[-5.22 ; -3.26]	100%	-4.75	0.50	[-5.74 ; -3.77]	100%	15%
Glibenclamide	-4.62	1.62	[-7.81 ; -1.45]	100%	-5.14	1.62	[-8.32 ; -1.97]	100%	13%
Pioglitazone 45mg QD	-5.67	0.49	[-6.62 ; -4.72]	100%	-6.19	0.49	[-7.14 ; -5.23]	100%	2%

Table 25 shows the fixed effects NMA results for weight loss measured in terms of change in BMI at 26 weeks on a metformin background. This network comprised nine trials with three direct comparisons: canagliflozin 300mg versus canagliflozin 100mg (two trials); sitagliptin versus placebo (two trials); pioglitazone versus placebo (two trials). There was no statistically significant heterogeneity for the direct comparisons. Canagliflozin 300mg and canagliflozin 100mg were the highest ranking treatments, with probabilities of being the best treatment of 98% and 82%, respectively.

Table 25: NMA results for dual therapy at 26 weeks, MET background (BMI)

	Cana 100mg				Cana 300 mg				SUCRA
	Median	sd	Crl	Prob	Median	sd	Crl	Prob	
Cana 300mg QD	0.20	0.06	[0.09 ; 0.31]	0%					98%
Cana 100mg QD					-0.20	0.06	[-0.31 ; -0.09]	100%	82%
Exenatide 10µg BID	-0.04	0.24	[-0.50 ; 0.42]	57%	-0.24	0.24	[-0.70 ; 0.22]	85%	81%
Vildagliptin 100mg QD	-0.44	0.22	[-0.88 ; 0.00]	97%	-0.64	0.22	[-1.08 ; -0.20]	100%	62%
Placebo	-0.74	0.10	[-0.92 ; -0.55]	100%	-0.93	0.10	[-1.13 ; -0.75]	100%	45%
Sitagliptin 100mg QD	-0.76	0.08	[-0.92 ; -0.60]	100%	-0.96	0.08	[-1.12 ; -0.80]	100%	43%
Glimepiride	-1.41	0.07	[-1.54 ; -1.27]	100%	-1.61	0.07	[-1.74 ; -1.47]	100%	18%
Pioglitazone 15mg BID	-1.49	0.44	[-2.35 ; -0.63]	100%	-1.69	0.44	[-2.55 ; -0.83]	100%	16%

Glibenclamide	-1.84	0.35	[-2.53 ; -1.15]	100%	-2.04	0.35	[-2.73 ; -1.36]	100%	5%

Canagliflozin 300mg was associated with a probability of improving BMI of 100% compared to placebo and all active comparators, including canagliflozin 100mg, except the GLP-1 analogue exenatide.

Table 26 reports the weight change results of the fixed effects NMA for the metformin background at 52 weeks. Eleven trials were included in this network, with one direct comparison (canagliflozin 300mg versus canagliflozin 100mg, two trials, no statistically significant heterogeneity). The GLP-4 analogue exenatide had the highest probability of being the best treatment, with a percentage of 94%. This was followed by canagliflozin 300mg (90%) and canagliflozin 100mg (79%). Canagliflozin 100mg and 300mg had probabilities of reducing weight of 16% and 23%, respectively, compared to exenatide, but both doses had probabilities of reducing weight of 100% compared to the DPP-4-inhibitors sitagliptin and vildagliptin, placebo, the SUs glimepiride, glibenclamide and glipizide, and pioglitazone.

Table 26: NMA results for dual therapy, MET background at 52 weeks (weight)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Exenatide 10µg BID	1.47	1.50	[-1.48 ; 4.41]	16%	1.11	1.50	[-1.84 ; 4.05]	23%	94%
Cana 300mg QD	0.37	0.28	[-0.17 ; 0.91]	9%					90%
Cana 100mg QD					-0.37	0.28	[-0.91 ; 0.17]	91%	79%
Dapagliflozin 10mg QD	-0.11	0.51	[-1.10 ; 0.89]	59%	-0.48	0.51	[-1.47 ; 0.53]	83%	77%
Liraglutide 1.2mg QD	-0.49	0.45	[-1.37 ; 0.38]	86%	-0.85	0.45	[-1.73 ; 0.02]	97%	68%
Sitagliptin 100mg QD	-2.12	0.28	[-2.66 ; -1.57]	100%	-2.48	0.28	[-3.03 ; -1.93]	100%	48%
Vildagliptin 100mg QD	-2.16	0.77	[-3.74 ; -0.69]	100%	-2.52	0.77	[-4.09 ; -1.05]	100%	48%
Placebo	-2.58	0.61	[-3.77 ; -1.35]	100%	-2.95	0.61	[-4.13 ; -1.71]	100%	41%
Glimepiride	-3.97	0.77	[-5.54 ; -2.48]	100%	-4.33	0.77	[-5.89 ; -2.85]	100%	25%
Pioglitazone 30mg QD	-4.57	0.85	[-6.28 ; -2.93]	100%	-4.93	0.85	[-6.64 ; -3.29]	100%	15%
Glipizide	-4.76	0.46	[-5.65 ; -3.85]	100%	-5.12	0.46	[-6.02 ; -4.21]	100%	14%
Glibenclamide	-10.83	2.17	[-15.06 ; -6.58]	100%	-11.19	2.17	[-15.43 ; -6.95]	100%	0%

Table 27 shows that the random effects NMA results for weight loss measured in terms of change in BMI at 52 weeks on a metformin background. This network comprised nine trials, with one direct comparison (canagliflozin 300mg versus canagliflozin 100mg, no statistically

significant heterogeneity). A similar pattern of results were found to the NMA of change in weight at 52 weeks. The GLP-1 analogue exenatide ranked first with a probability of 91% of being the best treatment, followed by canagliflozin 300mg with a probability of 72%. Canagliflozin 300mg had a probability of reducing weight of 16% compared to exenatide, but had a probability of at least 85% compared to the DPP-4-i sitagliptin, pioglitazone, and the SUs glimepiride and glibenclamide.

Table 27: NMA results for dual therapy, MET background at 52 weeks (BMI)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Exenatide 10µg BID	1.24	1.47	[-2.16 ; 3.95]	14%	1.10	1.47	[-2.30 ; 3.82]	16%	91%
Cana 300mg QD	0.14	0.96	[-1.86 ; 2.14]	38%					72%
Cana 100mg QD					-0.14	0.96	[-2.14 ; 1.86]	62%	67%
Vildagliptin 100mg QD	-0.19	2.00	[-4.63 ; 3.78]	57%	-0.32	2.00	[-4.75 ; 3.61]	62%	60%
Placebo	-0.49	1.47	[-3.82 ; 2.36]	73%	-0.63	1.47	[-3.94 ; 2.22]	77%	50%
Sitagliptin 100mg QD	-0.63	1.17	[-3.17 ; 1.75]	82%	-0.77	1.17	[-3.28 ; 1.60]	85%	45%
Pioglitazone 30mg QD	-1.53	2.28	[-6.01 ; 3.03]	78%	-1.66	2.28	[-6.13 ; 2.90]	80%	33%
Glimepiride	-1.72	1.17	[-4.08 ; 0.81]	95%	-1.85	1.17	[-4.22 ; 0.69]	95%	23%
Glibenclamide	-2.54	1.53	[-5.87 ; 0.52]	96%	-2.67	1.53	[-6.00 ; 0.40]	97%	9%

<u>Triple therapy</u>

In the triple therapy trial DIA3015, at 52 weeks, canagliflozin (300mg) combined with metformin and a sulfonylurea resulted in a statistically significant greater improvement in BMI than the DPP-4-i sitagliptin combined with metformin and a sulfonylurea (Table 28). Results for change in BMI from baseline were only reported in the NMAs (see Table 30 below).

Table 28: RCT results for triple therapy, MET and SU background at 52 weeks (weight [kg])

	Canagliflozin 300mg	**Active comparator**
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 LOCF - Mean (SD)	85.27 (22.537)	89.65 (23.023)
Change from baseline - LS Mean (SE)	-2.5 (0.2)	0.3 (0.2)
Diff. of LS Means (SE) (minus sitagliptin) P-Value 95% CI	-2.8 (0.3) <0.001 (-3.3;-2.2)	

mITT analysis

Superseded – see erratum

Table 29 reports the weight change results of the random effects NMA for the metformin and sulfonylurea background. The network comprised eight trials, with no direct comparisons. Canagliflozin 300mg ranked first, with a probability of being the best treatment of 78%. It was associated with a probability of reducing weight of at least 80% compared to the DPP-4-inhibitors linagliptin and sitagliptin, long-acting insulin and biphasic insulin. It was associated with a probability of 53% and 60% of reducing weight compared to the GLP-1 analogues exenatide 10µg BID (twice daily) and exenatide 5µg BID, respectively.

Table 29 – NMA results for triple therapy 26 weeks, MET and SU background (weight)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Cana 300mg QD	0.62	2.31	[-4.34 ; 5.57]	35%					78%
Exenatide 10µg BID	0.47	3.11	[-6.09 ; 7.24]	41%	-0.14	2.86	[-6.15 ; 6.08]	53%	76%
Exenatide 5µg BID	0.08	3.26	[-6.83 ; 7.11]	49%	-0.53	3.03	[-6.97 ; 5.94]	60%	69%
Cana 100mg QD					-0.62	2.31	[-5.57 ; 4.34]	65%	67%
Placebo	-1.02	2.31	[-5.95 ; 3.92]	73%	-1.63	1.96	[-5.83 ; 2.55]	85%	51%
Linagliptin 5mg QD	-1.35	3.32	[-8.47 ; 5.76]	72%	-1.96	3.09	[-8.59 ; 4.65]	80%	47%
Sitagliptin 100mg QD	-2.03	2.70	[-7.78 ; 3.76]	83%	-2.64	1.96	[-6.83 ; 1.56]	93%	35%
Long-acting insulin	-3.97	3.12	[-10.62 ; 2.70]	92%	-4.59	2.88	[-10.70 ; 1.57]	95%	16%
Biphasic insulin	-4.52	3.35	[-11.62 ; 2.63]	93%	-5.14	3.12	[-11.76 ; 1.52]	95%	11%

Table 30 shows the fixed effects NMA results for weight loss measured in terms of change in BMI on a metformin and sulfonylurea background. This network comprised just two studies, with no direct comparisons. Again, canagliflozin was the highest ranking treatment, with a probability of being the best treatment of 98%, followed by canagliflozin 100mg with a probability of 69%. Both canagliflozin 100mg and 300mg had probabilities of 100% of reducing weight compared to the DPP-4-i sitagliptin. Canagliflozin 300mg had a probability of 94% of reducing weight compared to canagliflozin 100mg.

Table 30: NMA results for triple therapy 26 weeks, MET and SU background (BMI)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Cana 300mg QD	0.22	0.14	[-0.06 ; 0.50]	6%					98%
Cana 100mg QD					-0.22	0.14	[-0.50 ; 0.06]	94%	69%
Placebo	-0.40	0.14	[-0.68 ; -0.12]	100%	-0.62	0.14	[-0.90 ; -0.34]	100%	32%

Sitagliptin 100mg QD	-0.70	0.16	[-1.02 ; -0.38]	100%	-0.92	0.08	[-1.08 ; -0.76]	100%	1%

Table 31 reports the weight change results of the fixed effects NMA for the metformin and thiazolidinedione (pioglitazone) background at 26 weeks. This network comprised just two studies, with no direct comparisons. Canagliflozin 300mg had the highest probability of being the best treatment, with a percentage of 94%. Canagliflozin 300mg was associated with a 99% probability of reducing weight compared to canagliflozin 100mg and a probability of 84% compared to the DPP-4-i sitagliptin.

Table 31: NMA results for triple therapy at 26 weeks, MET and TZD background (weight)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	Crl	Prob	Median	sd	Crl	Prob	
Canagliflozin 300mg	0.95	0.39	[0.18 ; 1.72]	1%					94%
Canagliflozin 100mg					-0.95	0.39	[-1.72 ; -0.18]	99%	59%
Sitagliptin 100mg	-2.70	3.75	[-10.10 ; 4.62]	76%	-3.65	3.75	[-11.07 ; 3.67]	84%	29%
Placebo	-2.54	0.40	[-3.31 ; -1.77]	100%	-3.49	0.40	[-4.27 ; -2.72]	100%	17%

An NMA was not conducted to analyse change in BMI for patients treated with triple therapy with canagliflozin on a background of thiazolidinedione (pioglitazone), because only one trial provided evidence for this network.

Insulin background

Table 32 reports the weight change results of the fixed effects NMA for the insulin background at 26 weeks. Ten trials were included in this network with no direct comparisons. Canagliflozin 300mg ranked joint first with the GLP-1 analogue exenatide, both with a probability of being the best treatment of 86%. Canagliflozin 300mg was associated with a probability of at least 81% of reducing weight compared to canagliflozin 100mg, the DPP-4-inhibitors sitagliptin, vildagliptin and saxagliptin, pioglitazone, dapagliflozin and the SU glibenclamide. It was associated with probabilities of 45% and 56% of reducing weight compared to the GLP-1 analogue exenatide and metformin, respectively.

An NMA was not conducted to analyse change in BMI for patients treated with canagliflozin on an insulin background, because no BMI data from the canagliflozin CANVAS trial (DIA3008 insulin sub-study) were available.[3]

Table 32: NMA results for insulin background at 26 weeks (weight)

	Cana 100mg				Cana 300 mg				SUCRA
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Cana 300mg QD	0.57	0.17	[0.24 ; 0.90]	0%					86%
Exenatide 10µg BID	0.64	0.55	[-0.44 ; 1.71]	12%	0.07	0.54	[-1.01 ; 1.14]	45%	86%
Metformin	0.34	1.55	[-2.69 ; 3.39]	41%	-0.23	1.55	[-3.26 ; 2.83]	56%	76%
Cana 100mg QD					-0.57	0.17	[-0.90 ; -0.24]	100%	70%
Dapagliflozin 10mg QD	-0.41	0.31	[-1.01 ; 0.20]	91%	-0.98	0.31	[-1.59 ; -0.38]	100%	60%
Placebo	-2.10	0.17	[-2.43 ; -1.77]	100%	-2.67	0.17	[-3.00 ; -2.34]	100%	36%
Sitagliptin 100mg QD	-2.10	0.29	[-2.67 ; -1.53]	100%	-2.67	0.29	[-3.24 ; -2.11]	100%	35%
Pioglitazone 30mg QD	-3.28	4.30	[-11.72 ; 5.14]	78%	-3.85	4.30	[-12.31 ; 4.57]	82%	34%
Saxagliptin 5mg QD	-2.31	0.31	[-2.92 ; -1.70]	100%	-2.88	0.31	[-3.49 ; -2.27]	100%	27%
Glibenclamide	-7.64	9.31	[-25.82 ; 10.69]	79%	-8.21	9.31	[-26.36 ; 10.11]	81%	25%
Vildagliptin 100mg QD	-2.80	0.46	[-3.69 ; -1.90]	100%	-3.37	0.46	[-4.25 ; -2.47]	100%	16%

3.3.4 Summary of results for lipid levels

NMAs were not conducted for any of the lipid outcomes, so we present only the results from the systematic literature review for these outcomes below.

Dual therapy

In DIA3009 and DIA3006, both doses of canagliflozin (100mg and 300mg) in combination with metformin were associated with statistically significant greater improvements (increases) in HDL-cholesterol from baseline levels than the SU glimepiride combined with metformin or the DPP-4-i sitagliptin combined with metformin at week 52 (Table 33). There were no statistically significant differences between the canagliflozin arms and the active comparators in change from baseline in LDL-cholesterol levels, except in DIA3009 in which patients treated with canagliflozin 300mg experienced a statistically significant greater increase in LDL-cholesterol at week 52 (Table 34) than patients treated with glimepiride. In DIA3009, both canagliflozin 100mg and 300mg resulted in a statistically significant decrease in triglycerides from baseline when compared to glimepiride at week 52. In DIA2009, there were no statistically significant

differences in changes in triglycerides from baseline between either of the canagliflozin arms and the sitagliptin arm (Table 35).

Table 33: RCT results for dual therapy MET background (HDL-cholesterol)

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 mmol/L LOCF Mean (SD)	1.29 (0.346)	1.30 (0.330)		1.19 (0.301)
% change from baseline LS Mean (SE)	7.9 (0.8)	9.0 (0.8)		0.3 (0.8)
Diff. of LS Means % (SE) (minus glimepiride) (95% CI)	7.5 (1.0) (5.6;9.5)	8.6 (1.0) (6.7;10.6)		
DIA3006	**(n = 367)**	**(n = 368)**	**(n = 183)**	**DPP-4 i (sitagliptin) (n = 366)**
Value at Week 26 mmol/L LOCF Mean (SD)	1.28 (0.357)	1.32 (0.331)	1.18 (0.295)	1.21 (0.289)
% change from baseline LS Mean (SE)	10.3 (0.9)	12.1 (1.0)	3.7 (1.3)	5.0 (1.0)
Diff. of LS Means % (SE) (minus placebo) P-Value (95% CI)	6.6 (1.6) <0.001 (3.6;9.7)	8.4 (1.6) <0.001 (5.3;11.5)		1.3 (1.6) N/A (-1.8;4.3)
	(n = 368)	**(n = 367)**		**(n = 366)**
Change from baseline LS Mean (SE) mmol/L week 52[b]	0.12 (0.01)	0.14 (0.01)		0.06 (0.01)
% change from baseline[b]	11.2 (1.0)	13.2 (1.1)		6.0 (1.1)
Diff. of LS Means % change (minus sitagliptin at week 52)[a,b] (95% CI)	5.2 (2.5, 7.9)	7.2 (4.4, 10.0)		

mITT analysis
[a] Data from trial publication[22]
[b] Statistical comparison not performed[22]

Table 34: RCT results for dual therapy MET background (LDL-cholesterol)

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 mmol/L LOCF Mean (SD)	2.71 (0.933)	2.98 (1.032)		2.70 (0.910)
% change from baseline LS Mean (SE)	9.6 (1.9)	14.1 (1.9)		5.0 (1.9)
Diff. of LS Means (SE) % (minus glimepiride) (95% CI)	4.5 (2.3) (-0.0;9.1)	9.0 (2.4) (4.4;13.7)		
DIA3006	**(n = 367)**	**(n = 368)**	**(n = 183)**	**DPP-4 i (sitagliptin) (n = 366)**
Value at Week 26 LOCF Mean (SD)	2.88 (0.871)	2.97 (0.928)	2.66 (0.912)	2.86 (0.909)
Change from baseline LS Mean % (SE)	6.5 (1.7)	10.7 (1.8)	-1.5 (2.4)	4.1 (1.8)
Diff. of LS Means % (SE) (minus placebo) (95% CI)	7.9 (2.8) (2.4;13.5)	12.2 (2.9) (6.6;17.8)		5.5 (2.9) (-0.1;11.2)
	(n = 368)	**(n = 367)**		**(n = 366)**
Change from baseline LS Mean (SE) mmol/L week 52[b]	0.11 (0.04)	0.11 (0.04)		0.08 (0.04)
% LS mean (SE) change[b]	7.7 (1.7)	8.8 (1.8)		6.0 (1.8)
Diff. of LS Means % (minus sitagliptin at week 52)[a b], (95% CI)	1.7 (−2.8, 6.2)	2.8 (−1.8, 7.4)		

mITT analysis
[a] statistical comparison not performed as it was not a pre-specified analysis
[b] Data from trial publication[22]

Table 35: RCT results for dual therapy MET background (triglycerides)

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Value at Week 52 mmol/L LOCF Mean (SD)	1.87 (1.250)	2.02 (2.095)		1.95 (1.167)
% change from baseline LS Mean (SE)	-3.7 (2.5)	2.3 (2.5)		9.5 (2.5)

Diff. of LS Means (SE) (minus glimepiride) (95% CI)	-13.2 (3.2) (-19.4;-7.0)	-7.2 (3.2) (-13.4;-1.0)		
DIA3006	**(n = 367)**	**(n = 368)**	**(n = 183)**	**DPP-4 i (sitagliptin) (n = 366)**
Value at Week 26 mmol/L LOCF Mean (SD)	2.04 (1.305)	1.84 (1.032)	2.09 (1.341)	1.95 (1.389)
% change from baseline LS Mean (SE)	1.6 (2.6)	-1.4 (2.6)	3.2 (3.6)	1.0 (2.7)
Diff. of LS Means (SE) (minus placebo) P-Value (95% CI)	-1.6 (4.2) 0.701 (-9.9;6.7)	-4.6 (4.3) 0.276 (-13.0;3.7)		-2.3 (4.3) N/A (-10.6;6.1)
	(n = 368)	**(n = 367)**		**(n = 366)**
Change from baseline LS Mean (SE) week 52[a]	−0.12 (0.05)	−0.19 (0.05)		−0.15 (0.05)
LS mean (SE) % change[b]	1.9 (2.4)	2.8 (2.4)		−0.4 (2.5)
Diff. of LS Means per cent change (minus sitagliptin at week 52)[b] P-value (95% CI)	2.3 NS (−3.9, 8.5)	3.2 NS (−3.1, 9.5)		

mITT analysis
NS = not statistically significant
[a] Data from trial publication[22]

Triple therapy

In the triple therapy trial DIA3015, canagliflozin (300mg) combined with metformin and a sulfonylurea was associated with a statistically significant greater improvement in HDL-cholesterol at 52 weeks than the DPP-4-i sitagliptin combined with metformin and a sulfonylurea (Table 36). However, it was also associated with statistically significant greater increases in LDL-cholesterol than sitagliptin (Table 37). There were no statistically significant differences between the arms in change in triglycerides at 52 weeks (Table 38).

Table 36: RCT results for triple therapy, MET and SU background (HDL-cholesterol)

	Canagliflozin 300mg	Active comparator
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 mmol/L LOCF - Mean (SD)	1.26 (0.327)	1.18 (0.297)
% change from baseline - LS Mean (SE)	7.6 (0.9)	0.6 (0.9)
Diff. of LS Means (SE) (minus sitagliptin) P-Value	7.0 (1.2) N/A (4.6; 9.3)	

95% CI		

mITT analysis

Table 37: RCT results for triple therapy, MET and SU background (LDL-cholesterol)

	Canagliflozin 300mg	Active comparator
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 mmol/L LOCF - Mean (SD)	2.75 (1.043)	2.50 (0.911)
% change from baseline - LS Mean (SE)	11.7 (1.8)	5.2 (1.8)
Diff. of LS Means (SE) (minus sitagliptin) P-Value 95% CI	6.4 (2.4) N/A (1.7; 11.2)	

mITT analysis

Table 38: RCT results for triple therapy, MET and SU background (triglycerides)

	Canagliflozin 300mg	Active comparator
DIA3015	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Value at Week 52 LOCF mmol/L - Mean (SD)	2.02 (1.414)	1.96 (1.211)
% change from baseline - LS Mean (SE)	9.6 (2.8)	11.9 (2.9)
Diff. of LS Means (SE) (minus sitagliptin) P-Value 95% CI	-2.3 (3.9) N/A (-9.8; 5.3)	

mITT analysis

3.3.5 Summary of results for hypoglycaemia

Dual therapy

There was a statistically significant lower rate of hypoglycaemia (5-6%) for both doses of canagliflozin compared to glimepiride (34%) in the DIA3009 trial (Table 39). Severe hypoglycaemia rates were low generally, with a slightly higher incidence for glimepiride. In the DIA3006 trial rates of hypoglycaemia were similar between the active treatments between week 26 and 52, but higher for both doses of canagliflozin (6.8% each) compared to sitagliptin over the whole 52 weeks (no statistical significance tests reported). The difference in hypoglycaemia rates between the active comparator arms of the two trials can be explained by the fact that sulfonylureas are known to be associated with hypoglycaemia.

Table 39: RCT results for dual therapy MET background (hypoglycaemia)

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3009	**(n = 483)**	**(n = 485)**		**SU (glimepiride) (n = 482)**
Documented hypoglycaemia, week 52	27 (6%)[a]	24 (5%)[a]		165 (34%)
Severe hypoglycaemia, week 52[b]	2 (<1%)	3 (<1%)		15 (3%)
DIA3006[b,c]	**(n = 368)**	**(n = 367)**	**(n = 183)**	**DPP-4 i (sitagliptin) (n = 366)**
Documented hypoglycaemia, week 26-52	4.2%	5.0%	1.6%	4.7%
Documented hypoglycaemia, week 0-52	6.8%	6.8%	2.7%	4.1%
Severe hypoglycaemia, week 0-52	1	0	0	1

mITT
[a] p<0.0001 for both canagliflozin groups compared to glimepiride
[b] Data from trial publication[21]
[c] statistical significance tests not reported for this outcome

Table 40 reports the hypoglycaemia results of the fixed effect NMA for the metformin background at 52 weeks. There were eight trials in the network, with one direct comparison (canagliflozin 300mg versus 100mg (two trials)) of which there was no statistical heterogeneity present). The percentage that the treatment always ranks first was highest for placebo, dapagliflozin and saxagliptin (all 87%). For canagliflozin 300mg and 100mg the probabilities were 45% and 41%, respectively. All three of the DPP-4-i drugs included in the network (saxagliptin, sitagliptin and vildagliptin) ranked higher than canagliflozin. Of the DPP-4-i drugs the probability that canagliflozin 100mg reduces all hypoglycaemic events was highest for vildagliptin at 21%. For canagliflozin 300mg the corresponding probability was 28%. Canagliflozin ranked higher than the three sulfonylureas included in the network (glipizide, glimepiride, and glibenclamide), with a 100% probability of reducing all hypoglycaemia events, though sulfonylureas are generally known to be associated with hypoglycaemia.

Table 40: NMA results for dual therapy, MET background at 52 weeks (all hypoglycaemic events)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Placebo	5.03	11.97	[0.62; 37.38]	7%	4.68	11.38	[0.58;35.22]	7%	87%
Dapagliflozin 10mg QD	3.65	2.30	[1.44;9.93]	0%	3.43	2.14	[1.34; 9.23]	1%	87%
Saxagliptin 5mg QD	3.62	2.22	[1.41;9.75]	0%	3.39	2.07	[1.31; 9.19]	1%	87%
Sitagliptin 100mg QD	1.77	0.65	[0.97;3.44]	3%	1.66	0.61	[0.90; 3.23]	5%	65%
Vildagliptin 100mg QD	1.27	0.42	[0.71;2.33]	21%	1.19	0.39	[0.66; 2.19]	28%	54%
Cana 300mg QD	1.07	0.23	[0.71;1.61]	38%					45%
Cana 100mg QD					0.94	0.20	[0.62; 1.40]	62%	41%
Glipizide	0.19	0.09	[0.09;0.42]	100%	0.18	0.08	[0.08; 0.40]	100%	21%
Glimepiride	0.11	0.02	[0.07;0.16]	100%	0.10	0.02	[0.07; 0.15]	100%	12%
Glibenclamide	0.05	0.03	[0.02;0.13]	100%	0.05	0.03	[0.02; 0.12]	100%	1%

Triple therapy

The rate of hypoglycaemia and severe hypoglycaemia was similar for canagliflozin and for sitagliptin, both in combination with metformin and sulfonylurea in DIA3015. The relatively high rate of hypoglycaemia in both trial arms is likely explained by the use of sulfonylurea in the background regimen.

Table 41: RCT results for triple therapy, MET+SU background (hypoglycaemia)

	Canagliflozin 300mg	**Active comparator**
DIA3015[a,b]	(n = 377)	DPP-4-i (sitagliptin) (n = 378)
Documented hypoglycaemia (at least 1 episode) week 0-52	43.2%	40.7%
Severe hypoglycaemia week 0-52	4%	3.4%

[a] Data from trial publication[23]
[b] Statistical significance tests not reported for this outcome

Table 42 reports the hypoglycaemia results of the fixed effect NMA for the metformin and sulfonylurea background at 26 weeks. Eight trials were included in the network, with one direct comparison (long-acting insulin versus biphasic insulin, two studies), with no statistically

significant heterogeneity detected. The percentage that the treatment always ranks first was highest for placebo, exenatide 5 µg and linagliptin (100%, 80% and 78%, respectively). For canagliflozin 300mg and 100mg the probabilities were 38% and 56%, respectively. Canagliflozin ranked higher than the two insulins included in the network (long-acting and biphasic), with a 97% to 100% probability of reducing all hypoglycaemia events, though insulin is known to be associated with hypoglycaemia.

Table 42: NMA results for triple therapy at 26 weeks, MET and SU background (all hypoglycaemic events)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Placebo	2.58	0.78	[1.50; 4.54]	0%	3.30	0.91	[2.03; 5.59]	0%	100%
Exenatide 5µg BID	1.55	0.66	[0.74; 3.29]	12%	2.00	0.80	[0.99; 4.08]	3%	80%
Linagliptin 5mg QD	1.51	0.56	[0.78; 2.97]	11%	1.94	0.68	[1.04; 3.68]	2%	78%
Cana 100mg QD					1.29	0.33	[0.80; 2.08]	15%	56%
Exenatide 10µg BID	0.96	0.40	[0.46; 2.00]	55%	1.23	0.48	[0.62; 2.48]	28%	50%
Cana 300mg QD	0.78	0.20	[0.48; 1.25]	85%					38%
Sitagliptin 100mg QD	0.75	0.22	[0.43; 1.29]	85%	0.96	0.14	[0.72; 1.29]	60%	35%
Long-acting insulin	0.32	0.19	[0.12; 0.82]	99%	0.41	0.23	[0.16; 1.04]	97%	13%
Biphasic insulin	0.24	0.13	[0.10; 0.60]	100%	0.31	0.16	[0.13; 0.75]	100%	1%

Table 43 reports the hypoglycaemia results of the fixed effect NMA for the metformin and thiazolidinedione background at 26 weeks. The network comprised just two trials, with no direct comparisons. The percentage that the treatment always ranks first was highest for placebo, followed by canagliflozin 100mg, and sitagliptin (66%, 63% and 53%, respectively). Canagliflozin 300mg ranked lowest at 18%.

Table 43: NMA results for triple therapy, MET and TZD background (all hypoglycaemic events)

	Cana 100mg				**Cana 300 mg**				**SUCRA**
	Median	sd	CrI	Prob	Median	sd	CrI	Prob	
Placebo	1.02	1.94	[0.17 ; 6.18]	49%	2.21	3.62	[0.55 ; 11.62]	14%	66%
Canagliflozin 100mg					2.15	3.46	[0.54 ; 11.25]	14%	63%
Sitagliptin 100mg	0.86	2.37	[0.10 ; 7.09]	55%	1.89	4.54	[0.30 ; 13.72]	25%	53%
Canagliflozin 300mg	0.46	0.50	[0.09 ; 1.87]	86%	1.00				18%

Insulin background

The manufacturer reported that it was not possible to construct an insulin background network for all hypoglycaemic events due to lack of data at 26 weeks from the DIA3008 insulin sub-study.

Pooled analysis

The manufacturer reports hypoglycaemia rates from the canagliflozin RCTs included in their clinical SLR. The analyses includes a numerically pooled analysis of a set of RCTs of patients not on a background therapy associated with hypoglycaemia (including, amongst others, DIA3009 and DIA3006) (MS Appendix Table 180), and a set of studies including patients receiving background therapy with insulin or insulin secretagogues (e.g. sulfonylurea) (including, amongst others, DIA3015) (MS Appendix Table 181). The latter is not a pooled analysis but results are presented separately for each trial. In each set of studies rates of hypoglycaemia and severe hypoglycaemia are reported for canagliflozin 100mg, 300mg, the two doses combined, placebo, and any active comparator. As would be expected, the incidence of hypoglycaemia was higher for patients on a background therapy associated with hypoglycaemia compared to those on a background not associated with hypoglycaemia. The ERG has not reported these results here as the data include a wider set of studies than considered relevant to the decision problem, and the pooled analysis was not a formal meta-analysis that preserves randomised comparisons. Details can be found in MS Appendix 10.8 (Tables 180 and 181).

3.3.6 Summary of health related quality of life

Results for HRQoL are not reported in the main MS document. However, MS Appendix 10.14 reports an analysis of HRQoL and PRO measures conducted to quantify the utility associated with changes in BMI. Limited results are provided on changes in measures for the respective canagliflozin trials in the MS (e.g. in terms of change from baseline). Table 203 in the appendix shows which HRQoL and PRO instruments were included at which time points in a sub-set of four of the canagliflozin trials (three of which are the trials that the ERG consider relevant to this report), and the rationale for their inclusion in their analysis. The EQ-5D is considered to be the reference case and is reported in MS Appendix Table 204 for dual therapy (based on DIA3009), and in MS Appendix Table 210 for pooled triple therapy (based on DIA3002 and DIA3015). Results are not given by randomised treatment group and therefore the ERG have not reported

them here. There were no other HRQoL data in MS Appendix 10.14 that the ERG considered relevant to report here.

3.3.7 Sub-group analyses results

MS Table 15 reports the range of sub-group analyses conducted across the canagliflozin trials included in the submission. The MS reports results only for the pre-specified sub-group analyses according to baseline HbA_{1c} %, for the outcome of change in HbA_{1c} %. Analysis of this sub-group was done in five of the included trials, including two of the three trials the ERG considers relevant to this submission (DIA3009 and DIA3015), covering treatment with metformin background (Table 44) and treatment with metformin and sulfonylurea background (Table 45). The sub-group analyses show that difference between treatments were maintained, but larger absolute reductions in HbA_{1c} were observed in patients with baseline HbA_{1c} ≥ 9%.

Table 44: Pre-defined sub-group analyses by baseline HbA_{1c} %, MET background (HbA_{1c} %)

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

mITT

Table 45: Pre-defined sub-group analyses by baseline HbA_{1c} %, MET + SU background (HBA1c %)

Trial no. (acronym)	Canagliflozin 100mg	Canagliflozin 300mg	Placebo	Active comparator
DIA3015		**(n = 185)**		**Sitagliptin (n = 174)**
Baseline HbA_{1c} <8%[a]				
Change from baseline: LS Mean (SE)		-0.57 (0.079)		-0.31 (0.079)
Diff. of LS Means (SE) (minus sitagliptin at week 52)		-0.26 (0.082)		
95%CI		(-0.418;-0.097)		
Baseline HbA_{1c} 8% to <9%[b]				
		(n = 125)		**(n = 122)**
Change from baseline LS Mean		-1.15		-0.73
Diff. of LS Means (minus sitagliptin at week 52)		-0.41		
[redacted][c]				
		[redacted]		[redacted]
[redacted]		[redacted]		[redacted]
[redacted]		[redacted]		
[redacted]		[redacted]		

mITT
[a] For this sub-group the MS did not mark the results as AIC in MS Appendix 10.6.4.
[b] There appears to be an error in MS Table 37, with results for the <8% sub-group transposed by mistake. The ERG has reported data from the trial publication instead.
[c] Marked as AIC in MS Appendix 10.6.4, though note that the LS mean values have been reported in the trial publication.
[redacted]

BMI was also specified in the NICE scope as a relevant sub-group. However, the MS did not report it as the manufacturer states that efficacy of canagliflozin in terms of change in HbA_{1c} and secondary endpoints was not impacted by baseline BMI levels (MS Table 7) (NB. they also state that the economic model does not differentiate between patients based on their baseline BMI level). The manufacturer supplied the results of this sub-group to the ERG on request.
[redacted]

3.3.8 Summary of results for adverse events

The MS reports a numerically pooled analysis of adverse events for a pre-specified safety

dataset (the 'broad dataset') (MS section 6.9), comprising all active- and placebo- controlled trials examining both doses of canagliflozin (therefore including DIA3009, DIA3006, but not DIA3015 as that trial included only one canagliflozin dose).

Table 46 reports the overall summary of adverse events for the broad dataset, reproduced from MS Table 22 (Note, there is a discrepancy between the figures in MS Table 22 and some of the figures presented in the text on MS page 143 (6.9.2.1. AEs reported) and MS page 152 (6.9.2.8. Serious AEs) as well as with the figures presented in MS reference 118 ('Summary of clinical safety', Table 41). The set of figures in MS Table 22 show higher rates of AEs, and it appears they are from a later follow up time point – 1st July 2012. The ERG considers these to be the more relevant figures for this report. The all non-canagliflozin column in the table comprises placebo, the DPP4-i sitagliptin and the SU glimepiride.

Table 46: Overall summary of adverse events (broad dataset: safety analysis set)

Number (%) of patients with a least one adverse event of the following types	**Canagliflozin 100mg**	**Canagliflozin 300mg**	**All Cana**	**All non-Cana**
	N = 3,092	**N = 3,085**	**N = 6,177**	**N = 3,262**
Any AEs	2,369 (76.6)	2,375 (77.0)	4,744 (76.8)	2,473 (75.8)
AEs leading to discontinuation	173 (5.6)	224 (7.3)	397 (6.4)	164 (5.0)
AEs related to study drug	910 (29.4)	1,037 (33.6)	1,947 (31.5)	711 (21.8)
AEs related to study drug and leading to discontinuation	110 (3.6)	142 (4.6)	252 (4.1)	70 (2.1)
Serious AEs	417 (13.5)	406 (13.2)	823 (13.3)	445 (13.6)
Serious AEs leading to discontinuation	63 (2.0)	52 (1.7)	115 (1.9)	71 (2.2)
Serious AEs related to study drug	35 (1.1)	33 (1.1)	68 (1.1)	27 (0.8)
Serious AEs related to study drug and leading to discontinuation	17 (0.5)	14 (0.5)	31 (0.5)	10 (0.3)
Deaths	25 (0.8)	24 (0.8)	49 (0.8)	37 (1.1)

As Table 46 shows, there was a similar rate of any adverse events between canagliflozin and non-canagliflozin (76.8% for all canagliflozin and 75.8% for all non canagliflozin). There was a slightly higher percentage of patients discontinuing due to adverse events for all canagliflozin, and a 10% higher rate of patients experiencing adverse events related to the study drug for all canagliflozin. The percentage of patients with serious adverse events and the percentage of deaths was similar between canagliflozin and non-canagliflozin. MS Table 22 also reports

percentage of patients experiencing a range of specific adverse events grouped by body system (the ERG have not reproduced these here for brevity).

Analysis of just the three trials that the ERG consider relevant to the decision problem (based on DIA3009[21], DIA3006[22], and DIA3015[23] – as reported in their respective journal publications) showed a similar pattern to the broad dataset, with generally similar incidence of adverse events between canagliflozin and comparators sitagliptin and glimepiride.

Table 47 reports the percentage of patients in the broad dataset (regardless of the use of rescue medication) who experienced specific adverse events known to be associated with canagliflozin due to its mode of action. These include genital mycotic infections (GMI) (vulvovaginal candidiasis (VVC) in women and Candida balanitis in men), urinary tract infections, osmotic diuresis-related adverse events (increased urine output), and volume depletion (adverse events related to reduced intravascular volume (e.g. dizziness, hypotension). As would be expected, a higher percentage of patients receiving canagliflozin experienced these adverse events than non canagliflozin patients. The MS reports that GMIs and osmotic diuresis-related AEs were mild to moderate in intensity and rarely led to study discontinuation (MS section 6.9.2). The manufacturer states that approximately 65% of the GMI events in women were treated by a health care professional and 35% of the events by the patient herself (MS p. 147). The ERG notes that this contradicts the manufacturer's statement earlier in the MS that GMIs in women would require no extra use of health care resources because women can obtain topical and oral azole therapies over the counter (MS p. 43).

Table 47: Summary of specific adverse events (broad dataset: safety analysis set)

Number (%) of patients with a least one adverse event of the following types	**Canagliflozin 100mg**	**Canagliflozin 300mg**	**All Cana**	**All non-Cana**
Any female GMI	N=1289	N=1319	N=2608	N=1338
	190 (14.7)	184 (13.9)	374 (14.3)	42 (3.1)
Any male GMI	N=1803	N=1766	N=3569	N=1924
	131 (7.3)	164 (9.3)	295 (8.3)	30 (1.6)
	N = 3,092	**N = 3,085**	**N = 6,177**	**N = 3,262**
Urinary tract infections	254 (8.2)	250 (8.1)	504 (8.2)	218 (6.7)
Osmotic diuresis-related AEs	227 (7.3)	243 (7.9)	470 (7.6)	79 (2.4)
Volume depletion AEs[a]	99 (3.2)	141 (4.6)	240 (3.9)	78 (2.4)

[a] MS Section 6.9.2.6 reports that this is based on the placebo-controlled dataset (i.e. only the trials that featured a placebo arm: DIA3002, DIA3005, DIA3006, and DIA3012) though the percentages given have been calculated on the broad dataset. MS Reference 117, Table 19, reports that this is the broad dataset.

The manufacturer reports a pre-specified meta-analysis of the composite endpoint MACE (cardiovascular death, non-fatal myocardial infarction, and non-fatal stroke) and hospitalised unstable angina (MACE-plus) (MS section 6.9.2.8). The meta-analysis was performed in a pooled patient population (n=9,632) from nine Phase 2/3 studies of ≥12 weeks in duration (excluding DIA3015). The ongoing CANVAS study (DIA3008) contributed 45% (n=4,327) of patients included in the meta-analysis. The Cox proportional hazards model Hazard Ratio (HR) of canagliflozin (both doses pooled) to non-canagliflozin from the primary analysis of time to first adjudicated MACE-plus event was 0.91 (95% CI, 0.68-1.22). (see MS section 6.9.2.8.) The manufacturer reports that this meets the US FDA guidance to demonstrate an upper bound <1.8, indicating that there is no evidence of increased CV risk with canagliflozin. MS reference 8 (Table 203) reports an analysis of MACE-plus events for the individual components. The HRs for the individual components were all less than 1 with the upper bound of the 95% CI less than 1.8 with the exception of fatal or non-fatal stroke where the HR was 1.47 (0.83, 2.59). Due to the low number of MACE-plus events confidence intervals around the HRs are wide and indicate limited precision. Further, these data are from an interim analysis and the DIA3008 (CANVAS) trial is on-going. Caution is therefore advised in the interpretation of these results.

3.4 Summary

The MS provides a generally unbiased estimate of the treatment effect within the stated scope of the decision problem, although there are some uncertainties (as we detail below). The MS was based on a reasonably well conducted systematic review, with an NMA providing supporting evidence. Three RCTs,[21-23] judged to be of a high quality, were identified that provided head-to-head evidence for the efficacy of canagliflozin relative to active comparators. The manufacturer also included another three RCTs that compared canagliflozin to placebo which the ERG considered did not directly meet the decision problem, though their placebo arms were used to connect canagliflozin to other treatments in the manufacturer's NMA. The MS states that canagliflozin 100mg was non-inferior and canagliflozin 300mg was superior to the SU glimepiride and the DPP-4-i sitagliptin in reducing HbA_{1c} when used as part of dual therapy on a background of metformin, and that canagliflozin 300mg was superior to sitagliptin in reducing HbA_{1c} when used in triple therapy on a background of metformin and SU (though see below for discussion of superiority). It states that canagliflozin also consistently reduced

weight and blood pressure. It concludes that canagliflozin was well tolerated, with GMIs in women the most common adverse event.

The manufacturer's interpretation of the evidence in is on the whole appropriate and justified. However, the ERG identified the following exceptions and uncertainties:

- The three relevant canagliflozin RCTs[21-23] only examined the efficacy of canagliflozin relative to sitagliptin (in dual and triple therapy) and glimepiride (in dual therapy). There is a lack of direct evidence for the efficacy of canagliflozin in comparison to other agents and regimens, including insulin.
- The per protocol analyses, reported in the trial journal publications and CSRs, show less favourable results for canagliflozin 300mg as dual therapy than the mITT analyses results reported in the MS. In the per protocol analyses in both the dual therapy trials, canagliflozin 300mg was only non-inferior, and not superior (as it was in the mITT analyses), to sitagliptin and glimepiride. The MS does not discuss the implications of differences between the mITT and per protocol results.
- The manufacturer states that the inclusion of supplementary evidence from the NMA means that there is "data to support the efficacy of canagliflozin across all lines of therapy relevant to the decision problem" (MS p. 158). However, the ERG considers the NMA results to only provide an indication of the efficacy of canagliflozin relative to other active comparators, and that the results should be interpreted with caution due to the low volume of direct evidence for some of the loops in the networks.
- The manufacturer's NMA assesses effects at 26 weeks duration for triple therapy, rather than the 52 weeks duration used for dual therapy. The MS justifies this by stating that there were limited data available for 52 weeks for the comparator treatments. They state that analyses at 52 weeks support the extrapolation of treatment effects from 26 to 52 weeks. The ERG notes that the manufacturer discarded studies from the NMA that did not report outcomes at 26 +/- 10 weeks for the MET and SU, and MET and TZD backgrounds. It is possible that some of these discarded studies could have been used to construct 52 week networks for triple therapy.
- The MS only provides evidence for the efficacy of canagliflozin among patients treated on therapy backgrounds that included metformin. This does not reflect all the potential uses of canagliflozin outlined within its licenced indication.
- There is a lack of evidence for the longer-term efficacy and safety of canagliflozin, as most outcomes were measured up to 52 weeks in the clinical trials.

- The MS states that “canagliflozin is well tolerated” (MS p. 157), but the data presented shows a 10% higher rate of adverse events related to the study drug among patients treated with canagliflozin relative those treated with sitagliptin, glimepiride or placebo.
- Similarly, the MS states that “canagliflozin demonstrated lower rates of discontinuation […] than active comparators (sitagliptin and glimepiride)” (MS p. 159), but the data presented show that a slightly higher proportion of patients treated with canagliflozin experienced AEs leading to discontinuation (6.4%) than patients treated with placebo or these other agents (5.0%).
- [redacted]

4 ECONOMIC EVALUATION

4.1 Overview of manufacturer’s economic evaluation

The manufacturer’s submission to NICE includes:

iii) A review of published economic evaluations of treatments for T2DM.

iv) A report of an economic evaluation undertaken for the NICE STA. The cost effectiveness of canagliflozin in dual therapy with metformin is compared with dapagliflozin, DPP-4-i, GLP-1-a, sulfonylurea and thiazolidinedione (pioglitazone). The cost effectiveness of canagliflozin in triple therapy with metformin and sulfonylurea or is compared with DPP-4-i, GLP-1-a and insulin. The cost effectiveness of canagliflozin in triple therapy with metformin and thiazolidinedione (pioglitazone) is compared with DPP-4-i. The cost effectiveness of canagliflozin as an add-on to insulin is compared with DPP-4-i, GLP-1-a, and dapagliflozin.

Manufacturer’s review of published economic evaluations

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of canagliflozin for T2DM, agents in the class of SGLT-2 inhibitors (dapagliflozin) and intended comparators (DPP-4-i, GLP-1-a, sulfonylureas) (MS p162). Details of the search strategy are reported in a secondary reference.[4] In addition the MS (p170) reports a systematic

search for economic simulation models (full details reported in a secondary reference[4]). As reported in section 3.1.1 of this report, the ERG considered the search strategy to be fit for purpose.

The inclusion and exclusion criteria for the systematic review are listed in section 7.1.1 of the MS, page 162. The MS included studies if they were economic evaluations, undertaken in UK, Canada or Australia, and were published after June 2003 (<10 years old). Studies were required to report incremental costs and at least one measure of efficiency, preferably as an incremental cost-effectiveness ratio (ICER) (from MS ref 123[4]). Studies were excluded if they were not T2DM, not written in English, used cost-minimisation or cost-consequence methods, or were limited to individual complications of T2DM. The searches identified 289 references of which 96 studies were selected for abstract review and 66 retrieved for full article review.[4] Fifty two studies covering 44 unique studies were included in the review. No studies of canagliflozin were included;[4] four studies of dapagliflozin, 23 of a DPP-4-i, 20 of a GLP-1-a, one a TZD (rosiglitazone), and one of insulin glargine were included. Three studies included more than one therapy. Fourteen studies retrieved as full articles were excluded, reasons for exclusion were not specified in the MS or the secondary reference.[4]

The search for existing simulation models[4] required these to be capable of estimating the cost-effectiveness of competing treatment alternatives for T2DM in the UK setting, be unique and have publically available details. For the search for models 1308 references were identified and 921 were selected for abstract review (MS Ref 123[4]). Seventy three studies met the inclusion criteria (of which 21 were included). Reasons for excluding the 848 studies selected for abstract review were not reported.

The MS assessed the quality of cost effectiveness studies using a series of questions which were based on the format developed by Drummond and Jefferson.[30] This assessment of quality was only undertaken on studies that were not Health Technology Assessment (HTA) reports. Very little interpretation or conclusions of this quality assessment were provided in the MS, reporting that ‘the 14 studies were generally of acceptable quality’ and that ‘the level of detail of reporting varied widely’, MS p167.

The systematic review of cost effectiveness studies appears to have been undertaken to fulfil the requirements of the submission but does not appear to have been used to inform the choice

of economic model used. The MS does not make any general conclusions about the systematic review. In the secondary reference[4] the general conclusions focus on the key issues that have previously been flagged up by HTA bodies with existing cost effectiveness studies and models. These key modelling features include the inclusion of non-significant treatment effects, unconventional (not conservative) disutility weights for weight gain and hypoglycaemic events, and undue uncertainty in clinical efficacy data generated with indirect methods such as NMA and mixed treatment comparisons (MS Ref 123[4] p40).

CEA Methods

The cost effectiveness analysis (CEA) uses a stochastic micro-simulation model (the ECHO-T2DM model) to estimate the cost-effectiveness of canagliflozin compared with SU, TZD, DPP-4-i, GLP-1-a, dapagliflozin and insulin in adults with T2DM (MS Table 32 p175). All simulations reported in the MS were run with both patient-level (first order) and parameter (second order) uncertainty. Base case simulations were run with 1000 cohorts of 1000 patients. The model adopts a lifetime horizon (40 years) and has an annual cycle length.

The results from the economic evaluation are presented for the base case assumptions, which are summarised in MS Table 37 (MS p187-9). The key assumptions appear to be that the baseline characteristics from the clinical trial programme are generalisable to the UK population; that the treatment effects from the NMA reported at 26 weeks were equivalent to what would be observed at 52 weeks; and that treatment effects are immediately reversed at discontinuation.

The MS also present sub-group analyses for patients with an HbA_{1c} exceeding 9% (75mmol/mol) at baseline (MS Section 7.9 p249). No other additional subgroup analyses are presented.

Overview of health states used in the model

The model has three parallel sets of micro-vascular complications and five types of macro-vascular complications represented as Markov states (MS Section 7.2.4 p171). There is a health state where the patient is free from complications, and then micro-vascular health states relating to chronic kidney disease (includes 7 stages); neuropathy (includes 5 conditions); and retinopathy (includes 4 conditions). Patients can also experience macro-vascular complications (ischaemic heart disease [IHD], myocardial infarction [MI], chronic heart failure [CHF], stroke,

and peripheral vascular disease [PVD]). IHD and MI can occur more than once. There is also a health state of death.

The model also incorporates heath states and events associated with pharmacological treatment, including (MS p173):

- Patient weight change
- Hypoglycaemic events (separately by mild, moderate, and severe)
- Upper and lower UTIs
- GMI
- Gastrointestinal upset, including nausea

Overview of treatment effects

The principal treatment effectiveness data (HbA_{1c} and BMI) are taken from the NMA for dual, triple and add-on insulin therapy (MS p178 and model data input spreadsheet). For dual therapy the NMA results at 52 weeks were used, for triple therapy 26 week data were used. The MS states that the reason for the difference in treatment durations was due to limitations in the data for the comparators. For canagliflozin the relevant RCTs were DIA3006 (canagliflozin + MET versus DPP-4-i + MET); DIA3009 (canagliflozin + MET versus SU + MET); DIA3002 (canagliflozin + MET + SU versus placebo + MET + SU), DIA3015 (canagliflozin + MET + SU versus DPP-4-i + MET + SU), DIA3012 (canagliflozin + MET + TZD versus placebo + MET + TZD), DIA3008 (canagliflozin + SU versus placebo + SU added onto insulin). Only three of these trials were deemed relevant to the NICE scope by the ERG (see Section 3). Hypoglycaemic events and selected adverse events were also included.

Overview of Health related quality of life

HRQoL is applied in the model using utility from a published source (CODE-2 EQ-5D modelling study [31]) as the clinical trial utility data was unable to be used due to the short duration of the studies and likely confounding. Baseline utility from the CODE-2 regression analysis has a range of disutilities for a range of co-morbidities, complications and adverse events applied over each modelling cycle, based on probabilities of occurrence of these events from published sources. When multiple co-morbidities and complications are experienced these disutilities are additive. MS Table 38, (p198-199) shows the disutilities for each co-morbidity, complication and adverse event (See Section 4.2.5). Disutilities for some states (fear of hypoglycaemia and aversion to injections) were excluded, and a justification for these exclusions is given on MS

p201. As canagliflozin is an oral agent and is not associated with higher rates of hypoglycaemia than other treatments, the lack of inclusion of such disutilities is a conservative assumption.

Overview of resources and costs

Resources and costs were sourced from NHS reference costs and payment by results tariffs and where these were not specific enough for T2DM, UK clinical guidelines and other published sources were also used (MS p202). A comprehensive literature search was conducted to identify published peer-reviewed studies that evaluated costs associated with common diabetes-related complications or co-morbidities that were relevant for the ECHO-T2DM model (MS p203).

Costs for the different health states associated with co-morbidities, complications and adverse events were applied in the model. An advisory group was used to inform decisions about the relevance of various adverse events (for example the degree of hypoglycaemic episodes, GMI, UTI, and hypovolaemic events) and self-monitoring of blood glucose costs.

No administration costs were applied for any treatment. There were no additional monitoring costs applied for canagliflozin for testing renal function using the estimated glomerular filtration rate (eGFR) as the MS states that these are covered in standard monitoring. The ERG clinical expert agrees that this is appropriate. The costs used for adverse events are in MS Table 40, p208.

Drug dosing was applied as per their respective recommended use and drug acquisition costs were taken from the British National Formulary (BNF) July 2013.[32] A single comparator drug within each class of treatment was chosen (full details on MS p205 and Table 39 (p206)). Treatments for complications and co-morbidities are presented in Table 41 (page 208). See also Section 4.2.7 for further detail and critique.

Assessment of uncertainty

Twenty one scenario analyses were performed and these are listed on MS p209-214 (the MS states 21 were undertaken although the numbering in MS Table 44 suggests there may have been more). The MS focuses on three of these; dose modification, weight loss and treatment rebound. The remainder are reported in MS Appendix 12. Deterministic sensitivity analyses

(DSA) are reported to have been undertaken on 35 parameters, or sets of parameters, which were varied by plus or minus 20% (MS p215-6).

Probabilistic sensitivity analysis (PSA) is a default feature of the ECHO-T2DM model. Table 46 (MS p217) reports the parameters and distributions for PSA. More detail is in the MS Technical report.

Results from the scenario, DSA and PSA are presented in MS pages 237-272.

The model validation (MS p274-278) uses five principal forms: face validity, verification, cross-validation, external and internal validation and predictive validation (modelling cross validation).

CEA Results

Results from the economic model are presented as incremental cost per QALY gained for canagliflozin 100mg and 300mg in dual therapy, triple therapy and add-on insulin therapy. Base case results are presented in MS Section 7.7.6 (p233-237). Subgroup analyses are presented in MS p279-281.

For the base case analyses pairwise comparisons were generated by the model.

For dual therapy combinations incremental costs per QALYs gained were calculated compared with TZD, SU, dapagliflozin, DPP-4-i and GLP-1-a. Canagliflozin 100mg was dominated by TZD. Other ICERs for canagliflozin 100mg were £1,537 compared with SU, £8,674 for dapagliflozin, £97 for DPP-4-i and £50,005 for GLP-1-a where canagliflozin was less expensive and less effective than GLP-1-a (see Table 48). In the base case canagliflozin 300mg is dominated by TZD and led to ICERs of £4,899 compared with SU, £18,349 compared with DPP-4-i, £27,419 compared with dapagliflozin, and £76,214 compared with GLP-1-a where canagliflozin was less expensive and less effective (see Table 48). Table 48 incorporates results from MS Table 63 (MS p234) where it is stated that canagliflozin 300mg dominates TZD. Canagliflozin 300mg is in fact dominated by TZD because it is more expensive and is associated with fewer total QALYs. This is shown in Table 48.

For triple therapy combinations with metformin and SU, canagliflozin 100mg dominates DPP-4-i and GLP-1-a. The ICER compared with insulin LA was £263 (see Table 49). Canagliflozin

300mg dominates GLP-1-a and has ICERs of £13,287 and £607 for DPP-4-i and insulin LA respectively (Table 49).

For triple therapy combinations with metformin and TZD, canagliflozin 100mg had an ICER of £1,095 compared with DPP-4-i and canagliflozin 300mg had an ICER of £21,430 compared with DPP-4-i as seen in Table 49.

Table 48: Manufacturer's base case results for canagliflozin in dual therapy combinations (MET background), with correction (see footnote b)

Comparator	Total costs	Total LYG	Total QALYs	Incremental costs	Incremental LYG	Incremental QALYs	ICER
Canagliflozin 100mg							
TZD	£23,236	13.667	9.119				
Canagliflozin 100mg	£26,069	13.669	8.960	£2,833	0.001	-0.159	Dominated
SU	£25,894	13.664	8.751				
Canagliflozin 100mg	£26,182	13.666	8.938	£288	0.003	0.188	£1,537
Dapagliflozin	£26,086	13.673	8.923				
Canagliflozin 100mg	£26,149	13.678	8.930	£63	0.005	0.007	£8,674
DPP-4-i	£26,100	13.688	8.928				
Canagliflozin 100mg	£26,102	13.692	8.941	£1	0.003	0.013	£97
GLP-1-a	£28,390	13.650	8.970				
Canagliflozin 100mg	£25,966	13.644	8.921	-£2,424	-0.006	-0.048	£50,005[a]
Canagliflozin 300mg							
TZD	£23,204	13.665	9.092				
Canagliflozin 300mg	£26,556	13.663	8.951	£3,353	-0.003	-0.141	Dominated[b]
SU	£25,753	13.671	8.761				
Canagliflozin 300mg	£26,729	13.690	8.961	£976	0.019	0.199	£4,899
DPP-4-i	£25,926	13.644	8.902				
Canagliflozin 300mg	£26,501	13.655	8.933	£576	0.011	0.031	£18,349
Dapagliflozin	£26,044	13.670	8.910				
Canagliflozin 300mg	£26,668	13.679	8.932	£625	0.009	0.023	£27,419
GLP-1-a	£28,515	13.652	8.972				
Canagliflozin 300mg	£26,624	13.657	8.947	-£1,892	0.005	-0.025	£76,214[c]

[a]Canagliflozin 100mg is both cheaper and less effective than GLP-1-a, therefore the ICER is for GLP-1-a to replace canagliflozin 100mg
[b]MS reports that canagliflozin 300mg dominates TZD however this should state that canagliflozin 300mg is dominated by TZD as canagliflozin 300mg is more costly and less effective than TZD.
[c]Canagliflozin 300mg is both cheaper and less effective than GLP-1-a, therefore the ICER is for GLP-1-a to replace canagliflozin 300mg

Table 49: Manufacturer's base case results for canagliflozin in triple therapy combinations

Comparator	Total costs	Total LYG	Total QALYs	Incremental costs	Incremental LYG	Incremental QALYs	ICER
Canagliflozin 100mg (Metformin+SU)							
DPP-4-i	£25,776	12.964	7.060				
Canagliflozin 100mg	£25,734	12.967	7.076	£-42	0.004	0.016	Dominates
Insulin LA	£25,577	12.972	5.193				
Canagliflozin 100mg	£25,712	12.982	5.707	£135	0.010	0.514	£263
GLP-1-a	£27,119	12.978	7.101				
Canagliflozin 100mg	£25,822	12.980	7.102	-£1,297	0.002	0.001	Dominates
Canagliflozin 300mg (Metformin+SU)							
DPP-4-i	£26,196	12.967	7.047				
Canagliflozin 300mg	£26,657	12.976	7.082	£461	0.009	0.035	£13,287
Insulin LA	£26,326	12.992	5.139				
Canagliflozin 300mg	£26,705	12.993	5.763	£379	0.002	0.624	£607
GLP-1-a	£27,153	12.936	7.068				
Canagliflozin 300mg	£26,468	12.944	7.072	-£685	0.008	0.004	Dominates
Canagliflozin 100mg (Metformin+TZD)							
DPP-4-i	£22,557	8.294	8.287				
Canagliflozin 100mg	£22,565	12.597	12.595	£7	0.002	0.007	£1,095
Canagliflozin 300mg (Metformin+TZD)							
DPP-4-i	£22,586	12.606	8.274				
Canagliflozin 300mg	£23,277	12.623	8.306	£691	0.017	0.032	£21,430

Table 50 shows the base case results for the 'add on' to insulin therapy combinations. Canagliflozin 100mg had an ICER of £1,340 compared with DPP-4-i where canagliflozin was both less expensive and less effective. Canagliflozin 100mg dominated dapagliflozin and had an ICER of £12,915 compared with GLP-1-a, where canagliflozin was both less expensive and less effective. Canagliflozin 300mg had ICERs of £7,975 compared with DPP-4-i; £5,992 compared with dapagliflozin; and £35,575 compared with GLP-1-a, where canagliflozin 300mg was both less expensive and less effective than GLP-1-a.

Table 50: Manufacturer's base case results for canagliflozin as 'add on' to insulin therapy

Comparator	Total costs	Total LYG	Total QALYs	Incremental costs	Incremental LYG	Incremental QALYs	ICER
Canagliflozin 100mg							
DPP-4-i	£29,423	10.858	5.612				
Canagliflozin 100mg	£29,410	10.855	5.602	-£13	-0.003	-0.010	£1,340[a]
Dapagliflozin	£29,580	10.852	5.550				
Canagliflozin 100mg	£29,508	10.846	5.553	-£72	-0.007	0.003	Dominates
GLP-1-a	£30,229	10.834	5.653				
Canagliflozin 100mg	£29,393	10.836	5.588	-£836	0.002	-0.065	£12,915[a]
Canagliflozin 300mg							
DPP-4-i	£29,328	10.829	5.585				
Canagliflozin 300mg	£29,650	10.838	5.625	£322	0.009	0.040	£7,975
Dapagliflozin	£29,618	10.853	5.572				
Canagliflozin 300mg	£29,946	10.863	5.627	£327	0.009	0.055	£5,992
GLP-1-a	£30,333	10.851	5.636				
Canagliflozin 300mg	£29,807	10.859	5.621	-£526	0.007	-0.015	£35,575[b]

[a]Canagliflozin 100mg is both cheaper and less effective than the other therapies, therefore the ICER is for canagliflozin 100mg to be replaced
[b]Canagliflozin 300mg is both cheaper and less effective than GLP-1-a, therefore the ICER is for GLP-1-a to replace canagliflozin 300mg

In their clarification letter to the NICE and the ERG the manufacturer presents revised base cases for canagliflozin versus dapagliflozin and GLP-1-a in dual therapy, and canagliflozin versus GLP-1-a and insulin in triple therapy (MET+SU background). These revised base cases are based on assumptions about BMI change rather than actual BMI changes which were observed in the key trials. The revised base cases show improvements in the cost-effectiveness of canagliflozin against the comparators, except in the case of canagliflozin 300mg in triple therapy compared to insulin, where the ICER increases slightly, by £64.

The MS summarises the results of the PSA (MS page 237-272) for each combination (dual, triple and add-on to insulin therapy).

For dual therapy the MS (p245) states that there is a 92.9% probability of canagliflozin 100mg and metformin being cost-effective, relative to metformin and SU, at a willingness-to-pay threshold (WTP) of £20,000 per QALY gained. For canagliflozin 300mg in combination with metformin the MS (p246) reports that there is a 89.5% probability of being cost-effective, relative

to metformin and SU, at a WTP threshold of £20,000 per QALY gained. Canagliflozin 100mg in combination with metformin is reported to have a 54.5% probability of being cost-effective at a WTP threshold of £20,000 per QALY gained compared with metformin and DPP-4-i (MS page 247). The probability of canagliflozin 300mg and metformin being cost-effective at a WTP threshold of £20,000 per QALY gained, relative to metformin and DPP-4-i, was reported in the MS as 51.6% (MS page 248).

For triple therapy the MS (p250) states that there is a 55.7% probability of canagliflozin 100mg with metformin and SU being cost-effective, relative to metformin, SU and DPP-4-i, at a WTP threshold of £20,000 per QALY gained. Canagliflozin 300mg in combination with metformin and SU was reported to have a 55.1% probability of being cost-effective, relative to metformin, SU and DPP-4-i, at a WTP threshold of £20,000 per QALY gained (MS page 251).

As an add-on to insulin, canagliflozin 100mg is reported to have a 44.7% probability of being cost-effective at a WTP threshold of £20,000 per QALY gained relative to insulin and DPP-4-i (MS page 252). Relative to insulin and DPP-4-i, canagliflozin 300mg has a 60.7% probability of being cost-effective at a WTP threshold of £20,000 per QALY gained (MS page 253).

Results at the WTP threshold of £30,000 per QALY gained for each comparison were also provided in the MS, pages 237-72. In general these are similar to the £20,000/QALY ranges.

The MS states (p161) that overall the results of the economic evaluation show a consistent picture across a range of comparisons, demonstrating that canagliflozin as an add-on therapy to other AHAs is a cost-effective treatment option.

4.2 Critical appraisal of the manufacturer's submitted economic evaluation

Manufacturer's review of published economic evaluations

The MS presents a summary overview of a systematic review provided in a secondary reference.[4] This systematic review appears to have used a comprehensive search for studies and methods appear to be reasonable, although only a selection of included studies was quality assessed. The publication[4] does not mention TZD as a possible comparator in the inclusion criteria but does include one study of TZD so the ERG assume this is an omission. As noted above no studies of canagliflozin were identified.[4]

Critical appraisal of manufacturer's submitted economic evaluation

The ERG has considered the methods applied in the economic evaluation in the context of the critical appraisal questions listed in Table 51 below, drawn from common checklists for economic evaluation methods (e.g. Drummond and colleagues[30]). The critical appraisal checklist indicates that overall the manufacturer follows recommended methodological guidelines.

Table 51: Critical appraisal checklist of economic evaluation

Item	Critical Appraisal	Reviewer Comment
Is there a well-defined question?	Yes	In line with NICE scope, although some comparators were not included in the different therapy combinations
Is there a clear description of alternatives?	Yes	Comparisons made with SU, TZD, DPP-4-i, GLP-1-a, dapagliflozin and insulin according to their respective marketing authorisations
Has the correct patient group / population of interest been clearly stated?	Yes	For base case population: • people with T2DM requiring either dual therapy, triple therapy or as an add-on therapy with insulin • as per marketing authorisation Patients with an HbA_{1c} exceeding 9% (75mmol/mol) at baseline were considered as subgroups. (Discussed in sections 4.2.2)
Is the correct comparator used?	Partial	Scoped comparators for dual therapy with SU were not included; two scoped comparators (TZD and insulin) for triple therapy with metformin and SU were not included; two scoped comparators (GLP-1-a and insulin) for triple therapy with metformin and TZD were not included. (Discussed in section 2.3 and 4.2.3)
Is the study type reasonable?	Yes	
Is the perspective of the analysis clearly stated?	Yes	Noted in the decision problem as NHS and Personal Social Services (PSS)
Is the perspective employed appropriate?	Yes	Discussed in sections 4.2.6/4.2.7 for costs and 4.2.5 for outcomes
Is effectiveness of the intervention established?	Yes	Discussed in section 4.2.4
Has a lifetime horizon been used for analysis (has a shorter horizon been justified)?	Yes	Discussed in section 4.2.1 A 40 year time horizon is used
Are the costs and consequences consistent with the perspective employed?	Yes	Discussed in sections 4.2.6/4.2.7 for costs and4.2.5 for outcomes
Is differential timing considered?	Yes	3.5% for costs and health benefits, tested in sensitivity analysis at 0% and 6%
Is incremental analysis performed?	Yes	Described in section 4.1. Comparators discussed in section 4.2.3
Is sensitivity analysis undertaken and presented clearly?	Yes	Described in section 4.1. Discussed in section 4.2.9

Superseded - see erratum

NICE reference case

The NICE reference case requirements have also been considered for critical appraisal of the submitted economic evaluation in Table 52. The submitted evaluation generally conforms with the NICE reference case, key issues are that two of the comparators stated in the NICE scope were excluded from the MS decision problem, and the population used as the source of preference data for utility may not be wholly representative of the population of England.

Table 52: NICE reference case requirements

NICE reference case requirements:	Included in submission	Comment
Decision problem: As per the scope developed by NICE	No	Two comparisons stated in the scope were not included in the decision problem, discussed more in Section 2.3 of this report
Comparator: Alternative therapies routinely used in the UK NHS	Yes	Discussed in sections 3.1/ 3.2/ 3.3/ 3.4/ 4.2.3
Perspective on costs: NHS and PSS	Yes	
Perspective on outcomes: All health effects on individuals	Yes	Discussed in section 4.2.5
Type of economic evaluation: Cost effectiveness analysis	Yes	
Synthesis of evidence on outcomes: Based on a systematic review	Yes	Discussed in section 4.2.4
Measure of health benefits: QALYs	Yes	Discussed in section 4.2.5
Description of health states for QALY calculations: Use of a standardised and validated generic instrument	Yes	Discussed in section 4.2.5
Method of preference elicitation for health state values: Choice based method (e.g. TTO, SG, not rating scale)	Yes	Discussed in section 4.2.5 Predominantly EQ-5D
Source of preference data: Representative sample of the public	Unclear	Discussed in section 4.2.5 Data from a European study, unclear whether tariff used representative of UK population.
Discount rate: 3.5% pa for costs and health effects	Yes	Described in section 4.1

Overall the methods in the MS appear to be reasonable and the methods and data inputs conform to NICE methodological guidance. The decision problem does not meet the NICE scope, but the MS does meet the decision problem as outlined on pages 46-49 of the MS.

4.2.1 Modelling approach / Model Structure

The MS notes that models of T2DM must be capable of modelling complex treatment intensification algorithms, be able to extrapolate the effects of interventions observed in short-term trials over long time horizons; and be able to model patients moving on to dual or triple (or

later) therapies (MS p169). The MS describes a systematic literature review of published T2DM models which was undertaken to assess the suitability of existing models (MS p170). Twenty one models were identified and the MS notes that a subset of these models is appropriate to estimating the cost-effectiveness of canagliflozin in the UK setting (MS p170).

The ECHO-T2DM model was selected by the manufacturer for use in the submission but no justification for this choice is given in the MS. The model Technical Report states that work began on ECHO-T2DM in 1997 and that funding for the development of ECHO-T2DM has primarily been provided by Janssen Pharmaceutical Services, LLC, although the final form of ECHO-T2DM is the responsibility of The Swedish Institute for Health Economics.[5]

ECHO-T2DM is a stochastic patient-level simulation model of competing treatments for T2DM. Cohorts of individual hypothetical patients are created and simulated over time using Monte Carlo (first order uncertainty) techniques. Parameter (second order) uncertainty is captured using many cohorts of patients in which each cohort is assigned unique values of the key parameters (MS p167). Markov health states reflect important microvascular and macrovascular events. It is possible to turn off second order uncertainty.[5] The model has over 800 user-definable input parameters.

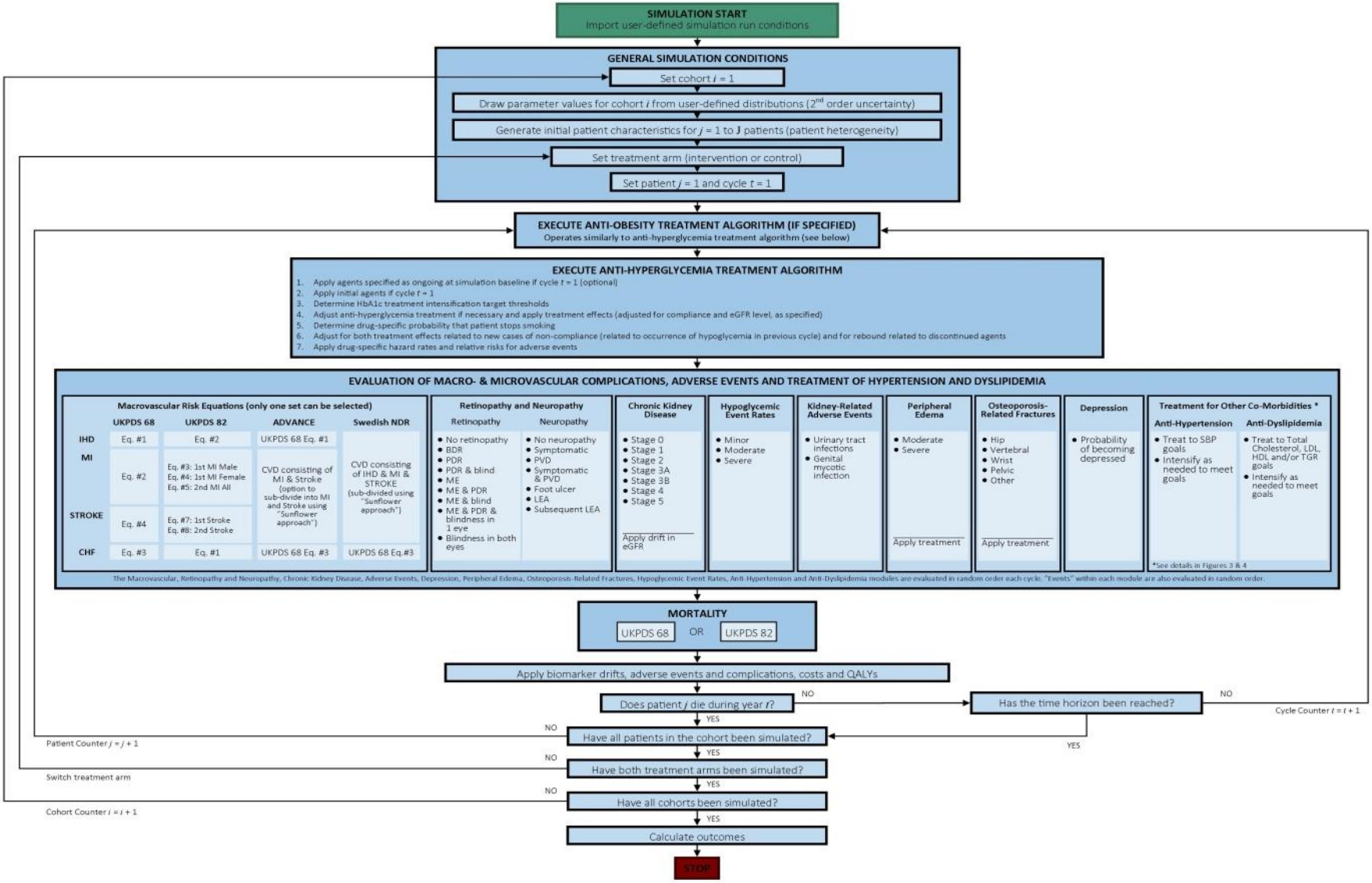
A patient-level model simulates treatment and response pathways for individual patients, and the outputs are mean costs, effectiveness or cost-effectiveness measures for a sample of individuals. Patient-level simulation models are commonly used in T2DM because it is a disease associated with multiple health states and sequences of treatment, and has a time dependent prognosis. There is also interaction between metabolic variables (e.g. HbA_{1c}) and comorbidities such as stroke and neuropathy. This can mean that the average of risks is not same as the risk for the patient with ‘average’ characteristics,[33] and consequently a patient-level simulation is preferred to a cohort model.

It is necessary to simulate a very large number of patients in a patient-level model in order for the simulated sample to give an accurate value for the population cost-effectiveness measure for each input parameter set.[34] Brennan and Chilcott [33] state that in practice they use 10,000 patients, although have used more when the treatment differences are small.

ECHO-T2DM is programmed in R with Microsoft Excel® front- and back-end interfaces. The complete set of files required for installation is described in the model User Guide[6] although one R file included in the ERG installation is missing from the User Guide list (evopred.R). The Technical Report states that ECHO-T2DM is built of 30 inter-linked modules in order to minimize the risk of programming error and to maximize the ease of modifying existing or adding new features.[5] The overarching relationship of these modules to each other, i.e. the processing flow, and where key functions and input values are located in the code, is not described in either the Technical Report or User Guide.[5;6] The ERG installation included a set of 29 R code files which in total contain around 20,000 lines of R code. The Excel front-end contains additional VBA code.

A schematic of the model is given in MS figure 59 (MS p168), reproduced as Figure 1 below. The sources of evidence used to develop and inform the structure of the model are described in the Technical Report[5] and the model structure is consistent with this evidence. The Technical Report notes that Janssen Pharmaceutical Services, LLC has reviewed and commented on the model and has provided access to clinical experts during model development.[5] The model has been presented at numerous conferences to experts in T2DM modelling, including the Fifth and Sixth Mount Hood Challenges (MS p274).

Figure 1: Schematic structure of the ECHO-T2DM model (reproduced from MS figure 59 p168)



Intermediate trial outcome measures such as HbA_{1c}, SBP, lipids, and BMI are linked to final clinical outcomes such as micro- and macro-vascular events, amputation, blindness and mortality through microvascular and macrovascular risk equations. The MS observes that this is consistent with other diabetes models (MS p182). Patients can begin with micro- and/or macrovascular complications at baseline and/or they can develop them during the course of the simulation. The complications are chronic and progressive except for the possible reversion from macroalbuminuria to microalbuminuria.[5]

The model uses risk equations from UKPDS 68 to simulate the macro-vascular complications of IHD, MI, stroke, and congestive heart failure, (CHF) given key biomarker values (e.g. age, BMI, HbA_{1c}, SBP etc).[35] It is possible to vary this equation set in scenario analysis. Neuropathy health states and allowable transitions are given in Figure 5 of the model Technical Report.[5] The model structure here is similar to Bagust and colleagues.[36] Retinopathy health states and allowable transitions are given in figure 6 of the Technical Report.[5] These take explicit account of left and right eye differences, again as in Bagust and colleagues.[36] The chronic kidney disease (CKD) module consists of kidney damage (measured by “None”, “Microalbuminuria”, “Macroalbuminuria”, and “ESRD” (end stage renal disease)) and kidney function as defined by the continuous covariate eGFR. This is consistent with the National Kidney Foundation K/DOQI stage definitions and with the diabetes sub-population of the Centers for Disease Control (CDC) Model of CKD, with some modifications.[5]

The model adopts a lifetime horizon of 40 years in the base case and has an annual cycle length. MS Table 31 (MS p174) notes that the comorbidities resulting from hyperglycaemia develop relatively slowly and that this justifies an annual cycle length. The ERG agrees with this assessment. A half-cycle correction was not used as the modelling duration is long relative to the cycle length (MS Table 31 p174).

The model assumes that the effects of treatment seen in clinical trials (typically at 26 and 52 weeks of follow-up) may not persist with continued treatment. This effect is implicitly subsumed in “annual drift” parameters.[5]

The model has parameters which concern the “Duration during which to apply the fading-off of the effect after withdrawal” [6] which apply to non-insulin treatments. The fade off duration is not

specified in the base case, indicating that any treatment rebound is immediate.[5] This applies to all treatment effects in the base case, including BMI. The ERG notes that the assumption of immediate rebound is beneficial to treatments which are less effective than their comparators, i.e. will improve their cost-effectiveness.

Modelling base case comparisons were carried out using 1000 patients in each cohort and 1000 cohorts (Table 53). It is unclear why 1000 cohorts were used in the base case as this means that all base case analyses include PSA. In the ERG's experience base case analyses do not usually reflect PSA. Given the complexity of the diabetes disease pathway it is also unclear if the use of 1000 patients is sufficient to robustly capture the variation in ICER. Variation to these model settings is examined by the ERG in Section 4.3.

ECHO-T2DM generates pairwise comparisons (MS p233). The MS argues that since each simulation run will generate slightly different cohorts of patients there can arise differences in the absolute magnitudes of the outcomes from run to run for identical scenarios and for this reason it is inappropriate to mix and match simulation results from different runs for fully incremental analysis. However such differences are minimised if a very large number of patients can be used and the ERG presents incremental results for cohorts of 100,000 patients (with no parameter uncertainty) in Section 4.3.

Table 53: Model settings (reproduced from MS Table 34 p184)

Variable	Value
Number of patient cohorts per model run: base case comparisons	1000
Number of patients in each cohort: base case comparisons	1000
Number of patient cohorts per model run: scenario analyses and sensitivity analyses	500
Number of patients in each cohort: scenario analyses and sensitivity analyses	1000
Discount Rate for Costs	3.5%
Discount Rate for Health	3.5%
Modelled time horizon	40
Maximum patient age	100

The Technical Report states that on a typical computer (Intel i7 8GB RAM) a single simulation of the model over a 25 year time horizon with 500 cohorts of 500 patients requires approximately four hours of computation time (p8). The ERG found that it takes approximately 15 hours to run a base case simulation of 1000 cohorts of 1000 patients over a 40 year time horizon on a high specification PC (Intel i7 16 GB RAM). It takes 1.5 hours to run a simulation with 100,000 patients and one cohort on the same high specification PC.

In summary, the modelling approach adopted in the submission appears reasonable although no justification was provided for the final choice of model from other models which also met the inclusion criteria of the systematic review. The model structure is comprehensive and draws on well-documented risk equations including UKPDS 68 to model disease progression. However no details of the programming architecture of the 30 component program and data files, i.e. how the separate programs relate to each other, and the processing flow, are given in the Technical Report or User Guide. This has hindered model checking by the ERG. The model takes an appreciable amount of time to run with base case assumptions (15 hours plus). It is unclear if the base case use of 1000 patients is sufficiently large to robustly estimate the ICERs, or how the base case results might change when parameter uncertainty is switched off.

4.2.2 Patient Group

The patient group included in the economic evaluation is adults with T2DM. The principal sources used to describe patients' initial characteristics comes from RCTs included in the SLR (data reported in MS page 67, MS Appendix 4 and MS Appendix 9), see also Section 3.1.3 above. The populations correspond to the licensed indication for canagliflozin for add-on therapy. There are some differences in the populations in terms of their duration of diabetes and baseline HbA_{1c} reflecting the differences between the treatment sequencing of the trials (dual therapy, triple therapy, add-on insulin) as outlined below. The populations generally reflect the target population in clinical practice. Inadequate control of T2DM is defined in current guidelines[7] as HbA_{1c} >7.5% and the mean HbA_{1c} of participants in the included RCTs was above this threshold at baseline (although in the dual therapy group there may have been some participants below this as the mean HbA_{1c} was 7.8 and SD was 0.843).

For dual therapy with metformin background two RCTs (DIA3009/DIA3006) were 'pooled' (method not described); the populations of these trials were adults with T2DM. The inclusion criteria of these trials differed slightly in terms of the baseline range of HbA_{1c} eligible (which depended on the type and doses of prior treatment) but the mean age of participants overall was 56 years, 50% were male, 74% white and the mean baseline HbA_{1c} was 7.9%. The duration of diabetes was 6.7 years. For triple therapy with metformin and SU background two RCTs were pooled (DIA3002/DIA3015). Participants were all adults with T2DM, inclusion criteria differed slightly but participants mean age was 57 years, 54% were male, and the mean baseline HbA_{1c} was approximately 8.1%. The proportions of white participants was 69% overall (Appendix 9, Table 183) although the rates differed between the two trials, with around 82% in

Superseded – see erratum

DIA3002 and 64% in DIA3015 (MS page 67). The duration of diabetes was 9.6 years. For triple therapy with metformin and TZD one RCT (DIA3012) was relevant. Mean age of the participants in this trial was 57 years, the proportion male was 63% and approximately two thirds were white (the MS reports 63% and Appendix 9 reports 79%). The mean baseline HbA_{1c} was approximately 8% and the duration of diabetes 10.6 years. For add-on to insulin, one trial (DIA3008 insulin sub-study) was used. Participants differed from those of the other studies, with a mean age of 63 years, a duration of diabetes of 16.6 years, and a mean baseline HbA_{1c} of 8.3%. 67% were men and 78% were white. These participants also had a high risk for CVD (MS page 67).

Sub-group analyses for patients with an HbA_{1c} exceeding 9% (75mmol/ml) at baseline are reported on MS page 279-81. Analyses were only conducted where there were direct comparative data from the canagliflozin trials (DPP-4-i and SU in dual therapy, DPP-4-i in triple therapy). Cost-effectiveness analyses by sub-groups of patients with an HbA_{1c} <8% or 8% to <9% were not reported. However, clinical effectiveness results for these sub-groups were given in the MS, and the results show that there were larger absolute reductions in HbA_{1c} in patients with baseline HbA_{1c} ≥ 9% (see section 3.3.7 of this report). No justification is provided for the HbA_{1c} sub-group though it was a pre-specified sub-group in the canagliflozin clinical trials. The MS reports that subgroup analyses on baseline BMI undertaken in the SLR (MS p107-8) show the response to canagliflozin was consistent and there was a lack of differential effectiveness for canagliflozin by BMI. The cost effectiveness of treatment for the subgroup of BMI (defined in the NICE scope) is therefore not analysed.

4.2.3 Interventions and comparators

The intervention is canagliflozin 100mg or canagliflozin 300mg. The recommended starting dose of canagliflozin is 100mg once daily. In patients tolerating canagliflozin 100mg once daily who have an eGFR ≥ 60mL/min/1.73 m^2 or CrCl ≥ 60mL/min and need tighter glycaemic control, the dose can be increased to 300mg once daily.

A range of comparators were used, all of which are relevant to current UK practice. Canagliflozin is compared to the treatment classes SU, TZD, DPP-4-i, GLP-1-a and insulin, as well as dapagliflozin, the only other currently available SGLT-2 inhibitor. The specific drugs selected for comparison within each of these classes are shown in Table 54. Treatments within each class are not wholly equivalent. Specifically, the ERG clinical expert notes that within the

SU class gliclazide is associated with a lower rate of hypoglycaemia than glimepiride and that within the GLP-1-a class liraglutide is more effective than exenatide BID (and more cost-effective [10]).

Table 54: Specific comparators selected within each pharmacological class (extract of MS Table 33 p176)

	Pharmacological class member selected	Justification
SU	Efficacy: glimepiride Cost: gliclazide	Glimepiride was the SU used in the trial programme Gliclazide is the SU with the highest level of UK prescribing
TZD	Pioglitazone	Rosiglitazone has no licence/ no other TZD available for prescribing in clinical practice in the UK
DPP-4-i	Sitagliptin	Most used UK DPP-4-i and comparator in two canagliflozin trials
GLP-1-a	a) Liraglutide	Market share data shows liraglutide has a wider use in UK clinical practice, and thus selected as the GLP-1-a of choice where data allow
	b) Exenatide BID	Exenatide BID has a broader recommendation for use in T2DM; Exenatide BID is the only GLP-1 present in the Triple (MET+SU) and add on to insulin NMA networks.
Insulin	insulin aspart	From market share data ASPART has a widest use for SA INS, in UK clinical practice
Legend: DPP-4-i: Dipeptidyl peptidase-4 inhibitor; GLP-1-a: Glucagon-like peptide-1 agonist; SU: sulfonylurea; TZD: Thiazolidinediones		

The manufacturer's clarification letter to NICE and the ERG also states that there is some evidence that the rate of hypoglycaemic events is lower for gliclazide than glimepiride.[37] This was examined by the manufacturer in sensitivity analysis by varying the hypoglycaemia rate by +/-20% but only had a small influence on results. The ERG examines greater uncertainty in the hypoglycaemia rate for gliclazide, and improved efficacy for liraglutide, in additional analysis described in Section 4.3.

Some comparators from the scope were excluded from the decision problem. These were:

- Dual therapy with a SU (comparators in the scope were TZD (with SU); DPP-4-i (with SU); GLP-1-a (with SU)). The MS justifies this on the basis that the rate of hypoglycaemic episodes is expected to be high and that they do not expect this combination to be used in clinical practice. Also, this is excluded because a proportion of patients would not be eligible to have SU in combination with canagliflozin because of the risks in those with an eGFR <45ml/min. The ERG clinical advisor suggests that the risk of renal impairment does not fully justify excluding this therapy combination because a reasonable proportion of patients on

SU do not have renal impairment (see section 2.3 and 4.4 of this report for a discussion of this).

- In triple therapy in comparison to insulin (alone or in combination with one or more oral anti-diabetic agents, such as a DPP-4-1 or GLP-1-a).

[redacted]

A summary of the economic modelling comparisons conducted is shown in Table 55.

Table 55: Economic modelling comparisons conducted (reproduced from MS Table 32 p175).

Comparator	Cana dose	Dual therapy	Triple therapy		Add-on to insulin
		MET	MET+SU	MET+TZD	
SU	100mg	✓	*n/a*		
	300mg	✓	*n/a*		
TZD	100mg	✓		*n/a*	
	300mg	✓		*n/a*	
DPP-4-i	100mg	✓	✓	✓	✓
	300mg	✓	✓	✓	✓
GLP-1-a	100mg	✓	✓		✓
	300mg	✓	✓		✓
Dapagliflozin	100mg	✓			✓
	300mg	✓			✓
Insulin	100mg		✓		
	300mg		✓		

Legend: DPP-4-i: Dipeptidyl peptidase-4 inhibitor; GLP-1-a: Glucagon-like peptide-1 agonist ; n/a: not applicable; SU: sulfonylurea; TZD: Thiazolidinediones

Superseded – see erratum

The comparators are routinely used in UK NHS, although clinical advice to the ERG suggests that pioglitazone (a TZD) is used infrequently in clinical practice now. Two of the specific comparators used in the economic model (the SU glimepiride and DPP-4-I exenatide BID) are not the most widely prescribed in their class in the UK, and are not wholly equivalent clinically to other drugs in their class.

4.2.4 Clinical Effectiveness

The primary outcome from the efficacy trials used by the model is the first year change from baseline in HbA_{1c} (%) (MS p67). Secondary outcomes used by the model are first year BMI

change from baseline (kg/m^2); SBP change from baseline (mmHg); change in total cholesterol (mg/dL); change in LDL cholesterol (mg/dL); change in HDL cholesterol (mg/dL); change in triglycerides (mg/dL); annual rate of mild, moderate and severe hypoglycaemic events; and rates of UTI and GMI.

The clinical effectiveness model parameter values are given in Appendix 10.9 Tables 185-188. MS Table 185 is reproduced as Table **56** below for dual therapy. The sources of these parameter values are identified in the footnotes to Tables 185-188 as the NMA and the various clinical trials, although some parameter values are assumed.

The model requires treatment effectiveness versus placebo to be specified, rather than treatment effectiveness relative to the comparator. Where NMA data were not available these data were taken from clinical trials which involved both intervention and comparator, and consequently one treatment can be associated with various efficacy estimates for a single outcome, depending on which treatment it is being compared to (Table 56 and Appendix 10.9 Tables 185-188). This applies to outcomes which are not informed by the NMA, for example first year change in total cholesterol, and rate of GMI.

Table 56: Treatment effects, Rates of AEs and 1st year discontinuation due to AEs: Dual therapy, Base cases; Canagliflozin, TZD, GLP-1-a, dapagliflozin (extract of MS Table 185 Appendix 10.9)

Parameter	Dual therapy +MET															
	CANA 100mg		Pioglitazone		Liraglutide		Dapagliflozin		CANA 300mg		Pioglitazone		Liraglutide		Dapagliflozin	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Treatment effects –LS mean change from baseline																
HbA1c (%)[a]	-0.63	0.26	-0.77	0.3	-1.02	0.35	-0.48	0.23	-0.76	0.23	-0.77	0.3	-1.02	0.35	-0.48	0.23
SBP (mmHg)	-3.72[b]	0.38[b]	-5.77[a]	2.69[a]	-0.22[a]	1.55[a]	-3.72[d]	0.38[d]	-4.94[b]	0.39[b]	-5.77[a]	2.6[a]	-0.23[a]	1.55[a]	-3.72‡	0.38‡
BMI (kg/m2)	-0.56[a]	1.47[a]	0.96[a]	2.41[a]	-0.74[e]	0.24[e]	-0.87[e]	0.19[e]	-0.69[a]	1.47[a]	0.96[a]	2.41[a]	-0.74[e]	0.24[e]	-0.87[e]	0.19[e]
Total Cholesterol (mg/dl)	5.54[b]	1.09[b]	0[f]	0[f]	0[f]	0[f]	5.54[d]	1.09[d]	10.44[b]	1.10[b]	0[f]	0[f]	0[f]	0[f]	5.54[d]	1.09[d]
LDL Cholesterol (mg/dl)	5.07[b]	0.95[b]	0[f]	0[f]	0[f]	0[f]	5.07[d]	0.95[d]	8.03[b]	0.96[b]	0[f]	0[f]	0[f]	0[f]	5.07[d]	0.95[d]
HDL Cholesterol (mg/dl)	3.57[b]	0.26[b]	0[f]	0[f]	0[f]	0[f]	3.57[d]	0.26[d]	4.35[b]	0.27[b]	0[f]	0[f]	0[f]	0[f]	3.57[d]	0.26[d]
Triglycerides (mg/dl)	-16.05[b]	3.09[b]	0[f]	0[f]	0[f]	0[f]	-16.05[d]	3.09[d]	-11.88[b]	3.13[b]	0[f]	0[f]	0[f]	0[f]	-16.05[d]	3.09[d]
Adverse events -event/pt-years																
Female genital mycotic infection	0.13[b]	-	0.01[c]	-	0.01[c]	-	0.13[d]	-	0.14[b]	-	0.013[c]	-	0.01[c]	-	0.13[d]	-
Male genital mycotic infection	0.07[b]	-	0.01[c]	-	0.01[c]	-	0.07[d]	-	0.07[b]	-	0.012[c]	-	0.01[c]	-	0.07[d]	-
Upper urinary tract infection	0.003[b]	-	0[c]	-	0[c]	-	0.003[d]	-	0.001	-	0[c]	-	0[c]	-	0.003	-
Lower urinary tract infection	0.09[b]	-	0.08[c]	-	0.08[c]	-	0.09[d]	-	0.08[b]	-	0.08[c]	-	0.08	-	0.09[d]	-
Severe hypoglycaemia	0.01[b]	-	0[c]	-	0[c]	-	0[g]	-	0.001[b]	-	0[c]	-	0[c]	-	0[g]	-
Non-severe symptomatic hypoglycaemia	0.07[b]	-	0.02[c]	-	0.02[c]	-	0.02[g]	-	0.05[b]	-	0.02[c]	-	0.02[c]	-	0.02[g]	-
Nausea	0.04[b]	-	NA	-	0.05[f]	-	NA	-	0.05[b]		NA	-	0.05[f]	-	NA	-
Proportion discontinuing due to AEs	0.05[b]	-	0.04[c]	-	0.04[c]	-	0.05[d]	-	0.05[b]	-	0.04[c]	-	0.04[c]	-	0.05[d]	-

a. Data from NMA[38]; b. Pooled DIA3006&DIA3009; c. Placebo data from DIA3006 (to week 26); d. Cana 100mg data from pooled DIA3006&DIA3009; e. Weight change from NMA, converted to BMI using UK gender-specific standardized heights weighted by the distribution of gender from pooled DIA3006 & DIA3009; f. Assumption (see 7.3.1); g. Odds ratio from NMA used to estimate event rates

The model estimates of the first-year change in HbA_{1c} % compared to placebo are derived from the NMA reported in MS Section 6.7 (MS p110), discussed in Section 3.1.7 of this report. NMA results are available for dual therapy at 26 and 52 weeks and for triple therapy at 26 weeks. Results observed at 26 weeks in the triple therapy and add-on to insulin trials are assumed to be equivalent to the effects that would have been observed at 52 weeks in these trials (MS p187). Given that the trial data for 52 week change in weight and HbA_{1c} do appear similar to 26 week data the ERG clinical expert considers that this is a satisfactory assumption.

The NICE Methods Guide states that data from head-to-head RCTs should be presented in the reference-case analysis, if available.[39] However the model sometimes uses input data from the NMA in preference to data from head-to-head trials against placebo, even when there is only one trial informing a comparison in the network of evidence. The MS states that this is for methodological consistency (MS p161). The ERG considers that this may not be an overall conservative approach as in some cases the direct data indicate than canagliflozin is less effective than the NMA data. For example the 26 week NMA data for canagliflozin 100mg in triple therapy with MET+SU, for change in HbA_{1c} versus placebo, show that canagliflozin is more effective against placebo than the direct trial data (DIA3002): a first year change in HbA_{1c} of -0.75 is given by the NMA (MS Appendix Table 187), but a change of -0.71 was found in the trial (MS reference 28, poster presentation).

The ERG observes that given that the MS only reports pair-wise cost-effectiveness analyses, it would have been more appropriate to use direct trial data to inform the pairwise comparisons rather than NMA data. Although the MS argues that the NMA data provide methodological consistency, the ERG considers that this was perhaps unnecessary given that a fully incremental cost-effectiveness analysis was not undertaken.

The model uses data from the 52-week NMA to inform the effectiveness of canagliflozin in reducing HbA_{1c} % compared to placebo in dual therapy with MET. This comparison is not available using the head-to-head data (MS Figure 35 p116). However head-to-head comparisons against placebo do exist at the 26 week time point (MS Figure 34 p116). It is not clear if these direct comparison data have been examined in scenario analysis. MS Table 44 (MS p211) indicates that treatment effect scenarios were run using direct comparative data from

RCTs and using the output of the 26 week NMA (scenarios 16 and 14 respectively), but the time point of the RCT data is not given.

The model Technical Report (p35) states that because treatment effect data for the first year change in HbA_{1c} (and other biomarkers) may implicitly have annual upward drift already included, it can sometimes be useful to subtract the annual drift for cycle 1 from the treatment effect in order to avoid undercounting the effects of treatment.[5] Consequently the estimates derived from the NMA are adjusted downwards in the model input data spreadsheet by an amount equivalent to the annual drift, and do not correspond exactly to the outputs of the NMA given in Appendix 10.7 (Manufacturer's clarification response letter). The drift assumptions for biomarkers for all lines of therapy are given in Table 57, by treatment. The effect of this calculation is that the first year annual drift of -0.14 for canagliflozin (Table 57) is added to the first year change in HbA_{1c} % reported by the NMA of -0.63 (Table 56) to give a first year change in HbA_{1c} of -0.77% as input to the model.

The ERG does not consider that these calculations are transparent as they are not documented in the MS or input data spreadsheet. A further issue is that the standard error of the drift assumption is not reflected in the standard error for the combined first year change in HbA_{1c} and drift estimate, which as specified in the model is simply the standard error associated with the NMA estimate of the first year change in the HbA_{1c} %.

The annual drift is also added to treatment effect estimates obtained from the NMA at 26 weeks. Although this is consistent with the MS assumption that results observed at 26 weeks are equivalent to the effects that would have been observed at 52 weeks, the ERG does not consider that it is reasonable to add back the entire annual drift to these estimates. However the impact on results is likely to be negligible as the model is not very sensitive to first year change in biomarker values (MS Section 7.7.7).

Values for annual drift are treatment specific and taken from the ADOPT study[40], UKPDS 68,[35] the dapagliflozin NICE TA288[8] and other sources, depending upon the biomarker (Table 57). It is reported elsewhere in the MS (MS p82) that the upward drift in HbA_{1c} % from weeks 26 to 104 in the DIA3009 trial was 0.16% for both canagliflozin 100mg and 300mg. Consequently the ERG considers that the assumed annual drift of 0.14% for canagliflozin for HbA_{1c} % appears

reasonable. The ERG clinical expert considers that the drift assumptions given in Table 57 are based on robust studies and are appropriate.

The drift assumptions for HbA_{1c} % seen in Table 57 have been checked with the source by the ERG. In the ADOPT study[40] 95% CIs were provided and these have been added by the ERG to Table 57 for completeness. Given that the manufacturer did not report these CIs, it is not clear to what extent uncertainty in annual drift in HbA_{1c} % has been considered in PSA (see also Section 4.2.9).

Table 57: Drift assumptions for biomarkers for all lines of therapy (reproduced from MS Table 36 p186, with CIs added)

	AHA	Annual Drift Value	Source
HbA_{1c}	Canagliflozin, Dapagliflozin, DPP-4-i, GLP-1-a	0.14% (95% CI 0.13, 0.16)	Assumed equal to value observed for the metformin arm in ADOPT study[40] [41]
	TZD	0.07% (95% CI 0.06, 0.09)	Assumed equal to value observed for the rosiglitazone arm of ADOPT study[40]
	SU	0.24% (95% CI 0.23, 0.26)	Assumed equal to value observed for the SU arm of ADOPT study[40]
	Insulin	0.15%	
SBP	All AHAs	0.30 mmHg	Derived from UKPDS 68[35]
Lipids	All AHAs	0.03 mg/dl	Derived from UKPDS 68[35]
Weight*	All AHAs	0.1kg/year	Value from Dapagliflozin NICE FAD
eGFR	All AHAs	Varies by eGFR category and whether patient micro- or macro-albuminuria	Technical Report

The MS notes that the non-linear HbA_{1c} evolution equation (Equation 11 from UKPDS 68) has been used by some analysts to model upward drift in HbA_{1c}, rather than data from the ADOPT study (MS p180). The MS states that this specification is confounded in ECHO-T2DM by the intermittent addition of rescue medication in the UKPDS and resultant HbA_{1c} -lowerings, and that because ECHO-T2DM applies the consequences of rescue medication explicitly, using the evolution equation would amount to double-counting of the benefits of rescue medication.

Apart from HbA1c % and annual drift in HbA_{1c} %, other input parameters affected by the intervention in the model are change in SBP; change in total cholesterol; change in LDL cholesterol; change in HDL cholesterol; change in triglycerides; change in BMI; annual rate of

hypoglycaemic events; hazard ratio for GMI and UTI; probability of discontinuing treatment given certain complications; and the multiplier on treatment effects given eGFR values. These parameters are considered in MS Sections 6.5 (MS p78), 6.7 (MS p110), and 6.9 (MS p140).

Some of the secondary outcome parameters (change in SBP; change in total cholesterol; change in triglycerides; change in BMI) enter the model as differences from baseline values and are drawn either from NMA (BMI change) or the key trials (SBP and blood lipid change).

The MS notes that where BMI was not available in the NMAs, but data on weight were, values were converted to BMI using a weighted average of male and female heights observed in the UK general population (MS p179). The averages used were 1.75m for males and 1.613m for females (manufacturer's clarification letter to NICE and the ERG). The manufacturer uses the proportion of males and females from the pooled trial data together with the UK height averages to estimate the change in BMI in these cases. The ERG considers that this assumption could lead to an over- or under-statement of BMI change depending on the heights of the individuals enrolled in a trial, and that the pooling of the trial data to calculate the proportion of males and females is an additional simplification. In their clarification letter the manufacturer presents revised base cases which extend the use these height assumptions to more treatments, but the ERG does not consider that the uncertainty associated with using alternative methods of calculating BMI change with more specific (unpooled) trial data have been fully explored.

Hypoglycaemic events are modelled using observed rates which are taken from the key trials (for canagliflozin), but with assumptions for some of the other comparators. For example, notes on the input data spreadsheet for the gliclazide hypoglycaemic event rate state "assume same as Glimepiride…adjusting GLIM so it's relative to CANA100 from 3009 52wks". The adjustment which is used here is not clear.

HRs are used to model the incidence of GMIs and UTIs with data drawn from the relevant canagliflozin trials for the first year of treatment and based upon clinical assumption thereafter. Absolute probabilities are used to model treatment discontinuation in the case of adverse events or particular health states, again with data drawn from the key trials or based upon clinical assumption. Multipliers are used to modify treatments effects given eGFR status which either use trial data for canagliflozin, or are assumed to be one when there are no data available.

The discontinuation rates due to adverse events for all treatments come from the included RCTs of canagliflozin. These are generally higher than the rates used for treatments in NICE TA288[8] for dapagliflozin which were taken from an NMA and identified as being potentially too low, as they were lower than rates from the associated dapagliflozin clinical trials. Given that the discontinuation rates for canagliflozin and its comparators are drawn from clinical trials the ERG considers that they are reasonable.

Change in HbA_{1c} % and the other biomarkers reflect intermediate outcomes. Disease progression and mortality are defined in the model by sets of macro- and microvascular risk equations which take the biomarker values (HbA_{1c}, SBP etc) as risk factors. The equations are used to calculate the probability that a patient will experience a particular complication in a given cycle. The choice of microvascular risk equations is made by entering an appropriate profile name for the microvascular complication parameter set in the input data spreadsheet. The parameter values associated with this profile are specified elsewhere in the input data spreadsheet. The choice of macrovascular risk equations is made by entering an appropriate profile name for the "CVD Risk Equation" profile in the input data spreadsheet. The associated parameter values for the base case profile (UKPDS 68) are specified in the model R programs rather than in the input data spreadsheet.

Macrovascular risk equations from UKPDS 68 are used in the model base case. UKPDS 68 equation 1 is used to predict the risk of IHD; equation 2 is used to predict the risk of MI; equation 3 used to predict risk of CHF; and equation 4 is used to predict the risk of stroke.[35] It is possible to select from UKPDS 82, ADVANCE or Swedish NDR equations instead of UKPDS 68 equations if desired.

The annual probabilities of worsening microvascular complications are derived primarily from the Wisconsin Epidemiologic Study of Diabetic Retinopathy, The Rochester Epidemiology project and the CDC model of CKD. The model uses the power function approach outlined in Eastman and colleagues [42] to link the risk of developing microvascular complications to individual patient HbA_{1c} values (during each cycle). The link between microvascular risk and SBP uses data from Brown and colleagues.[43] The ERG clinical expert considers that these are quite old sources of data and that it is likely that progression to retinopathy and nephropathy is now longer because of better treatment and earlier diagnosis. However the ERG clinical expert is not aware of better data.

Risk of mortality is based on equations 8, 9 and 10 of the UKPDS 68 Outcomes model in the base case, although equations from UKPDS 82 may be used as an alternative.

No search for evidence for the macrovascular and microvascular risk equations is described in either the Technical Report or MS. However the use of the UKPDS 68 equations is common to many other models of T2DM (MS p171) and the ERG clinical expert considers that the base case use of the UKPDS 68 data is reasonable.

Due to time constraints it has not been possible for the ERG to cross-check all of the clinical effectiveness values used in the model data input spreadsheet against their clinical and NMA sources. However there is a discrepancy in the input value used for change in LDL cholesterol for canagliflozin 100mg at 26 weeks in the DIA3002 trial. The trial poster reports a change of -0.6 mg/dL (MS reference 28), whilst the input data spreadsheet uses a value of -0.8 mg/dL attributed to the same source, for the treatment profile 'NMA TRIPLE 26WKS (MET+SU BACKGROUND) CANA100 vs. SITA'. Use of a value of -0.8 instead of -0.6 will lead to an improvement in the cost-effectiveness of canagliflozin 100mg in this line of therapy.

Overall the ERG considers that model clinical effectiveness data are appropriate although sometimes NMA estimates are used in preference to head-to-head trial data and that this is not well justified given that a fully incremental cost-effectiveness analysis was not reported in the MS. In some cases input calculations are not fully reported. Equations used for the extrapolation of biomarker outcomes are well-established and appropriate to the UK population.

4.2.5 Patient outcomes

The cost-effectiveness model incorporated the impact of the different treatments on HRQoL into QALYs. HRQoL enters the model with each patient having a baseline utility. As participants progress through the model, disutility values are attributed according to specific characteristics. These include patient characteristics (age, sex, duration of diabetes), macro-vascular complications (IHD, MI, CHF, Stroke); micro-vascular complications (retinopathy, blindness, gross proteinuria, end-stage renal disease, symptomatic neuropathy, PVD, symptomatic neuropathy and PVD, diabetic foot ulcer, one lower extremity amputation, two lower extremity amputation); hypoglycaemic episodes (minor, moderate, severe); obesity; and adverse events (lower UTI male/female, upper UTI male/female; balanoposthitis, candida vulvovaginitis), see

Table MS 38, p198-9. The model appears to be consistent with the model used in the dapagliflozin STA (NICE TA288).[8]

Utility data were taken from the CODE-2 EQ-5D modelling study in most cases.[31] For hypoglycaemic episodes utilities were taken from a published source (Currie et al 2006[44]) and these appear to correspond to estimates applied in NICE TA288[8] of dapagliflozin. Adverse events utilities were taken from a time trade off (TTO) study commissioned by Janssen[45] as no data were identified in the literature. The CODE-2 EQ-5D[31] study was identified in a systematic review although the MS does not directly make this link between the systematic review and the study. The second main source of utility data[44] was also identified in the systematic review of HRQoL.

For the manufacturer's systematic review of HRQoL a systematic search for evidence was undertaken on 20 databases, inclusion criteria were defined, and the review identified 106 possible utility studies (a PRISMA flow chart with full details is on MS page 193). Full details of the systematic review are provided in a subsequent report (MS reference 180). The systematic review provides summary data of utilities for a range of patient characteristics and complications, however, there is no justification provided in the MS as to why these were not used in the model and why the CODE-2 study was chosen for the source of the majority of utilities. In response to a clarification question from the ERG about the choice of the CODE-2 study the manufacturer states that the CODE-2 study was chosen because it used well-validated techniques to solicit HRQoL, patients were relevant to a UK setting, prevalence data were captured for a large number of characteristics and the study had a large sample size. The manufacturer also states that the UKPDS QoL were rejected because they included a smaller sub-set of the health states and did not include a disutility for the important BMI covariate (whereas CODE-2 does include such a coefficient). The ERG believes this is reasonable; however, no discussion is made of whether any alternative utilities from the other included studies could have informed sensitivity analysis.

As stated, the MS draws EQ-5D disutility values from a modelling study based on CODE-2 data[31] for most of the disutilities applied in the model. The MS states that previous diabetes related submissions to NICE have used the UKPDS trial data, but that this does not have the level of detail required for the present model in terms of micro-vascular complications. The CODE-2 modelling study is reported to be a better match to the micro-vascular health states as

it provides a wider range of estimates for different severity levels. The MS also justifies the choice of the CODE-2 study because there is an estimate of disutility related to BMI which reduces the risk of confounding (BMI was found to be a significant predictor of utility values at baseline in the canagliflozin clinical trials for dual and triple therapy groups).

Although the original CODE-2 study is not an RCT (it is a non-intervention cohort study) the MS states that the populations are more broadly representative of T2DM than an RCT population (MS p197).

Other sources were used to determine disutility weights associated with AEs (MS page 196). The MS states that where possible utilities were measured using the EQ-5D questionnaire and taken from a UK population with T2DM. It is assumed this relates to the CODE-2 study which was a European population based study, but this is not explicitly stated.

The baseline utilities used in the MS model are reproduced in Table 58 below. These have been checked with the source by the ERG and while in most cases these appear to be correct, it is observed that:

- the model uses the same disutility for two lower extremity amputations as is used for one lower extremity amputation (the ERG have been unable to source any other suitable data however);
- as stated in the MS (see Table 58), the disutilities applied for macro-vascular complications of IHD, MI and CHF were based on a disutility for CHD in the CODE-2 study using an assumption that the same disutility can be applied across these different states. This applies consistency with the other disutilities used, however, the ERG note that in the UKPDS study (used in previous diabetes NICE TAs) these macro-vascular complications all have different utility estimates (the utilities are also higher in the UKPDS, IHD -0.141, MI-0.081, CHF -0.058);
- the health states for adverse events in the MS (see Table 58) do not match the descriptions of the health states used in the TTO study,[45] however, the ERG have assumed that upper UTI=severe UTI and lower UTI=mild to moderate UTI. Disutilities given for adverse events in the TTO report have been adjusted to reflect the likely duration of the infections in days (MS p197). While it is not clear how long the MS assumes each type of infection will last this can be calculated and the ERG estimate that it is assumed that a mycotic infection lasts for 1 week, upper UTI lasts for 2 weeks and a lower UTI lasts for 5 days.

Table 58: Baseline utilities and sources used in the MS model (MS Table 38)

State	Utility	SE	Reference	Justification
Baseline	1.027	0.027	CODE 2	Values were estimated using multivariate regression techniques
Patient Characteristics				
Age (per 10 Years)	-0.0235	0.001	CODE 2	HRQL has been shown to gradually decrease over time and may change with age independent of the disease progression. MS ref 183
Female	-0.093	0.009	CODE 2	Gender has been show to impact patient reports of HRQL, often with women reporting lower scores (MS refs 130, 184)
Duration of T2DM (per 10 Years)	-0.0163	0.001	CODE 2	This captures any effects of having T2DM over time, above and beyond the separate effects of aging, treatment intensification, and the effects of the development and worsening of complications.MS ref 185
Macro-vascular Complications				
IHD	-0.028	0.01	Assumed from CHD in CODE 2	Identified in a systematic literature review, CODE-2 was deemed the most appropriate study on HRQoL related to T2DM complications, controlling for many potential confounding factors simultaneously. CODE-2 combines macro-vascular diseases together.
MI	-0.028	0.01	Assumed from CHD in CODE 2	
CHF	-0.028	0.01	Assumed from CHD in CODE 2	
Stroke	-0.115	0.017	CODE 2	
Micro-vascular Complications				
Retinopathy (BDR, ME, PDR, & combinations)	0	1E-99	CODE 2	CODE-2 controls for many potential confounding factors simultaneously.
Blindness (one or both eyes, incl.combinations)	-0.057	0.022	CODE 2	
Gross Proteinuria	-0.048	0.022	CODE 2	
ESRD	-0.175	0.028	CODE 2	
Symptomatic Neuropathy	-0.084	0.014	CODE 2	
Peripheral Vascular Disease (PVD)	-0.061	0.015	CODE 2	
Symptomatic Neuropathy & PVD	-0.085	0.018	CODE 2	
Diabetic Foot Ulcer	-0.17	0.019	CODE 2	
One lower extremity amputation	-0.272	0.029	CODE 2	
Two lower extremity amputations	-0.272	0.029	CODE 2	
Hypoglycaemic Events				
Minor Hypoglycaemic Episode (per Episode)	0		Assumption	Advised in advisory board- MS ref 163
Moderate Hypoglycaemic Episode	-0.0036		Currie et. al 2006. annualised	The approach taken by NICE in the evaluation of exenatide (CG-

State	Utility	SE	Reference	Justification
(per Episode)			disutility per symptomatic no help event; not chronic	87) also used utility estimates from Currie et al. 2006.MS ref 60; 128. Curie et al (2006) relates to a study that collected data from UK patients using both the hypoglycaemia fear survey and EQ-5D. The utility values were converted to annual values as follows: -0.047 x (3/12) year = -0.0118 per year.
Severe Hypoglycaemic Episode (per Episode)	-0.0118		Currie et. al 2006. annualized disutility per severe event; not chronic	
Obesity				
Excessive weight (disutility weight for each 1 kg/m^2 above a threshold of 25 kg/m^2. No disutility is applied to lower values)	-0.0061	0.001	CODE 2	Values used to inform utility values in the NICE CG-87
Adverse Events				
Lower urinary tract infections (male) (per Episode)	-0.0012		Janssen UK Study (regression analysis corrected for state duration)	No utility values related to urinary tract infections were identified in the systematic literature review. Therefore, a TTO utility study was conducted.(MS ref 164; 165)
Lower urinary tract infections (female) (per Episode)	-0.0012			
Upper urinary tract infections (male) (per Episode)	-0.0073			
Upper urinary tract infections (female) (per Episode)	-0.0073			
Balanoposthitis (male) (per Episode)	-0.0046			No utility values related to genital mycotic infections were identified in the systematic literature review. Therefore, a TTO utility study was conducted. (MS ref 164; 165)
Candidal vulvovaginitis (female) (per Episode)	-0.0046			

Superseded – see erratum

No HRQoL data for the canagliflozin clinical trials were reported in the SLR although MS Appendix 14 reports limited data on EQ-5D from three RCTs. These were for dual therapy (DIA3009) where the baseline utility was 0.798 (0.2049) and the week 52 utility was 0.814 (0.2140); and for triple therapy (DIA3002 + DIA3015 pooled) where the baseline utility was 0.788 (0.1929) and the week 52 utility was 0.797 (0.1957). These data appear to be for all patients, rather than for just canagliflozin, although this is not clearly stated.

The health benefits were measured and valued consistent with NICE's reference case and the utilities incorporated into the model appear to be appropriate. While the EQ-5D was used, no details of the value set adopted was reported in the MS, although as noted above the MS states

that where possible utilities were measured using the EQ-5D questionnaire and taken from a UK population with T2DM.

4.2.6 Resource use

The cost analysis was conducted from an NHS and PSS perspective and overall the assumptions appear to be reasonable. The basis for estimating dosage and frequency of administration of drug acquisition resource was not stated explicitly as resource estimates but were taken from the SmPC (MS Table 39 p206). Similarly, assumptions over dosing and frequency of treatment for the comparators was not explicitly stated as resource estimates, however, the MS refers to using various national guidelines, a systematic review of resources and costs, and the BNF for the dosing and frequency of treatment with comparators. These assumptions appear to be consistent with the canagliflozin RCTs in the manufacturer's SLR where these can be checked (the doses and frequency appear to be consistent, MS Table 6, Appendix 3) and with current clinical practice. In the case of insulin the resource use was based on a cost per body weight per patient in the model and changed as weight changes. These resources do not appear to have been tested in sensitivity analysis.

No drug administration resources or costs were applied in the model. Monitoring resources associated with checking renal function in canagliflozin were not included specifically as the MS (p206) states that these will also be covered in standard monitoring. The ERG clinical advisor agrees that this would be the case.

Little information on the estimates of resource use associated with disease progression was provided in the MS. A systematic review of resource and cost was undertaken and three studies were identified which provide costs used in the economic evaluation (although these did not come from the systematic review (MS p204 & Appendix 10 Table 191)). A Janssen advisory board recommended the approach to take for the resource implications for GMIs and UTIs. The advisory board stated that female patients will most commonly self-treat non-recurring GMIs with over the counter (OTC) medicines; whereas a male patient who experienced a balanitis will most likely require a GP appointment and prescription. Mild to moderate UTIs were thought to require a GP visit for both sexes and a prescription of antibiotics, such as trimethoprim, would be given (MS p205). In addition, the Janssen advisory board recommended that no resource was required for self-monitoring glucose testing strips, or any costs associated with mild adverse events, were required.

Few details of the Janssen advisory board were provided (e.g the methods, number of experts, how estimates were pooled, how uncertainty was addressed).

The MS also states that current guidelines were used but is not explicit about the details. Sources of information for resources associated with the adverse event health states were specified in MS Table 40 (see below for more details).

Overall, the estimates in the model appear to cover all relevant resource use. One possible exception is the GP visit for woman with a GMI where the ERG clinical expert considers it likely that a woman with GMI would visit a GP. In their SLR the manufacturer states that approximately 65% of the GMI events in women were treated by a health care professional and 35% of the events by the patient herself (MS p. 147). Despite the Janssen advisory board appearing to recommend no GP visit, the ERG on checking the costs of complications in the model suggest that a GP visit for females with GMI is costed.

4.2.7 Costs

A systematic review is reported in the MS, which appeared to meet required methods of search, selection of data and quality assessment and details of 10 studies were summarised in a separate report[19] and MS Appendix 10 Table 191. None of the findings of the review were deemed relevant by the manufacturer for use in the model.

Drug acquisition costs

The acquisition cost of canagliflozin 100mg and 300mg is currently indicative and is marked CIC by Janssen in the MS. The anticipated annual drug acquisition costs for canagliflozin 100mg is £477.26 (30-tab pack at £39.20) per patient. For canagliflozin 300mg the anticipated cost is £608.63 (30-tab pack at £49.99) per patient, MS Table 39, p206 and ERG Table 59.

The costs for comparators were taken from the BNF (July 2013) (MS Table 39 p206). These can be seen in Table 59. The data have been checked by the ERG, cross checking with the source of costs (BNF) and the calculations. While in most cases these data appear to be correct, it is observed that:

- For glimepiride the MS reports that the daily dose is 6mg. In the BNF the usual maximum dose is reported to be 4mg daily, where 6mg is stated to be used exceptionally.
- Pioglitazone is assumed to have a daily dose of 30mg. In the BNF the dose is recommended to be 15–30 mg once daily initially, increased to 45 mg once daily according to response. However, this is a conservative assumption by the manufacturer.
- There is a typo for the cost of gliclazide (reports cost of £1.16 which should be £1.61), however, the annual cost appears to be correct.
- The costs of insulin treatments is unclear from the information provided. For NPH (Neutral Protamine Hagedorn) insulin there is no reference to NPH insulin in the model input data spreadsheet and so it is unclear what this row refers to in Table 59. For insulin glargine the table indicates a dose range of 10-60 IU/day and for insulin aspart a dose range of 5-200 IU/day. The annual cost is reported as the cost per IU per year which has been worked out on the basis of a cost per IU multiplied by 365. For insulin aspart the cost does not appear to relate to the NovoRapid vial as stated but to the cost of Lispro 5 × 3-mL cartridge (for Autopen® Classic or HumaPen®) at £28.31.

Table 59: Drug acquisition costs applied in the model for all lines of therapy (MS Table 39).

Therapy	**Daily dose (mg)**	**Annual cost/pt**	**Description**
Anti-hyperglycaemic agents			
Canagliflozin	100	£477.26	100mg, 30-tab pack at £39.20
Canagliflozin	300	£608.63	300mg, 30-tab pack at £49.99
Glimepiride	6	£30.44	2mg, 30-tab pack at £1.24; 4mg, 30-tab pack at £1.39
Gliclazide	160	£19.60	80mg, 60-tab pack at £1.16
Pioglitazone	30	£36.13	30mg, 28 tab-pack at £2.77
Sitagliptin	100	£433.86	100mg, 28-tab pack at £33.26
Liraglutide	1.2	£955.49	6mg/mL, 2 x 3-mL prefilled pens at £78.48
Exenatide	0.2	£830.82	10µg/dose prefilled pen (60 doses) at £68.24
Dapagliflozin	10mg	£477.30	10mg, 28-tab pack at £36.59
Insulin therapy	**Daily dose**	**Cost/IU/year**	**Description**
NPH insulin	See insulin glargine and insulin aspart as applicable	£4.26	100 units/mL, 5 × 3-mL cartridge (ClikSTAR® and Autopen® 24) at £17.50: Price per unit is £0.01
Insulin glargine	10-60 IU/day	£10.11	5 × 3-mL Lantus® SoloStar® at £41.50
Insulin aspart	5-200 IU/day	£6.89	100 units/mL NovoRapid®, 10-mL vial at £14.08

Health state costs

Costs of the range of health states were presented in MS Appendix 11 and have been reproduced here for ease of reference. Direct first year costs are presented in Table 60, direct complication and co-morbidity follow-up costs are presented in Table 61 and direct event costs in Table 62. The sources of each of the costs were provided in the MS Appendix 11 and have been checked where possible by the ERG (the references within the Appendix do not correspond to the reference list provided by the MS but the ERG have mostly been able to source these). The descriptions of the source information are not reproduced here but can be found in MS Appendix 11. In general the resource requirement and cost are accurate, there are some instances where the ERG have been unable to source the same information. These relate to:

- For various eye related complications and co-morbidities the MS report the cost of an ophthalmology visit, however, no source for these costs are reported. The ERG have checked the cost of ophthalmology visits in the National Schedule of Reference Costs and are unable to replicate the costs used by the MS for ophthalmology from this data source.
- For various complications and co-morbidities the costs of a biochemistry tests is reported in the MS to be £1 and the source of this cost is reported to be DAP841 in the National Schedule of Reference Costs.[46] The ERG is unable to find this cost for code DAP841 which appears to cost biochemistry at £16.
- For various renal related complications and co-morbidities the cost of nephrologist is reported in the MS to relate to service code 361 in the National Schedule of Reference Costs. The ERG hs been unable to replicate the costs used for nephrology visits using this code. In addition, some renal related complications and co-morbidities have had a cost applied for a 'prescription' at a rage of £16.16 but the ERG have been unable to check what this relates to and whether the cost is accurate.
- A number of procedure related costs for vitreous retinal procedures have been applied in the MS but these have not been identified in the source document referred to.

In addition, the MS reports using an inflation rate to 2011/12 using the Health and Community Health Services (HCHS) index published in 2012, however, the range of dates fall outside the date ranges in the 2012 HCHS publication in some instances and an older version of the HCHS

must have been used. The ERG have sourced the correct inflation rate to check costs included in the model and presented in the MS in Appendix 11.

Table 60: Direct first year treatment costs associated with complication and co-morbidity (MS Appendix 11)

Health state	First year cost (£)
BDR	160.00
ME	674.33
ME and Blindness in One Eye	2402.67
PDR	674.33
PDR and Blindness in One Eye	2402.67
ME and PDR	674.33
ME, PDR and Blindness in One Eye	2402.67
Blindness in Both Eyes	2600.44
Symptomatic Neuropathy	361.60
PVD	4415.91
Neuropathy and PVD	4774.51
Diabetic Foot Ulcer	14,760.20
First LEA	13,828.89
Second LEA	21,448.09
Cost of IHD	3,607.34
Cost of MI	6,354.87
Cost of CHF	4,016.27
Cost of Stroke(s)	5,505.92

BDR: Background Diabetic Retinopathy; CHF: Coronary Heart Disease; IHD: Ischaemic Heart disease; LEA: Lower Extremity Amputation; MI: myocardial infarction; PDR: Proliferative diabetic retinopathy; PVD: Peripheral Vascular Disease

Table 61: Direct complication and co-morbidity annual follow-up costs (MS Appendix 11)

Health state	Annual cost (£)
No Retinopathy	43.00
BDR	43.00
ME	86.00
ME and Blindness in One Eye	768.10
PDR	86.00
PDR and Blindness in One Eye	768.10
ME and PDR	255.00
ME, PDR and Blindness in One Eye	768.10
Blindness in Both Eyes	1038.91
eGFR 60-89 without kidney damage	14.69
eGFR 45-59 without kidney damage	29.38
eGFR 30-44 without kidney damage	29.38
eGFR 15-29, no kidney damage	656.00
eGFR <15 without kidney damage	1312.00
eGFR ≥90, with microalbuminuria	0.00
eGFR 60-89, with microalbuminuria	0.00
eGFR 45-59, with microalbuminuria	88.54
eGFR 30-44, with microalbuminuria	88.54
eGFR 15-29,with microalbuminuria	508.16
eGFR <15, with microalbuminuria	1336.16
eGFR ≥90, with GPR	0.00

eGFR 60-89, GPR	0.00
eGFR 45-59, GPR	88.54
eGFR 30-44, GPR	88.54
eGFR 15-29, GPR	508.16
eGFR <15, GPR	1336.16
eGFR ≥90, ESRD	0.00
eGFR 60-89, ESRD	0.00
eGFR 45-59, ESRD	0.00
eGFR 30-44, ESRD	0.00
eGFR 15-29, ESRD	0.00
eGFR <15, ESRD	24,917.38
No Neuropathy	0.00
Symptomatic Neuropathy	361.60
PVD	908.85
Symptomatic Neuropathy and PVD	1270.45
Diabetic Foot Ulcer	244.59
One LEA	798.19
Two (or more) LEAs	1244.79
History of IHD	1189.36
History of MI	143.43
History of CHF	1407.91
History of Stroke	802.94

eGFR: estimated Glomerular Filtration Rate; BDR: Background Diabetic Retinopathy; CHF: Coronary Heart Disease; IHD: Ischaemic Heart disease; LEA: Lower Extremity Amputation; MI: myocardial infarction; PDR: Proliferative diabetic retinopathy; PVD: Peripheral Vascular Disease

Table 62: Direct event costs associated with complication and co-morbidity (MS Appendix 11)

Health state	First year cost (£)
first MA diagnosis	221.69
first GPR diagnosis	221.69
first ESRD diagnosis	164.00
eGFR dropping below 90	14.69
eGFR dropping below 60	44.08
eGFR dropping below 45	44.08
eGFR dropping below 30	236.38
eGFR dropping below 15	492.00

eGFR: estimated Glomerular Filtration Rate;ESRD: End-stage Renal Disease;

In addition to costs of complications and co-morbidity the MS reports using a range of costs for the adverse events considered in the model. These are seen in Table 40 of the MS (page 208) and have been replicated here (Table 63). These costs and the resources leading to these costs have been checked by the ERG and there are no issues of note.

Table 63: Summary of adverse event costs (MS Table 40)

Health state	First year cost (£)
Minor hypoglycaemic event	0.00
Moderate hypoglycaemic event	0.00
Severe hypoglycaemic event	621.41
lower UTIs (male)	86.82
lower UTIs (female)	43.35
upper UTIs (male)	87.64
upper UTIs (female)	87.64
balanoposthitis	53.60
Vulvovaginitis	45.00

Assessment of uncertainty

Disease progression cost estimates in DSA were applied to adverse events of severe hypoglycaemic events, lower UTI (males), lower UTI (female), balanoposthitis, and candidal vulvovaginitis and these were varied by plus or minus 20%. The costs of complications and comorbidities as given in MS Table 192 (Appendix 11) were subject to PSA.

4.2.8 Consistency/ Model validation

The MS notes that the latest joint International Society for Pharmacoeconomics and Outcomes Research/ Society for Medical Decision Making (ISPOR/SMDM) good research guidelines describe five principal forms of model validation and that these were followed (MS p274). Specifically, the MS notes that the model has been thoroughly tested and de-bugged, including artificial simulations designed to reveal errors in both logic and programming, and that idiosyncratic results were investigated and any identified errors in programming or logic were corrected (MS p274).

An internal validation exercise is reported on MS p276 which assessed the importance of stochastic variability in the case in which small changes versus comparators were observed. The MS concludes that, even in cases with modest incremental benefits, stochastic variability does not override the signal generated in the model simulations (MS p276). The ERG notes that if this is the case then it might also be reasonable to incrementally compare the base case simulation results given in MS Section 7.7.6 (MS p233), but this was not done. Rather, the MS cautions here that because of stochasticity it is inappropriate to mix and match simulation results from different model runs (MS p233).

The ERG explores stochastic variability further in Section 4.3.

External consistency

The MS notes that the external validity of ECHO-T2DM was evaluated by simulating the key features of a broad set of clinical trials and comparing the model predictions with the corresponding actual observed outcomes for a variety of clinical end-points (MS p275). ECHO-T2DM has also been validated against the ADVANCE, ACCORD and ASPEN trials as part of the Fifth Mt. Hood Challenge, observational data from Kaiser Permanent Northwest (KPNW), and the Swedish NDR as part of the Sixth Mt Hood challenge (MS p276). A cross-validation against the CDM model v8.5 is described (MS p277) and a peer-reviewed model validation paper has recently been published.[47] This paper notes that macrovascular outcomes are over-predicted but that this is common in health-economic models of diabetes, and may be related to a general over-prediction of event rates in the UKPDS Outcomes Model.[47]

For the cross-validation exercise against the CDM model the MS notes that three differences were important in explaining the results (given in MS Table 103 p278): time on agent before discontinuation was shorter in the CDM simulations; differences in HbA_{1c} and SBP lowering by agent were reversed immediately in the ECHO-T2DM simulations; and QALYs were applied differently.

Overall the ERG considers that the model is internally consistent and very well-validated, and that the results of the model make intuitive sense.

4.2.9 Assessment of uncertainty

Methodological uncertainty has been addressed by running alternative versions of the model with different assumptions. MS Table 44 (MS p211) gives details of 21 scenario analyses including variations to discount rate and time horizon. Structural uncertainty has also been examined, for example MS Table 44 (MS p211) gives details of scenario analyses around the HbA_{1c} metabolic drift assumption for canagliflozin, and specification of alternative macrovascular risk equations. DSA and PSA were conducted and are described in MS Sections 7.6.2 and 7.6.3 (MS p215-216).

Heterogeneity in the patient population has been examined for one subgroup, those with an HbA_{1c} exceeding 9% (75 mmol/mol) at baseline (MS Table 104 p280 and Table 105 p281).

One-way sensitivity analyses

MS Table 45 (MS p215) lists the DSA which were conducted. Some of these analyses are one-way, whilst others vary sets of parameters at a time in order to reduce the number of DSAs overall. A total of 35 parameters or sets of parameters were varied. DSAs were conducted for each of the systematic drug comparisons shown in MS Table 43 (MS p210), namely comparison against SU and DPP-4-i in dual therapy and against DPP-4-i in triple therapy and add-on to insulin therapy. The MS indicates the model was run with 500 patient cohorts and 1000 patients in each cohort in sensitivity analysis (Table 53), i.e. the number of patient cohorts is halved compared to the base case. The MS indicates that this was a pragmatic decision based upon the number of comparators, lines of therapy, the time taken for each pairwise comparison to run, and the finite computing resource available (MS p209). Parameter uncertainty was not switched off in DSA (i.e. by specifying only one patient cohort). All DSA results are consequently confounded with PSA and do not only reflect the uncertainty in results associated with just one parameter.

The MS states that parameter values were varied by plus or minus 20% in DSA (MS p215), but this range is not justified. The DSA findings indicate that in dual therapy comparisons against SU the most influential parameter on the ICER is the metabolic drift of HbA_{1c} (both for SU and for canagliflozin) (MS Figures 66 & 68). For canagliflozin 300mg versus DPP-4-i in dual therapy the greatest impact on the base case ICER is metabolic drift, incremental cholesterol and SBP therapeutic effects and discontinuation rate of DPP-4-I (MS Figure 69 p242). In add-on to insulin only increasing the DPP-4-i HbA_{1c} drift, or decreasing the canagliflozin 100mg HbA_{1c} drift switched the result from not cost-effective to cost-effective (MS p238). For canagliflozin 300mg versus DPP-4-i in combination with insulin, HbA_{1c} drift was also the major source of variation in ICER (MS p238, MS Figure 71 p244).

The MS notes that dominant and dominated results have been removed from the tornado diagrams (MS p239). Tabulated results are presented in MS Tables 195 to 202 (Appendix Section 10.13). The ERG considers that the tornado diagrams and tabulated results would have better represented the uncertainty in the decision problem if the outcome measure had

been net monetary benefit rather than ICER, as it would not have been necessary to remove any results had this measure been used.

Tables detailing the input parameter values for the DSAs are given in MS Appendix Section 10.13 (Tables 195 to 202). The ERG has examined these tables and found that the variation in the values given is not always plus or minus 20%, and that the level of deviation from 20% is not consistent. This is shown in the shaded column in Table 64 below. The discrepancy applies to some of the efficacy parameters which are described in Tables 195 to 202 as being incremental changes between canagliflozin and the comparators (Table 64). However the values given in these tables for, for example incremental HbA_{1c}, appear to relate to the absolute difference of canagliflozin versus placebo (with annual drift added back as noted in Section 4.2.4). The incremental base values which have been used to obtain the plus and minus 20% values for treatment effects in DSA are not given in Tables 195 to 202 and their source is not identified in the MS discussion of the DSA (MS p215).

Table 64 indicates that the variation which has been examined in DSA for the incremental HbA_{1c} change between CANA and SU is only 2.1% of the base case difference of CANA versus placebo. The ERG considers that as the key parameter in the model here is change in HbA_{1c} versus placebo (MS Appendix Section 10.9 Tables 185 to 188), that +/-20% uncertainty in this value should have been examined, rather than +/-20% of the incremental difference between treatments. Elsewhere the DSA does vary parameter values by +/-20% of each treatment baseline, rather than by +/-20% of an incremental difference between treatments – for example for HbA_{1c} drift % - so the DSA is inconsistent in this respect. The range examined in DSA is also much smaller than the range incorporated in the PSA distribution which is Normal with parameters (-0.77, 0.26) (input data spreadsheet). However DSA results indicate that the model is not unduly sensitive to first year change in HbA_{1c} % (MS Section 7.7.7) and the ERG does not consider that examination of a bigger first year change in HbA_{1c} % will substantively change these results.

The manufacturer concludes from the DSA that in dual therapy comparisons against SU, the ICERs of both doses of canagliflozin vary little (MS p237). For canagliflozin 300mg versus DPP-4-i in dual therapy, the base case ICER is higher and there are a greater number of DSAs that change the results and bring them closer to conventional acceptability thresholds of £20,000 to 30,000 (MS p238).

Table 64: Treatment effect values used in DSA for canagliflozin 100mg vs. SU in dual therapy (extract of MS Table 195 Appendix p322 with additional column)

Parameter	Lower input	Base case	Upper input	Lower input ICER	Base case ICER	Upper input ICER	Difference	% of ICER variability	Absolute variation from base case (ERG calculation)
Efficacy parameters (52 weeks)									
Δ incremental HbA1c between Cana and SU	-0.786	-0.77	-0.754	1,526	1,537	1,629	103	6.70%	2.1%
Δ incremental SBP between Cana and SU	-2.876	-3.57	-4.264	1,410	1,537	1,359	51	3.30%	19.4%
Δ incremental LDL between Cana and SU	4.01	4.56	5.11	1,793	1,537	1,554	239	15.50%	12.1%
Δ incremental HDL between Cana and SU	2.458	3.13	3.802	1,545	1,537	1,600	55	3.60%	21.5%
Δ incremental triglycerides between Cana and SU	-15.426	-19.054	-22.674	1,661	1,537	1,474	187	12.20%	19.0%
Δ incremental total cholesterol between Cana and SU	3.788	4.27	4.752	1,520	1,537	1,863	344	22.40%	11.3%
Δ incremental BMI between Cana and SU	-0.253	-0.595	-0.937	1,497	1,537	1,259	237	15.50%	57.5%

In triple therapy (MET+SU), canagliflozin 100mg is consistently dominant when compared to DPP-4-i across the majority of the parameters (MS p238). In add-on to insulin, canagliflozin 100mg was dominated by DPP-4-i in most DSAs as in the base case. Only increasing the DPP-4-i HbA_{1c} drift, or decreasing the canagliflozin HbA_{1c} drift switched the result (MS p238). For canagliflozin 300mg versus DPP-4-i in combination with insulin, as with all the other DSAs, HbA_{1c} drift was a major factor, and was the only factor that changed the result from cost-effective to not cost-effective (MS p238 & MS figure 71 p244).

The ERG agrees with the manufacturer's conclusions from the DSA analyses given the results presented. However the ERG does not consider that the uncertainty in the decision problem has been fully explored, or that the results have been presented in the most informative way to reflect all results. It is unclear why second order uncertainty was not switched off in DSA as this means that results are confounded to some extent by stochasticity in other parameters.

Scenario Analysis

MS Table 44 (MS p211) gives details of 21 scenario analyses which were carried out in varying detail on the various treatment comparisons. Not all analyses were carried out on every comparison; MS Table 42 and Table 43 (MS p210) together identify the number of scenario analyses associated with each comparison. Modelling the canagliflozin dose increase from 100mg to 300mg was specifically explored as a scenario rather than being incorporated into the base case modelling due to the late finalisation of the posology in the SmPC (Appendix 1, Section 10.1) (MS p272).

The MS indicates the model was run with 500 patient cohorts and 1000 patients in each cohort in scenario analysis (Table 53). Second order uncertainty was not switched off in scenario analysis (i.e. by specifying only one patient cohort) which means that results are to some extent confounded with parameter stochasticity.

The 21 scenario analyses and assumed alternative values are given in MS Table 44 (MS p211). Some justification is given in the MS main report for the dose modification scenario and the weight loss and rebound scenarios. Justification for the other scenarios is provided in Appendix Section 10.12. Appendix Section 10.12 reports that scenario 16 uses head to head RCT data for the dual therapy comparisons against SU and DPP-4-i, and triple therapy (MET+SU)

comparison against DPP-4-i (300mg canagliflozin only, as the DIA3015 only studies the 300mg dose). It is not clear why the data used in scenario 16 were restricted to direct trials against active comparators, as the model uses change versus placebo for key biomarkers. Results of the scenarios are given in MS Tables 78-100 (MS p256-271).

The manufacturer concludes that the canagliflozin dose increase from 100mg to 300mg improves the cost-effectiveness in each line of therapy and against each comparator, with the exception of dual therapy versus SU, where the ICER changes to being slightly less cost-effective (from £1,537 to £1,721) (MS p272 and MS Table 78 p256). The ERG notes that cost-effectiveness was not improved in the dose modification scenario for canagliflozin 100mg versus TZD in dual therapy, where canagliflozin remains dominated by TZD (as in the base case).

In combination with insulin, where the base case ICERs versus DPP-4-i or GLP-1-a were not cost-effective, the manufacturer concludes that the dose modification step improves the ICER so that canagliflozin 100mg is estimated to be cost-effective versus DPP-4-i or GLP-1-a in combination with insulin (MS Table 96 and MS Table 100) (MS p272). The ERG notes that MS Table 96 indicates that canagliflozin in add on to insulin is cheaper and less effective than DPP-4-i in the base case, but in the dose modification scenario is associated with an ICER of £503 per QALY. MS Table 100 indicates that canagliflozin in add on to insulin is cheaper and less effective than GLP-1-a in the base case and remains so in the treatment modification scenario, albeit at relatively higher incremental cost and with relatively more incremental QALYs.

The MS states that the conclusion of the dose modification scenario analysis (scenario 4 in MS Table 44 p211) is that the dose intensification schedule from 100mg to 300mg is cost-effective (MS p272). As noted above, the ERG does not consider that this conclusion applies to all treatment comparisons.

The MS highlights the importance of HbA_{1c} drift on cost-effectiveness indicated by the scenario analyses (MS scenario 8). The use of a higher HbA_{1c} metabolic drift in scenario analysis results in an ICER of £30,358 for canagliflozin 100mg versus SU in dual therapy (MS Table 78 p256); of £69,464 for canagliflozin 300mg versus SU in dual therapy (MS Table 80 p258); and of £698,330 for canagliflozin 100mg versus TZD in dual therapy.

The ERG observes that whilst the scenario for no weight/BMI rebound following treatment discontinuation (MS scenarios 12 and 13) generally renders canagliflozin more cost-effective versus comparators (lower ICER than base case - for example MS Tables 78, 79, 80 and 81), this does not happen in the scenarios for canagliflozin 100mg and canagliflozin 300mg versus dapagliflozin in dual therapy (MS Tables 84 and 85 p261). In these scenarios the ICER is £24,697 compared to £8,674 in the base case for canagliflozin 100mg (MS Table 84), and £33,368 compared to £27,419. These results bear out the clinical finding that dapagliflozin is better at controlling BMI than both canagliflozin 100mg and canagliflozin 300mg (MS Table 185, Appendix p299).

Results of the scenario analysis for canagliflozin 300mg versus dapagliflozin in dual therapy show that canagliflozin 300mg is not cost-effective at a WTP threshold of £30,000/QALY in the scenarios for HbA_{1c} switch threshold; patient baseline characteristics; BMI rebound (as noted above); and disutility from weight gain (MS Table 85 p261).

The MS does not discuss the results of most of the scenario analyses specifically but concludes that the incremental gain offered by canagliflozin 100mg and 300mg is consistent across a broad set of conditions as the ICERs in the vast majority of cases remain cost-effective (MS p272). The ERG agrees with this assessment in general, with the exceptions noted above. Dapagliflozin is more cost-effective than canagliflozin 300mg in several scenarios. Parameter uncertainty was not switched off in scenario analysis.

Probabilistic sensitivity analysis

As stated earlier, a model PSA using 1000 cohorts of 1000 patients with a lifetime horizon of 40 years takes over 15 hours to run on a high specification desktop computer (Intel i7 16GM RAM). The MS states that PSA is a default feature of the ECHO-T2DM model and was run on all simulations conducted (MS p216). All base case simulations thus incorporate PSA and their results are given in Table 48, Table 49 and Table 50. Cost-effectiveness planes and acceptability curves are presented in MS Figures 72 to 87 (MS p245-254), for selected base case comparisons. Table 65 summarises the PSA results for the comparisons which are reported in the MS.

Superseded – see erratum

Table 65. Summary of PSA results presented in MS

Intervention	Comparator	Probability cost-effective at £20,000/QALY	Probability cost-effective at £30,000/QALY
Dual therapy with MET			
Canagliflozin 100 mg	SU	92.9	93.6
	DPP-4-i	54.5	54.9
Canagliflozin 300 mg	SU	89.5	92.1
	DPP-4-i	51.6	54.9
Triple therapy with MET and SU			
Canagliflozin 100 mg	DPP-4-i	55.7	55.6
Canagliflozin 300 mg	DPP-4-i	55.1	58.2
Add-on to insulin plus standard of care			
Canagliflozin 100 mg	DPP-4-i	44.7	45.2
Canagliflozin 300 mg	DPP-4-i	60.7	63.7

MS Table 46 (reproduced as Table 66 below) shows the stochastic variables and associated distributions. The ERG notes that some of these variables concern patient-level (first order) uncertainty rather than parameter (second order) uncertainty, and that some of these variables are not stochastic (e.g. number of cohorts). Some variables are noted to account for both first and second order uncertainty. Table 66 indicates that all appropriate variable groupings have been included in PSA but individual variables are not specified in the table.

The MS does not report the parameters of the distributions for variables included in PSA. Mean values for the stochastic variables are only supplied in the input data spreadsheet, not in the MS. The standard error for the HbA1c % treatment effects only reflects the NMA estimate variation, and not the variation in annual drift, as noted in Section 4.2.4.

Standard errors (SE) for many parameters with normal distributions are not given in the input data spreadsheet. Table 66 indicates that SE are optional in some cases. It is not clear what the default standard error is in these cases as it is not specified in the input data or described in the Technical Report.{17} Table 66 notes that the distribution for probabilities is Bernoulli/Beta, to account for 1st/2nd order uncertainty, but no beta distribution parameters are identified in the input data spreadsheet or the MS. The beta distributions are found in the model R code in various locations. The R code does not make clear how the distributions were parameterised.

The submission does not identify the locations where the various PSA distributions are specified in the R code and due to time constraints the ERG did not check their coding.

Correlation is addressed for initial age and biomarker values. For the UKPDS 68 mortality and macro-vascular risk equations second order (parameter) uncertainty is captured using a set of 1,000 boot-strap replications of the regression coefficients (Technical Report p67). Correlation for the UKPDS 82 equations has not yet been applied as these have not yet been made available by the study authors (manufacturer's clarification letter to NICE and ERG).

In summary, although all relevant variables appear to have been included in PSA, the distributional assumptions for these variables are not well described in the MS and are not transparent from the input data spreadsheet or R code. Consequently it is not clear to what extent uncertainty in the decision problem has been explored. The standard errors used for the HbA1c % treatment effect parameters are too small. Cost-effectiveness acceptability data are not presented for all base-case comparisons.

PSA using 1000 iterations is included in the model base-case results (Section 4.1). The ERG reports base-case results with no second-order uncertainty, and cost-effectiveness acceptability data for key comparisons not reported in the MS, in Section 4.3.

4.2.10 Comment on validity of results with reference to methodology used

The ERG considers that the economic model captures all important aspects of the disease pathway and extrapolates intermediate outcomes to final outcomes in a robust and consistent manner, drawing upon standard sources from the literature. The model has been very well validated against external data.

The model is run with 1000 patients in the base case. It is unclear if this is sufficient to robustly estimate the ICERs for all treatment comparisons. The base case ICERs incorporate parameter uncertainty. Base case results are presented as pairwise treatment comparisons. Fully incremental analyses are not supplied.

Parameter distributions used in PSA are not well documented and the parameter ranges used for some variables in one-way sensitivity analysis do not reflect much uncertainty. PSA results are not presented for all base case treatment comparisons.

Table 66: Parameter distributions for PSA in ECHO-T2DM (reproduced from MS Table 46 p217)

Parameters	Quantity	Distribution	First Column			Second Column		
			Parameter Type	Range of Permissible Values	Comment	Parameter Type	Range of Permissible Values	Comment
Baseline biomarker values, age at baseline, HbA1c at diagnosis	Mean, SD	(Multivariate) Normal	Mean	[0, ∞)	Accounts for 1st order uncertainty	SD	[0, ∞)	
Correlation coefficients for initial age and bio-marker values	Correlation coefficients	Part of above multivariate normal	ρ	(-1, 1)	Accounts for 1st order uncertainty	N/A	N/A	
Treatment effects, annual bio-marker drift, treatment target thresholds	Mean, SE	Normal	Mean	(-∞, ∞)	2nd order uncertainty	Optional SE	(-∞, ∞)	
Patient-level (1st order uncertainty) variation in treatment effects and intensification targets	SD		SD	(-∞, ∞)	Accounts for 1st order uncertainty	N/A		
# of cohorts, # of patients per cohort, modelled time horizon	Constant Integers (Positive)	N/A	Integer	[1, ∞)	N/A	N/A	N/A	
Maximum treatment duration, duration of discontinuation rebound, time until treatment effects completely depreciated	Mean	N/A	Mean	[0, ∞)	N/A	N/A	N/A	
Baseline T2DM duration, logical bounds on initial bio-marker values	Minimum, Maximum	Uniform	Minimum	[0, maximum]	Accounts for 1st order uncertainty	Maximum	[Minimum, ∞)	Sometimes optional
Proportions with co-morbidities, ethnicity, gender, proportion non-responders	Proportion	Bernoulli or categorical	Proportion	[0, 1]	Accounts for 1st order uncertainty	N/A	N/A	
Probabilities (e.g., AE's, non-compliance, dose-reduction, discontinuation, smoking cessation, contraindications)	Probability, SE	Bernoulli/Beta	Probability	[0, 1]	1st order/2nd order uncertainty	Optional SE	[0, ∞)	SE restricted by distribution assumption
Hypoglycaemic event rates	Mean, SE	Poisson/ Lognormal	Mean	[0, ∞)	1st order/2nd order uncertainty	Optional SE	[0, ∞)	
Hazard Ratio (HR) for AE's, rates associated with declining eGFR	Mean, SE	Lognormal	Mean	[0, ∞)	2nd order uncertainty	Optional SE	[0, ∞)	
Relative Risks (RR) for AE's and complications	Mean, SE	Lognormal	Mean	[0, ∞)	2nd order uncertainty	Optional SE	[0, ∞)	
Multipliers for eGFR drift, eGFR-related treatment effect multipliers	Mean, SE	Normal	Mean	(-∞, ∞)	2nd order uncertainty	Optional SE	[0, ∞)	
Unit costs	Mean, SE	Normal	Mean	(-∞, ∞)	2nd order uncertainty	Optional SE	[0, ∞)	
QALY disutility weights	Mean, SE	Normal	Mean	(-∞, ∞)	2nd order uncertainty	Optional SE	[0, ∞)	
Bio-marker logical limits	Minimum, Maximum	N/A	Optional Minimum	(-∞, maximum]	N/A	Optional Maximum	[Minimum, ∞)	Must be larger than minimum

4.3 Additional work undertaken by the ERG

The ERG undertook additional work to:

a) Examine the variation in final ICER arising through re-running some of the base case analyses

b) Examine the variation in final ICER arising with no parameter uncertainty and 100,000 patients

c) Produce incremental analyses for dual therapy

d) Identify the probability that canagliflozin is cost effective in comparison with dapagliflozin

e) Examine the effect of modifications to the clinical efficacy of glimepiride and exenatide

a) Reproduction of base case analyses

The ERG re-ran three of the dual therapy base-case analyses for both canagliflozin 100 mg and canagliflozin 300 mg in order to examine the uncertainty in incremental costs, QALYs and final ICER which can arise from patient and parameter stochasticity. Results are given in Table 67 and Table 68.

Table 67 shows that incremental costs for canagliflozin 100mg versus SU vary from £288 in the manufacturer's base case to £303 in the ERG analysis, a difference of £15. Differences of £20 and £44 in incremental costs were found for the dapagliflozin and DPP-4-i comparisons respectively. Differences in incremental QALYs were also small. However because the ICER is a ratio, it tends to infinity as its denominator (i.e. incremental QALYs) tends to zero. The resulting differences in ICER between the two sets of base cases thus appear larger than these small incremental differences suggest, particularly in the case of canagliflozin 100 mg and DPP-4-i, where the ICER increases from £97 in the MS base case to £3,926 in the ERG base case (Table 67).

The ICERs from the ERG re-running of the canagliflozin 300mg base cases show slight increase for the comparisons with dapagliflozin and DPP-4-i (Table 68), but are broadly comparable to the MS base cases. Changes in incremental costs and QALYs are also small.

Table 67: Base case incremental cost-effectiveness results for canagliflozin 100mg in dual therapy, ERG analysis compared to MS analysis (1000 patients and 1000 cohorts)

MS results						ERG analysis				
Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
TZD	£23,236	9.119				Not run by ERG				
Cana 100mg	£26,069	8.96	£2,833	-0.159	Dominated					
SU	£25,894	8.751				£25,742	8.754			
Cana 100mg	£26,182	8.938	£288	0.188	£1,537	£26,046	8.947	£303	0.194	£1,566
Dapagliflozin	£26,086	8.923				£26,207	8.959			
Cana 100mg	£26,149	8.93	£63	0.007	£8,674	£26,250	8.965	£43	0.007	£6,685
DPP-4-i	£26,100	8.928				£26,163	8.950			
Cana 100mg	£26,102	8.941	1	0.013	£97	£26,208	8.961	£45	0.012	£3,926
GLP-1-a	£28,390	8.97				Not run by ERG				
Cana 100mg	£25,966	8.921	-£2,424	-0.048	£50,005 #					

ICER is for GLP-1-a to replace canaglilozin 100 mg

Table 68: Base case incremental cost-effectiveness results for canagliflozin 300mg in dual therapy, ERG analysis compared to MS analysis (1000 patients and 1000 cohorts)

MS results						ERG analysis				
Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
TZD	£23,204	9.092				Not run by ERG				
Cana 300mg	£26,556	8.951	£3,353	-0.141	Dominated					
SU	£25,753	8.761				£25,614	8.744			
Cana 300mg	£26,729	8.961	£976	0.199	£4,899	£26,572	8.941	£957	0.197	£4,857
DPP-4-i	£25,926	8.902				£25,986	8.931			
Cana 300mg	£26,501	8.933	£576	0.031	£18,349	£26,507	8.956	£521	0.0250	£20,836
Dapagliflozin	£26,044	8.91				£26,088	8.957			
Cana 300mg	£26,668	8.932	£625	0.023	£27,419	£26,683	8.978	£595	0.0205	£28,975
GLP-1-a	£28,515	8.972				Not run by ERG				
Cana 300mg	£26,624	8.947	-£1,892	-0.025	£76,214 #					

ICER is for GLP-1-a to replace canagliflozin 300 mg

b) Base case analyses with no parameter uncertainty and more patients

The MS reports base cases for model simulations with 1000 patients and 1000 cohorts. Given the small incremental benefits arising from the various treatments, and the potential complexity

of the diabetes disease pathway, the ERG considers that the model should ideally be run with more than 1000 patients. This will provide a more robust estimate of the final ICER. The ERG ran the model with 100,000 patients and at the same time explored the effect of using no parameter uncertainty on model results. (This also avoided a very long simulation time.) Results are given for canagliflozin 100 mg in dual therapy in Table 69 and canagliflozin 300 mg in dual therapy in Table 70.

Results from Table 69 can be compared with results from the original MS base case in Table 67. Canagliflozin 100 mg remains dominated by TZD whilst the ICER compared to SU is very similar to the original base case (£1,579 compared to £1,537). The ICER compared to dapagliflozin increases to £100,719 but this is driven largely by a very small incremental difference in total QALYs, as the incremental costs are only £130 higher. The ICER compared to DPP-4-i has increased to £12,938 compared with £97 in the original base case, again driven by small incremental changes.

Table 69: Base case incremental cost-effectiveness results for canagliflozin 100mg in dual therapy, 100,000 patients and no parameter uncertainty (1 cohort)

Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
TZD	£23,336	9.153			
Cana 100mg	£26,264	8.987	£2,929	-0.166	Dominated
SU	£25,941	8.793			
Cana 100mg	£26,216	8.967	£274	0.174	£1,579
Dapagliflozin	£26,100	8.956			
Cana 100mg	£26,294	8.958	£193	0.002	£100,719
DPP-4-i	£25,988	8.974			
Cana 100mg	£26,199	8.990	£211	0.016	£12,938
GLP-1-a	£28,576	9.016			
Cana 100mg	£26,195	8.981	-£2,381	-0.035	£67,414 #

ICER is for GLP-1-a to replace Cana 100 mg

Results in Table 70 compared to the original base case (Table 68) show a slightly worse ICER for canagliflozin 300 mg compared to SU, but more cost-effective ICERs against DPP-4-i and dapagliflozin. Again the incremental costs and QALYs associated with these comparisons are very small.

Table 70: Base case incremental cost-effectiveness results for canagliflozin 300mg in dual therapy, 100,000 patients and no parameter uncertainty (1 cohort)

Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
TZD	£23,371	9.145			
Cana 300mg	£26,616	9.003	£3,245	-0.14	Dominated
SU	£25,773	8.797			
Cana 300mg	£26,894	9.006	£1,121	0.21	£5,368
DPP-4-i	£26,225	8.972			
Cana 300mg	£26,637	9.016	£412	0.04	£9,246
Dapagliflozin	£26,121	8.965			
Cana 300mg	£26,872	9.008	£751	0.04	£17,161
GLP-1-a	£28,702	9.008			
Cana 300mg	£26,836	8.987	-£1,867	-0.0216	£86,412 #

ICER is for GLP-1-a to replace canagliflozin 300 mg

c) Fully incremental analysis for dual therapy

The MS provides many pairwise analyses of canagliflozin versus the key comparators in the various lines of therapy (Table 48,Table 49 and Table 50). The MS states that this is because of stochasticity in the model and that it is therefore inappropriate to mix and match simulation results from different runs (MS p233). The ERG notes in addition that many of the clinical input parameters (e.g. for first year change in total cholesterol) come from trials which directly compare intervention and comparator, and that the values used for these parameters can therefore vary, depending on the treatment comparison being made. Thus mixing and matching of results would in some cases also entail mixing and matching of individual arms of separate trials, without adjusting for trial populations. However the MS argues that data from the NMA are used where available for methodological consistency (MS p161). Given this consistency the ERG considers that it should to some extent be possible to compare simulation results from different model runs, particularly if each model run uses a large number of patients (to offset stochasticity in the patient population).

A fully incremental analysis based upon the MS original base case for canagliflozin 100 mg in dual therapy is given in Table 71 and plotted in Figure 2 (NB. dual therapy was prioritised for analysis over triple therapy due to time constraints and data availability.) An incremental analysis for canagliflozin 100mg in dual therapy using no parameter uncertainty and 100,000

patients is given in Table 72 and plotted in Figure 3. TZD is excluded from these analyses; it dominates all treatments but because of concerns about adverse events its use is declining in UK clinical practice.

Table 71: Fully incremental analysis based upon MS original base case, canagliflozin 100mg in dual therapy. Excludes TZD.

Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER compared to next cheapest option (excluding dominated)
SU	£25,894	8.751			
Dapagliflozin	£26,086	8.923	£192	0.172	£1,116
Cana 100mg*	£26,094	8.938	£8	0.015	£507
DPP-4-i	£26,100	8.928	£6	-0.01	Dominated
GLP-1-a	£28,390	8.97	£2,296	0.032	£71,763

* Data for canagliflozin 300 mg averaged across the four pairwise comparisons

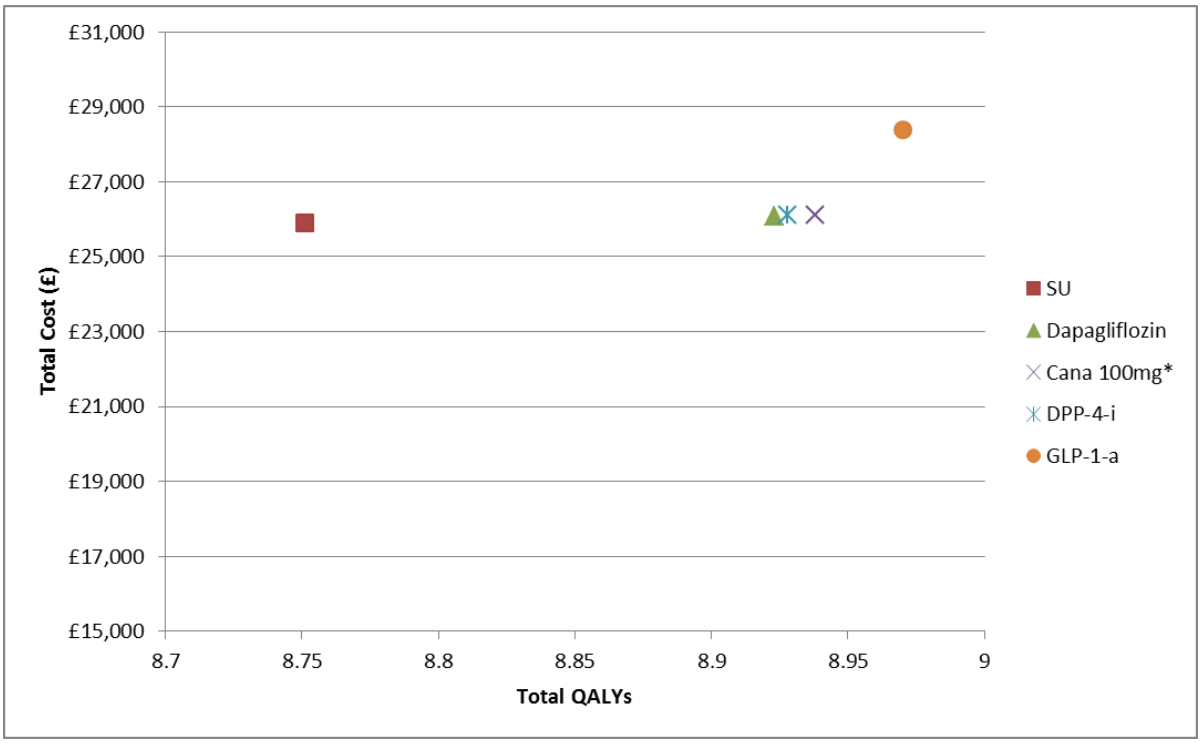


Figure 2: Cost-effectiveness of canagliflozin 100mg and other treatments in dual therapy based upon MS original base case. Excludes TZD.

Table 72: Fully incremental analysis based upon ERG base case with 100,000 patients and no parameter uncertainty, canagliflozin 100mg in dual therapy. Excludes TZD.

Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER compared to next cheapest option (excluding dominated)
SU	£25,941	8.79			
DPP-4-i	£25,988	8.974	£47	0.180	£261
Dapagliflozin	£26,100	8.956	£112	-0.018	Dominated
Cana 100mg*	£26,234	8.977	£245	0.003	£84,800
GLP-1-a	£28,576	9.016	£2,343	0.040	£59,214

* Data for canagliflozin 100 mg averaged across the four pairwise comparisons

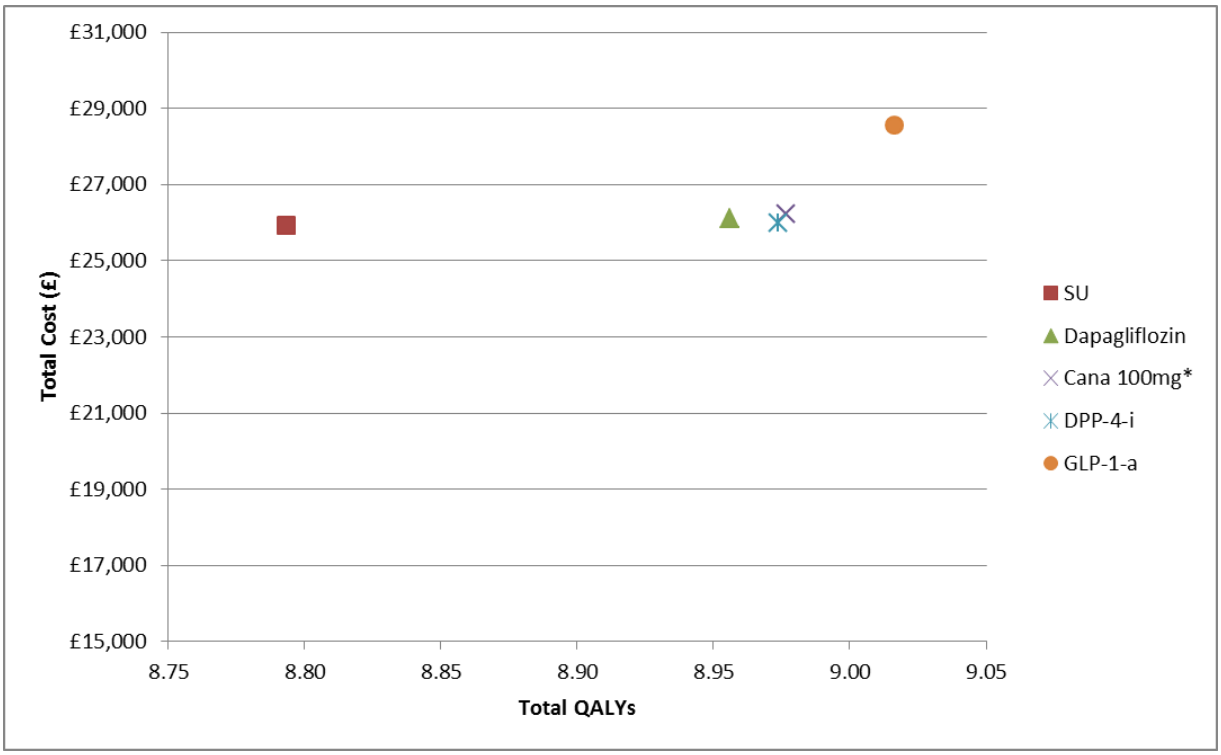


Figure 3: Cost-effectiveness of canagliflozin 100mg and other treatments in dual therapy based upon ERG base case with 100,000 patients and no parameter uncertainty. Excludes TZD.

It can be seen from Figure 2 and Figure 3 that dapagliflozin, canagliflozin 100mg and DPP-4-i are associated with a similar total cost and similar overall total QALYs. In the MS base case canagliflozin 100 mg has an ICER of £507 in comparison with the next best option (dapagliflozin) whilst in the ERG analysis (

Table 72) canagliflozin 100 mg has an ICER of £84,800 in comparison with the next best option (DPP-4-i). The uncertainty in final ICER is driven by the very small incremental benefits and costs associated with these three treatments over the time horizon (Table 71 and Table 72).

An incremental analysis based upon the MS original base case for canagliflozin 300 mg in dual therapy is given in Table 73 and plotted in Figure 4. An incremental analysis for canagliflozin 300mg in dual therapy using no parameter uncertainty and 100,000 patients is given in

Table 74 and plotted in Figure 5. In these analyses canagliflozin 300mg is associated with somewhat higher total QALYs than DPP-4-i or dapagliflozin, at higher cost. The ICER for canagliflozin 300 mg compared to the next best option (dapagliflozin) is £17,639 per QALY using data from the MS original base case (Table 73), and £17,903 per QALY compared to the next best option (DPP-4-i) using data from the ERG analysis with more patients and no parameter uncertainty (Table 74).

Table 73: Fully incremental analysis based upon MS original base case, canagliflozin 300mg in dual therapy. Excludes TZD.

Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER compared to next cheapest option (excluding dominated)
SU	£25,753	8.761			
DPP-4-i	£25,926	8.902	£173	0.141	£1,227
Dapagliflozin	£26,044	8.910	£118	0.008	£14,750
Cana 300mg*	£26,631	8.943	£587	0.033	£17,639
GLP-1-a	£28,515	8.972	£1,885	0.029	£65,548

* Data for Cana 300 mg averaged across the four pairwise comparisons

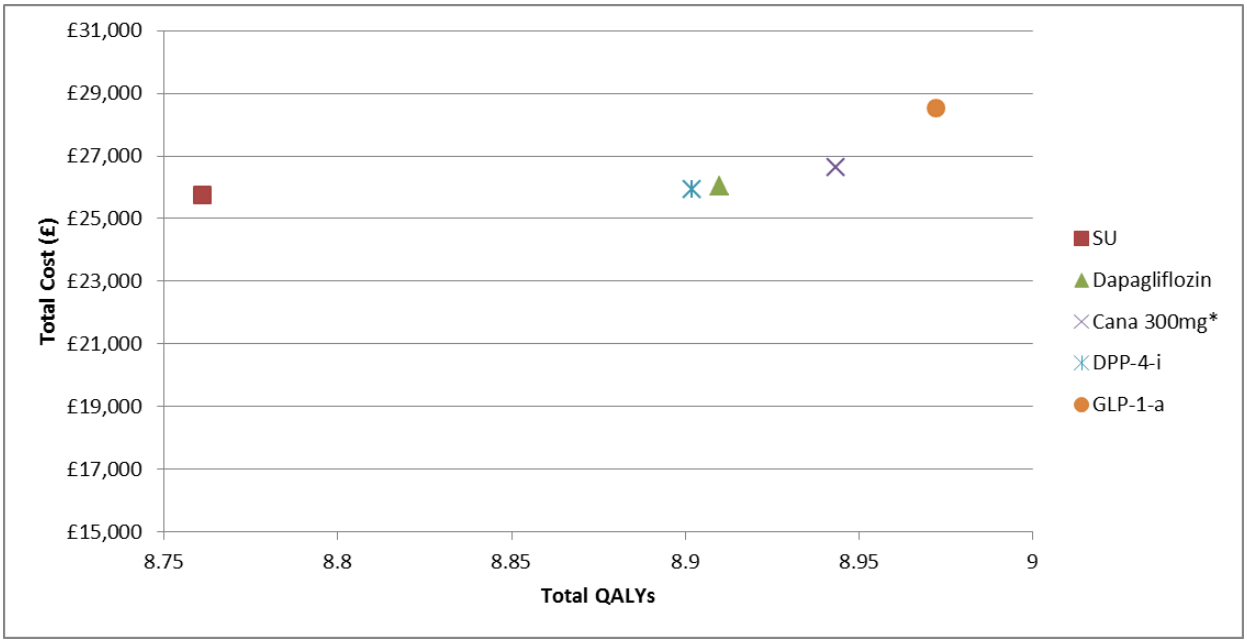


Figure 4: Cost-effectiveness of canagliflozin 300mg and other treatments in dual therapy based upon MS original base case. Excludes TZD.

Table 74: Fully incremental analysis based upon ERG base case with 100,000 patients and no parameter uncertainty, canagliflozin 300mg in dual therapy. Excludes TZD.

Comparator	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER compared to next cheapest option (excluding dominated)
SU	£25,773	8.797			
Dapagliflozin	£26,121	8.965	£348	0.168	£2,078
DPP-4-i	£26,225	8.972	£104	0.007	£15,024
Cana 300mg*	£26,809	9.004	£585	0.033	£17,903
GLP-1-a	£28,702	9.008	£1,893	0.004	£492,690

* Data for Cana 300 mg averaged across the four pairwise comparisons

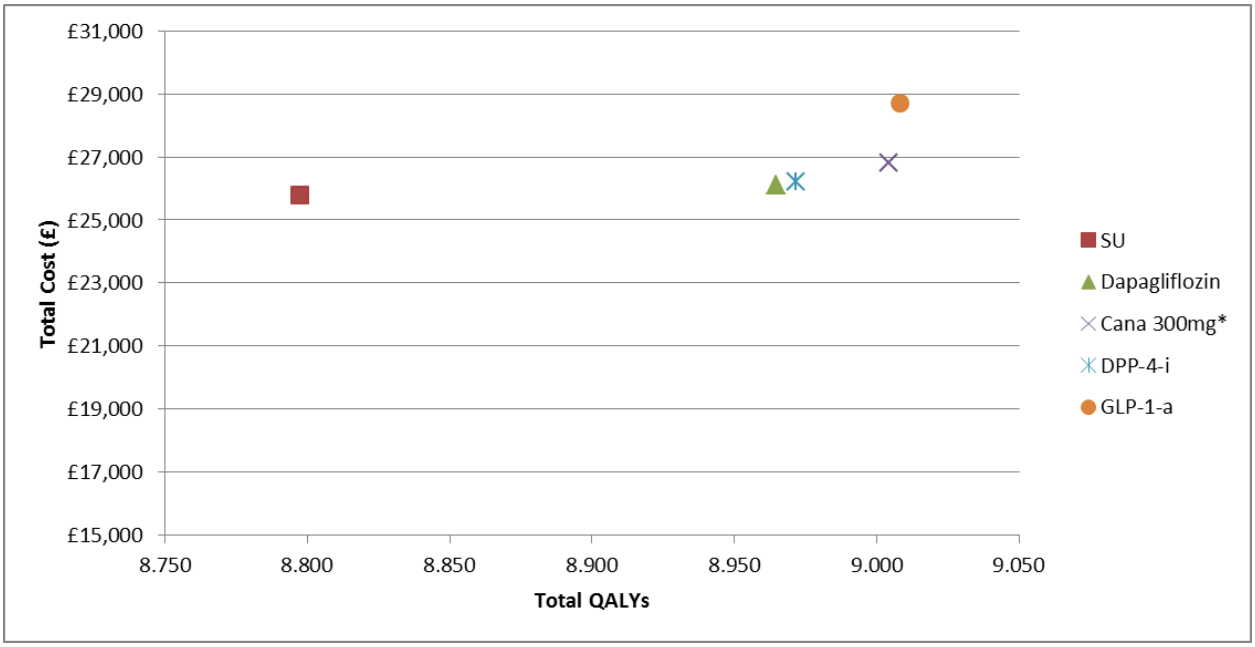


Figure 5: Cost-effectiveness of canagliflozin 300mg and other treatments in dual therapy based upon ERG base case with 100,000 patients and no parameter uncertainty. Excludes TZD.

d) Probability that canagliflozin is cost effective in comparison with dapagliflozin

The MS does not report the probability that canagliflozin is cost-effective compared to the other drug in the SGLT-2 class, dapagliflozin. These probabilities are given in Table 75 for WTP thresholds of £20,000 and £30,000 per QALY, for dual therapy with metformin. The probabilities are taken from ERG model simulations with base case assumptions.

Table 75 indicates that canagliflozin 100 mg has a probability of being cost-effective of 52.5% compared to dapagliflozin in dual therapy with metformin, at both £20,000 and £30,000 WTP. Canagliflozin 300 mg has a probability of being cost-effective of 46.7% compared to dapagliflozin in dual therapy with metformin at £20,000 WTP; this probability increases to 50% at £30,000 WTP.

Table 75: Summary of ERG PSA results for canagliflozin in comparison with dapagliflozin in dual therapy with metformin (1000 patients and 1000 cohorts)

Intervention	Comparator	Probability cost-effective at £20,000/QALY	Probability cost-effective at £30,000/QALY
Dual therapy with MET			
Cana 100 mg	Dapagliflozin	52.5	52.5
Cana 300 mg	Dapagliflozin	46.7	50.0

e) Modifications to the clinical efficacy of gliclazide and exenatide

The manufacturer's clarification letter to NICE and the ERG notes that gliclazide is associated with a lower rate of hypoglycaemia than glimepiride, another drug in the SU class which was used to inform the clinical efficacy of this class of treatments in the economic model. The manufacturer examined variation of +/- 20% to this hypoglycaemia rate in sensitivity analysis and found that it did not have great influence on model results. The ERG notes that gliclazide is associated with approximately half of the hypoglycaemia episodes of glimepiride[37] and so examined variation of -50% in the hypoglycaemia event rate of glimepiride. The ICER increased to £2,516 compared to £1,537 in the base case (Table 48). The ERG agrees with the manufacturer that this parameter is not unduly influential.

The manufacturer uses efficacy data for exenatide BID in cases where efficacy data for liraglutide were not available. However liraglutide is more effective than exenatide BID, although at higher cost (Section 4.2.3). The ERG found that using liraglutide cost and efficacy data reduced the ICER for the GLP-1-a class of treatments compared to canagliflozin (in add on to insulin), but that canagliflozin still remained cheaper and less effective. Again this efficacy parameter does not appear to be unduly influential. This finding is in agreement with the results of the DSA reported in the MS.

Summary of ERG analyses

Re-running of the manufacturer's base case analyses produces slightly different cost and QALY outcomes to those reported for the base cases in the MS, because of patient and parameter stochasticity. Results also vary slightly when a bigger patient population is used, and parameter

uncertainty is turned off. Overall the differences in total cost and QALYs between the various model runs are small (typically hundreds of pounds or less, and thousandths of a QALY). However the small differences in incremental QALYs in some cases drive an apparently large variation in the ICER.

Fully incremental analysis conducted by the ERG for dual therapy in combination with metformin found that, with no parameter uncertainty and 100,000 patients, canagliflozin 100 mg is associated with an ICER of £84,800 compared to the next best option, DPP-4-i. The incremental differences for this comparison are £245 in total costs and 0.003 total QALYs over the lifetime horizon of the model. Canagliflozin 300 mg is associated with an ICER of £17,903 compared to the next best option, DPP-4-i, with incremental total costs of £585 and incremental total QALYs of 0.033 over the lifetime horizon of the model.

Canagliflozin 300 mg has a lower probability of being cost-effective than dapagliflozin in dual therapy with metformin at a WTP threshold of £20,000 per QALY (46.7% probability). It has a 50% probability of being cost-effective compared to dapagliflozin at the £30,000 WTP threshold. Canagliflozin 100 mg has a 52.5% probability of being cost-effective compared to dapagliflozin at both the £20,000 and £30,000 WTP thresholds.

4.4 Summary of uncertainties and issues

The ERG considered whether it is possible to provide estimates of clinical-effectiveness and cost-effectiveness for canagliflozin and sulfonylurea compared to other treatments, given that the manufacturer did not include this intervention in their decision problem. Comparative trials of sulfonylurea in combination with other antidiabetic drugs (e.g. GLP-1-a) with active drug and placebo comparisons have been published.[48] However, an NMA would be problematic as there appears to be only one available RCT of canagliflozin in combination with sulfonylurea (the DIA3008 SU sub-study), and this trial included patients with a higher risk of CVD. This would likely increase heterogeneity in the network (though the manufacturer nonetheless constructed an insulin background network based on the DIA3008 insulin sub-study noting this caveat). The primary outcome assessment time point in the DIA3008 SU sub-study was 18 weeks and therefore would not meet the manufacturer's inclusion criteria for the NMA (trials reporting results at 26 weeks ± 4 weeks), though again the DIA3008 insulin sub-study, which also had a primary outcome assessment time point of 18 weeks, was included in the insulin background network (the MS states that post hoc analyses were performed at the 26 week time point to

support the NMA). Furthermore, incorporating a sulfonylurea background into the economic model would require not only the identification and input of efficacy data appropriate to this background, it would also be necessary to verify the suitability of each of the other model inputs to this background (>800 inputs in total). Whilst, as stated earlier, there may be some patients for whom sulfonylurea, rather than metformin, would be the most appropriate monotherapy and which canagliflozin could be added as a second line treatment, given the limitations noted above, it does not appear to be feasible to estimate clinical-effectiveness and cost-effectiveness for this therapy currently.

Base case results are presented as pairwise treatment comparisons in the MS. Fully incremental analyses were not supplied. The ERG has carried out some incremental analyses but due to time and data constraints these are limited to dual therapy.

Parameter distributions used in PSA are not well documented. The extent to which PSA explores uncertainty in the decision problem is thus not clear. The DSA identifies that the model is particularly sensitive to changes in the value of the HbA_{1c} drift parameter. However the standard deviations used for the drift parameters in PSA are not documented either in the MS or input data spreadsheet.

5 End of life

NICE end of life treatment criteria were not applicable and not included in the MS.

6 Innovation

The manufacturer suggests that canagliflozin is innovative due to its insulin-independent mode of action and limited inherent risk of hypoglycaemia. They also note that canagliflozin is efficacious in meeting the challenges of diabetes management, notably reducing blood pressure, weight and achieving glycaemic control. Furthermore, it is noted that canagliflozin 300mg has demonstrated some degree of local inhibition of intestinal SGLT-1, and this is under further investigation.

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

The manufacturer identified 11 clinical trials of canagliflozin, three of which the ERG considers to be relevant to the decision problem. Two of these evaluated canagliflozin dual therapy (one compared to the SU glimepiride, and the other to the DPP-4-i sitagliptin), and one evaluated canagliflozin as part of triple therapy with MET and SU, compared with the DPP-4-i sitagliptin. All three were non-inferiority trials and were judged to be of to be of generally good methodological quality. Canagliflozin was non-inferior to comparators and superior in terms of reducing HBa_{1c} in the mITT and per protocol analysis. However, per-protocol analysis did not support the finding of superiority. Canagliflozin was generally statistically significantly better than comparators across other outcomes including blood pressure, weight and BMI, and lipids. The ERG clinical advisor commented that the improvements in HBa_{1c} and weight exceed minimally clinically important differences.

The rate of adverse events between canagliflozin and non-canagliflozin was similar but there was a 10% higher rate of patients experiencing adverse events related to the study drug for all canagliflozin. The percentage of patients with serious adverse events and the percentage of deaths was similar between canagliflozin and non-canagliflozin

A number of outcomes were included in the NMA, though it is only HbA_{1c} and BMI from the NMA that appear to be consistently used as input data in the economic model. The probability of canagliflozin ranking as best treatment varied across the networks and by outcome measure. For the outcome of HBa_{1c} canagliflozin ranked highly across the separate networks created for different background therapies. It is worth noting that whilst the NMA appears to have been constructed according using standard methods in a number of cases treatments are linked together by only one trial. Where more than one trial linked treatments there was statistical heterogeneity present. Some networks were small, particularly those used for triple therapy.

The MS only provides evidence for the efficacy of canagliflozin among patients treated on therapy backgrounds that included metformin. This does not reflect all the potential uses of

canagliflozin outlined within its licenced indication, though at present metformin is the most widely used antidiabetic drug in T2DM.

7.2 Summary of cost effectiveness issues

The base case results reported in the MS generally show few incremental QALYs accruing to canagliflozin over a lifetime horizon of 40 years, at small incremental cost. Canaglilozin 100mg generally appears cost-effective in these results, within the WTP thresholds of £20,000 and £30,000 per QALY. Because of its relatively small incremental benefit even slight variation to parameter assumptions can lead to large changes in the ICER against its closest comparators (DPP-4-i and dapagliflozin in dual therapy). Canagliflozin 300 mg is associated with higher ICERs than canagliflozin 100 mg and has a lower probability of being cost-effective than dapagliflozin in dual therapy at a WTP threshold of £20,000 per QALY.

The MS does not report a fully incremental analysis and the place of canagliflozin in the cost-effectiveness hierarchy is not clear in all lines of therapy.

The extent to which PSA has been explored in the model not readily apparent as the parameter distributions which are used in PSA are not well documented.

8 REFERENCES

(1) Canagliflozin Improves Glycaemia in Subjects With Type 2 Diabetes on Metformin Plus Sulphonylurea - Poster presented at CODHy. Congress on Controversies to Consensus in Diabetes, Obesity and Hypertension (CODHy) - 4th World Congress; 12 Nov 8; 2012.

(2) Canagliflozin, a Sodium Glucose Co-Transporter 2 Inhibitor, Improves Glycaemia inSubjects With Type 2 Diabetes Inadequately Controlled With Metformin Plus Sulphonylurea - Poster presented at EASD and ADA. European Association for the Study of Diabetes(48th Annual meeting) and Ammerican Association for Diabetes; 12 Oct; 2012.

(3) Amaris. Systematic literature review and meta-analysis of treatments of type 2 diabetes - Study report. 2013.

(4) The Swedish Insitute for Health Economics (IHE). Systematic Literature Review Of Economic Simulation Models Of T2dm And Cost-Effectiveness Studies Of Canagliflozin, Sglt-2 Inhibitors, And Relevant Treatment Comparators From The Uk Perspective. 2013.

(5) Willis M, Johansen P, Calara PSQ, Asseburg C. ECHO-T2DM 1.4 Technical Report. 10-9-2013. Swedish Institute for Health Economics.

(6) Johansen P, Willis M, Calara PSQ. ECHO-T2DM 1.4 User Guide. 10-7-2013. Swedish Institute for Health Economics.

(7) NICE. The management of type 2 diabetes. 87. 2008.

(8) NICE. Dapagliflozin in combination therapy for treating type 2 diabetes. 288. 2013. Manchester, NICE.

(9) NICE. Extenatide prolonged-release suspension for injection in combination with oral antidiabetic therapy for the treatment of type 2 diabetes. 248. 2012. Manchester, NICE.

(10) NICE. Liraglutide for the treatment of type 2 diabetes mellitus. 203. 2010. London, NICE.

(11) NICE. Continuous subcutaneous insulin infusion for the treatment of diabetes mellitus. 57. 2008. London, NICE.

(12) NICE. Managing type 2 diabetes (diabetes pathway). 2013. Manchester, NICE.

(13) NICE. Blood-glucose-lowering therapy for type 2 diabetes. NICE Pathways . 2013.

(14) MHRA. New advice on risk of bladder cancer with the anti-diabetic drug pioglitazone (Actos, Competact). 2011.

(15) Azoulay L, Yin H, Filion KB, Assayag J, Majdan A, Pollak MN et al. The use of pioglitazone and the risk of bladder cancer in people with type 2 diabetes: nested case-control study. *BMJ* 2012; 344(e3645).

(16) CEGEDIM. Patient Data Report: Diabetes Treatment Combinations - Report AH (JAN2364513_ADH.DN2). 2013.

(17) Janssen. Canagliflozin Summary of Product Characteristics - Draft. 2013.

(18) Janssen Research and Development. A randomized, multicenter, double-blind, parallel, placebo-controlled study of the effects of JNJ 28431754 on cardiovascular outcomes in adult subjects with type 2 diabetes mellitus (Clinical Study report). 2012.

(19) Janssen, Schroeder MD, Girod I. The Cost of Co-morbidities and Complications in Type 2 Diabetes Mellitus. Results of a Systematic Literature Review: Resource identification, measurement and valuation studies. 2013.

(20) Bouma M, Dekker JH, de Sonnaville JJ, van der Does FE, de Vries H, Kriegsman DM et al. How valid is fasting plasma glucose as a parameter of glycemic control in non-insulin-using patients with type 2 diabetes. *Diabetes Care* 1999; 22:904-907.

(21) Cefalu WT, Leiter LA, Yoon KH, Arias P, Niskanen L, Xie J et al. Efficacy and safety of canagliflozin versus glimepiride in patients with type 2 diabetes inadequately controlled with metformin (CANTATA-SU): 52 week results from a randomised, double-blind, phase 3 non-inferiority trial. *Lancet* 2013; 382(9896):941-950.

(22) Lavalle-Gonzalez FJ, Januszewicz A, Davidson J, Tong C, Qiu R, Canovatchel W et al. Efficacy and safety of canagliflozin compared with placebo and sitagliptin in patients with type 2 diabetes on background metformin monotherapy: a randomised trial. *Diabetologia* 2013; 56(12):2582-2592.

(23) Schernthaner G, Gross JL, Rosenstock J, Guarisco M, Fu M, Yee J et al. Canagliflozin compared with sitagliptin for patients with type 2 diabetes who do not have adequate glycemic control with metformin plus sulfonylurea: a 52-week randomized trial. *Diabetes Care* 2013; 36(9):2508-2515.

(24) NICE. NICE Menu of Indicators (NM14 QOF ID: DM007). 2013.

(25) Janssen Research and Development. A randomized, double-blind, 3-arm parallel-group, 2-year (104-week), multicenter study to evaluate the efficacy, safety, and tolerability of JNJ-28431754 100 mg and JNJ-28431754 300 mg compared with glimepiride in the treatment of subjects with type 2 diabetes mellitus not optimally controlled on metformin monotherapy The CANTATA-SU Trial (CANagliflozin Treatment And Trial Analysis Sulfonylurea) 52-week core double-blind period (Clinical Study Report). 2012.

(26) Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011. The Cochrane Collaboration.

(27) Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. 2011. Decision Support Unit, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA.

(28) Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. NICE DSU Technical Support Document 4: inconsistency in networks of evidence based on randomised controlled trials. 2011. Decision Support Unit, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA.

(29) Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomised controlled trials. *Journal of Clinical Epidemiology* 1997; 50(6):683-691.

(30) Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; 313(7052):275-283.

(31) Bagust A, Beale S. Modelling EuroQol health-related utility values for diabetic complications from CODE-2 data. *Health Econ* 2005; 14:217-30.

(32) British Medical Association RPS. British National Formulary No.65. 2013.

(33) Brennan A, Chilcott J. Sheffield Experiences of Patient Level Simulation. 2013. Centre for

Bayesian Statististics in Health Economics, University of Sheffield UK. 29-7-2013.

(34) O'hagan A, Stevenson M, Madan J. Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA. *Health Econ* 2007; 16(10):1009-1023.

(35) Clarke PM, Gray AM, Briggs A, Farmer AJ, Fenn P, Stevens PJ et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). *Diabetologia* 2004; 47:1747-59.

(36) Bagust A, Hopkinson PK, Maier W, Currie CJ. An economic model of the long-term health care burden of Type II diabetes. *Diabetologia* 2001; 44(12):2140-2155.

(37) Schernthaner G, Grimaldi A, Di MU, Drzewoski J, Kempler P, Kvapil M et al. GUIDE study: double-blind comparison of once-daily gliclazide MR and glimepiride in type 2 diabetic patients. *Eur J Clin Invest* 2004; 34(8):535-542.

(38) Amaris. Study report - systematic literature review and meta-analysis of tretments of type 2 diabetes. 2013.

(39) NICE. Guide to the methods of technology appraisal 2013. 2013. NICE. 29-11-2013.

(40) Kahn SE, Haffner SM, Heise MA, Herman WH, Holman RR, Jones NP et al. Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med* 2006; 355(23):2427-2443.

(41) Kahn SE, Haffner SM, Heise MA, Herman WH, Holman RR, Jones NP et al. Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med* 2006; 355(23):2427-2443.

(42) Eastman RC, Javitt JC, Herman WH, Dasbach EJ, Zbrozek AS, Dong F et al. Model of complications of NIDDM. I. Model construction and assumptions. *Diabetes Care* 1997; 20(5):725-734.

(43) Brown JB, Russell A, Chan W, Pedula K, Aickin M. The global diabetes model: user friendly version 3.0. *Diabetes Res Clin Pract* 2000; 50 Suppl 3:S15-S46.

(44) Currie CJ, Morgan CLI, Poole CD, Sharplin P, Lammert M, McEwan P. Multivariate models of healthrelated utility and the fear of hypoglycaemia in people with diabetes. *Current Medical Research and Opinion* 2006; 22:1523-34.

(45) Janssen Cilag. Confidential report: Utilities for Type 2 Diabetes Mellitus and addendum: Utilities for Type 2 Diabetes Mellitus: Regression Analysis Results. 2013.

(46) Department of Health. National Schedule of Reference Costs 2011-12 for NHS trusts and NHS foundation trusts. 2013.

(47) Willis M, Asseburg C, He J. Validation of economic and health outcomes simulation model of type 2 diabetes mellitus (ECHO-T2DM). *J Med Econ* 2013; 16(8):1007-1021.

(48) Shyangdan DS, Royle P, Clar C, Sharma P, Waugh N, Snaith A. Glucagon-like peptide analogues for type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2011;(10):CD006423.