Empagliflozin for the treatment of type 2 diabetes: a single technology assessment

Produced by Warwick Evidence

Authors

Deepson Shyangdan

Ruth Pulikottil Jacob

Martin Connock

Rhona Johnston

Ewen Cummins

Norman Waugh

Correspondence to;

Professor Norman Waugh

Division of Health Sciences

Warwick Medical School norman.waugh@warwick.ac.uk 02476 151585

Date completed: 1st July 2014

Source of funding: this report was commissioned by the NIHR HTA Programme as project number 13/60

Competing interests: none.

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, who commissioned the report.

This report should be referenced as;

Shyangdan D, Jacob R, Connock C, Johnston R, Cummins E, Waugh N. Empagliflozin for the treatment of type 2 diabetes: a single technology appraisal. Warwick Evidence, 2014.

Acknowledgement: We would like to thank Dr Pamela Royle for undertaking additional searches and providing feedback on the report.

Table of Contents

LIST OF TABLES AND FIGURES

Tables

Figures

LIST OF ABBREVIATIONS

ACCORD	Action to Control Cardiovascular Risk in Diabetes
ADDITION	Anglo-Danish_Dutch Study of Intensive Treatment In People with Screen Detected Diabetes in Primary Care
ADVANCE	Action in Diabetes and Vascular Disease: Preterax and Diamicron MR Controlled Evaluation
AE	Adverse event
BG	Blood glucose
BMI	Body mass index
BNF	British National Formulary
CADTH	Canadian Agency for Drugs and Technologies in Health
CEA	Cost effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CG	Clinical guidance
CHF	Congestive heart failure
CHMP	Committee for Medical Products of Human Use
CI	Confidence interval
CrI	Credible interval
DAWN	Diabetes Attitudes, Wishes and Needs
DPP-4	Dipeptidyl peptidase-4
DSU	Decision support unit
ECEM	Empagliflozin Cost Effectiveness Model
EMA	European Medicines Agency
Empa	Empagliflozin
EQ-5D	Euroqol 5 dimensions
ERG	Evidence review group
FAS	Full analysis set
FDA	Food and Drug Administration
FE	Fixed effect
GI	Genital infection
GLP-1	Glucagon-like peptide-1
HbA1c	Glycosylated haemoglobin
HDL	High density lipoprotein
HRQoL	Health related quality of life
Hypo	Hypoglycaemia
ICER	Incremental cost effectiveness ratio
IHD	Ischaemic heart disease
INS	Insulin
IU	International unit
LDL	Low density lipoprotein
LOCF	Last observation carried forward
MDG	Mean daily glucose
MDI	Multiple daily injection
Met	Metformin
MetSU	Metformin plus sulphonylurea
MI	Myocardial infarction
MS	Manufacturer submission
MTA	Multiple technology assessment
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NPH	Neutral Hagedorn Insulin
PG	Plasma glucose

Pio	Pioglitazone
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
PSSRU	Personal Social Services Research Unit
QALY	Quality adjusted life year
QOF	Quality and Outcomes Framework
QoL	Quality of life
RCT	Randomised controlled trial
RE	Random effect
RR	Relative risk
SBP	Systolic blood pressure
SE	Standard error
SGLT2	Sodium glucose co-transporter 2
SLC5A2	Solute carrier family 5 sodium/glucose cotransporter
STA	Single technology appraisal
SU	Sulphonylurea
T2DM	Type 2 diabetes mellitus
TC	Total cholesterol
TIA	Transient ischaemic attack
TTO	Time trade off
TZD	Thiazolidinedione
UKPDS	UK prospective diabetes study
UTI	Urinary tract infection
VAS	Visual analogue scale
VB	Visual basic

1 SUMMARY

Empagliflozin is the third in a new class of drugs for type 2 diabetes to be evaluated by NICE, referred to hereafter as the flozins. The first two, dapagliflozin and canagliflozin, have already been appraised by NICE and recommended for use, subject to certain restrictions. These drugs act by reducing conservation of glucose in the kidney, leading to loss of glucose in the urine, which helps to reduce plasma glucose and also causes a loss of calories.

Scope of manufacturer submission

The industry submission from Boehringer Ingelheim covers use of empagliflozin in;

- dual therapy in people with diabetes that is not sufficiently controlled on metformin
- triple therapy in people whose diabetes is not well controlled on dual therapy with metformin and either a sulphonylurea or pioglitazone
- people whose diabetes is not well-controlled despite therapy with insulin and one or two oral agents

The submission identifies the main comparators as the other flozins, and the dipeptidyl peptidase-4 (DPP-4) inhibitors, referred to hereafter as the gliptins. In the modelling, the gliptin used as comparator is sitagliptin, because that is the one most commonly used in the UK.

The main difference between the scope of the Boehringer submission and the final scope issued by NICE was the omission of any comparison with the glucagon-like peptide-1 (GLP-1) analogue. Boehringer argue, and the ERG agrees, that the GLP-1 analogues belong at a different place in the treatment pathway. The ERG also agrees with the omission by Boehringer of insulin as a comparator, despite its listing as a comparator in the NICE scope.

Clinical evidence submitted by the manufacturer.

The manufacturer submitted data from seven trials and an extension study. The ERG regarded the most important of the trials as being;

- empagliflozin in dual therapy with metformin
- empagliflozin in triple therapy with metformin and a sulphonylurea (unspecified)
- empagliflozin used with regimens containing basal insulin

- empagliflozin used with regimens with multiple daily insulin injections (MDI)

The trials used two doses of empagliflozin, 10 mg and 25 mg daily.
Compared to placebo, empagliflozin reduced HbA1c;

- in dual therapy by 0.57% and 0.64% for 10 mg and 25 mg respectively, at 24 weeks
- in triple therapy, by 0.64% and 0.59% at 24 weeks
- in basal insulin regimens, by 0.46% and 0.62% at 78 weeks
- in MDI insulin regimens, by 0.38% and 0.46% at 52 weeks

In a trial in patients with renal impairment, HbA1c was reduced by 0.52% and 0.68% in those with mild renal impairment, and by 0.42% by the 25 mg dose in those with moderate renal impairment.

Empagliflozin was also associated with weight loss (via losing glucose and hence calories in the urine);

- in dual therapy, by 1.6 kg for the 10mg dose and 2.0 kg for the 25 mg dose, compared to placebo
- in triple therapy, by 1.8 and 2.0 kg
- in basal insulin regimens, by 3.6 kg and 3.1 kg (i.e. those on the larger dose lost less)
- with MDI insulin, by 1.5 and 1.6 kg

There were reductions in systolic blood pressure (SBP) ranging from 1.4mm Hg in the MDI trial, to 4.8 mg Hg in the dual therapy trial. In a short-duration (12 weeks) trial in patients with hypertension, SBP reductions were 3.4 and 4.2mm Hg for the two doses.
There were only small differences between the two doses of empagliflozin, with the lower dose sometimes reported to have greater effects.

Because there were no head to head trials of empagliflozin against the gliptins or other flozins, the submission provided data for modelling from a series of network meta-analyses. In brief, the results showed roughly equal effectiveness in glycaemic control amongst the flozins and the gliptins.

ERG Commentary

The main weakness in the evidence base was that all but one of the trials compared empagliflozin with placebo rather than active comparators such as a gliptin. One trial

compared empagliflozin with glimepiride, a sulphonylurea, in dual therapy with metformin, but the ERG considered that this was less relevant because sulphonylureas should be a precursor to flozins, given the very low cost of the former. In the trial against glimepiride there was little difference in HbA1c (0.1%) but those on the sulphonylurea gained weight where those on the flozin lost weight, giving a difference of 4.5 kg at 2 years.

The ERG had some concerns with the network meta-analysis (NMA) in which some errors were detected, but correcting these made little difference – no difference in HbA1c results and only slight differences in hypoglycaemic episodes. The ERG therefore agrees with Boehringer that empagliflozin is comparable in clinical effectiveness with canagliflozin, dapagliflozin and sitagliptin.
No data on lipid changes were included in the clinical effectiveness submission.

The main adverse effects were urinary tract infections (UTIs) and genital infections, both seen mainly in women (women with UTIs about 12% on empagliflozin versus 8% on placebo). Hypoglycaemia was reported infrequently, and the definition used was $\leq$ 3.9mml/l which includes some of the normal range for plasma glucose. The ERG thinks it would be reasonable to say that empagliflozin does not cause hypoglycaemia.

The ERG had access to an independent academic NMA which confirmed that the three flozins were similar in effectiveness.

Economic model used by Boehringer: ERG critique
The model submitted was the Empagliflozin Cost Effectiveness Model (ECEM) written in visual basic, which is not on the NICE approved software list. As far as we know, this model has never been used in any previous NICE appraisals.

The ERG has cross checked a number of elements of the visual basic (VB) implementation of the ECEM. This has identified what may be a number of serious issues: random sampling at the patient level, modelling of the evolution of the risk factors, model convergence, model sensitivity to the random seeds chosen, questionable handling of the application of quality of life values to weight changes and a possible halving of the quality adjusted life year (QALY) decrements associated with adverse events and the complications of diabetes. If the manufacturer confirms that many of these are indeed errors, it will largely invalidate the submitted results. The ECEM has also been constructed so that it can only simulate 100 individual patients if 300 probabilistic sensitivity analysis (PSA) iterations are being

conducted. These are unusually low numbers and may limit the ability of the ECEM to reliably discriminate between the overall impacts of different therapies.

Due to the extent and complexity of the coding of this new model, the ERG has not had time to parse all of VB code and there may be other problems not detected. It appears that there may have been a lack of validation and stress testing of the model, which may call into question the robustness and reliability of the remaining code.

These problems have implications for both the economic modelling included in the Boehringer submission, and for ERG analyses. If the problems are confirmed, neither can be regarded as reliable.

Summary of cost-effectiveness evidence submitted by the manufacturer

The Boehringer submission compares the cost-effectiveness of empagliflozin with sitagliptin, canagliflozin and dapagliflozin, in dual and triple oral therapy, and in insulin-containing regimens. Data on clinical effectiveness was taken from the NMA. Both doses of empagliflozin, canagliflozin 100mg daily, and dapagliflozin were costed at £477 per annum, with sitagliptin at £433 and canagliflozin 300 mg daily at £608. Modelling involved treatment effects for the evolutions of HbA1c, systolic blood pressure and weight.

Because the clinical effectiveness data from the NMA showed similar clinical effectiveness, differences in QALYs gained were very small, and often too small to matter. A QALY difference of 0.01 represents 3.65 days of perfect health. Cost differences over the 40 years modelled were also usually small, for example a few hundred pounds. Incremental cost-effectiveness analyses (ICERs) were therefore subject to considerable uncertainty and empagliflozin fluctuated from being dominated by sitagliptin to being dominant over sitagliptin.

ERG commentary on cost-effectiveness analysis

Serious problems with the model raise doubts about the estimates of cost-effectiveness and the uncertainty surrounding them. This does not necessarily mean that the conclusion of equivalence (based on clinical trial data and the NMA, and similar pricing) is incorrect, but the model is incapable of showing this in a robust way. For example, there is an error in converting utility per BMI point change into utility per 1 kg change, and this affects the estimation of the effect of weight change in the model.

The Boehringer submission states that;

"*The overall differences in QALYs and costs were marginal in all analyses and no treatment was clearly the optimum choice."*

The ERG agrees with this summary. A few changes were made by the ERG and some model re-runs carried out, but differences were unimportant, with differences in QALYs ranging from 0.001 to 0.019.

Conclusions

The evidence from the trials of empagliflozin show that it is clinically effective in improving glycaemic control, though not dramatically so, with mean reductions in HbA1c ranging from 0.38 to 0.64%, when 0.5% is usually regarded as clinically meaningful. Empagliflozin also provides modest reductions in blood pressure and weight. Its clinical effectiveness is similar to other drugs already approved, canagliflozin, dapagliflozin, and sitagliptin.

Costs are similar, except for the higher dose of canagliflozin. Given that and the similar clinical effects, and despite concerns with model and modelling, the ERG expects empagliflozin to be as cost-effective as the comparators.

2 BACKGROUND

Diabetes (T2DM) affects more than 3.4 million people in England, with a prevalence of 7.9%. The prevalence in Wales is higher at about 9%, with about 219,000 people affected.[1] About 90% of these people with have type 2 diabetes. The prevalence of type 2 diabetes has been increasing, partly due to demographic change, partly due to better detection, but mainly due to increased prevalence of overweight and obesity. Diabetes is increasingly costly to the National Health Service (NHS), with a recent study estimating that 10% of all NHS expenditure is on diabetes.[2]

The guidelines on the management of T2DM from the UK's National Institute for Health and Care Excellence (NICE), recommend that if lifestyle intervention is insufficient, the first line of drug treatment is metformin, followed by a sulphonylurea (SU), or sometimes pioglitazone, before commencing on insulin.[3] However sulphonylureas, pioglitazones and insulin all cause weight gain which may worsen insulin resistance. Sulphonylureas and insulin can also cause hypoglycaemia. Pioglitazone, now the only glitazone left in use in the UK, can cause oedema, heart failure and fractures, and there is increasing concern about whether its use is associated with bladder cancer. Pioglitazone use has now been discontinued in France.[4]

The NICE Clinical Guideline 87[3] on T2DM contains a flowchart reproduced here (Figure 1).

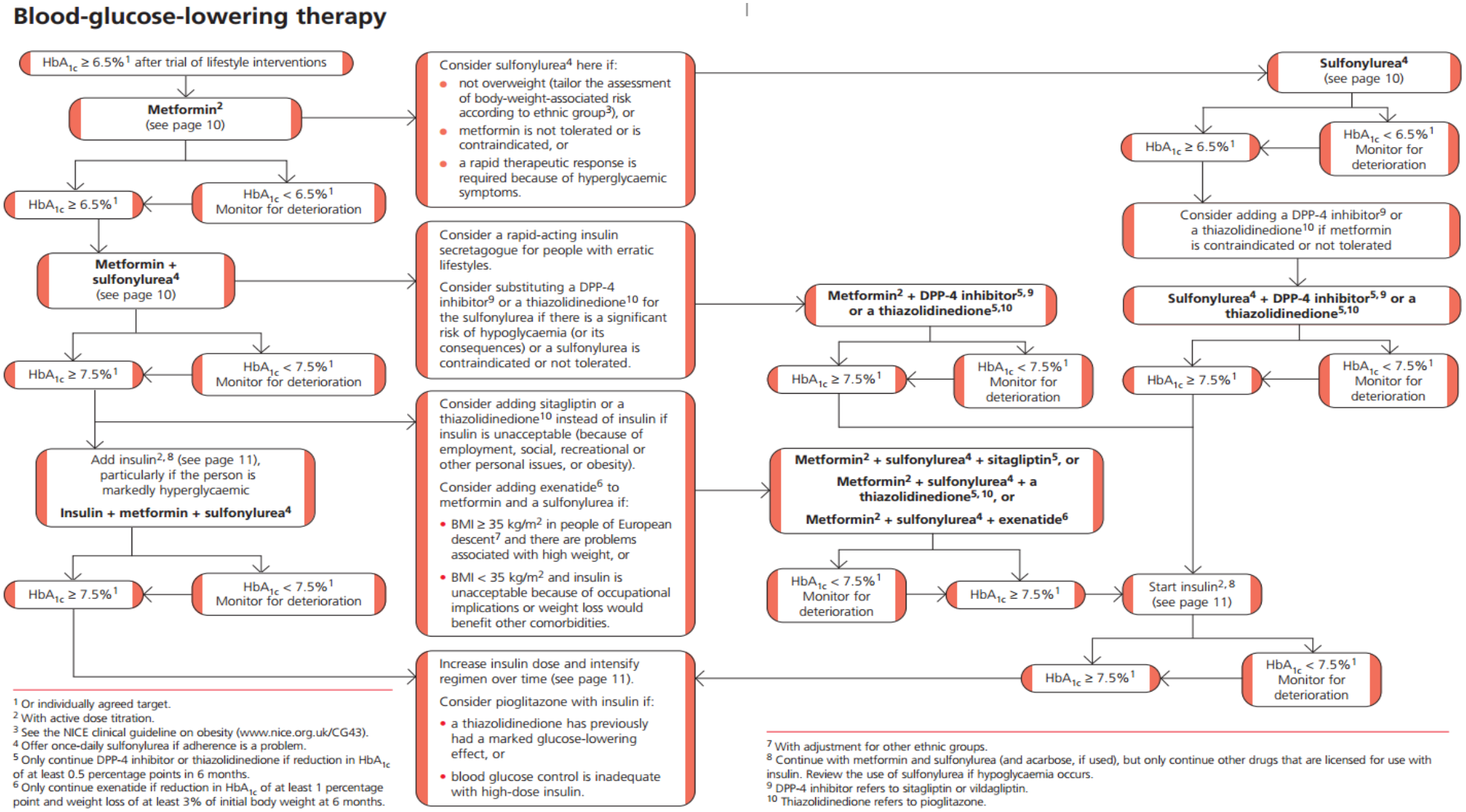


Figure 1. Flow diagram of blood-glucose-lowering treatments for the management of type 2 diabetes (source: NICE Clinical Guidelines 87)

We now have eight classes of glucose-lowering drugs for T2DM, though some contain only a single drug:

- Biguanides: metformin
- SUs: gliclazide, glimepiride and glipizide
- Thiazolidinediones (TZDs): pioglitazone
- Acarbose
- Meglitinides: nateglinide and repaglinide
- The glucagon like peptide-1 (GLP-1) analogues: exenatide (now with a once a week form) and liraglutide (once daily)
- The dipeptidyl peptidase-4 (DPP-4) inhibitors, also known as the 'gliptins'
- Insulins. In T2DM, insulin treatment starts with a once daily basal insulin (NICE recommends NPH as first choice) but if intensification is needed, short-acting insulins may be added at mealtimes, or twice daily biphasic insulin may be used.

However, there is still a need for drugs that that will lower glucose without causing hypoglycaemia or weight gain, and that can improve cardiovascular outcomes.

Empagliflozin is one of the newest class, the sodium glucose co-transporter 2 (SGLT2) receptor inhibitors, hereafter referred to as the flozins. Glucose is allowed through the filter in the renal glomeruli but is reabsorbed in the renal tubules. Glycosuria (glucose in the urine) occurs when the renal threshold for glucose (blood glucose of approximately 10 mmol/l) has been reached. At this threshold the kidney cannot reabsorb all of the filtered glucose. 90% of the urinary glucose is transported across the membrane of the proximal tubule by sodium glucose co-transporter 2 (SGLT2).[5] The SGLT2 protein in humans is encoded by the gene solute carrier family 5 sodium/glucose cotransporter (SLC5A2). Some people have a mutation in the SLC5A2 gene that causes a defective SGLT2 protein, resulting in glycosuria. Individuals who have this mutation do not have significant problems related to the glycosuria, such as urinary tract infections (UTIs). [6] This implies that blocking the transport mechanism should not cause problems.

The SGLT2 inhibitors block the transport system and so mimic the effect of the SLC5A2 mutation and reduce the reabsorption of renal filtered glucose back into the bloodstream, thereby reducing hyperglycaemia, without the side-effects of weight gain or hypoglycaemia.[7]

There is also a SGLT1 transport mechanism, which is present both in the kidney and the gut. In the kidney, it is much less important than SGLT2. Inhibition of gut SGLT1 reduces absorption of glucose

there, and it has been suggested, albeit by a group linked to Janssen who market it, that canagliflozin may have a dual action.[8]

Dapagliflozin has already been approved by NICE.[9] The guidance is reproduced in Box 1. Dapagliflozin is a highly selective inhibitor of SGLT2 and has little effect on SGLT1.

Box 1. NICE guidance on dapagliflozin for the treatment of type 2 diabetes mellitus

1 Guidance on dapagliflozin

1.1 Dapagliflozin in a dual therapy regimen in combination with metformin is recommended as an option for treating type 2 diabetes, only if it is used as described for dipeptidyl peptidase-4 (DPP-4) inhibitors in Type 2 diabetes: the management of type 2 diabetes (NICE clinical guideline 87).

1.2 Dapagliflozin in combination with insulin with or without other antidiabetic drugs is recommended as an option for treating type 2 diabetes.

1.3 Dapagliflozin in a triple therapy regimen in combination with metformin and a sulfonylurea is not recommended for treating type 2 diabetes, except as part of a clinical trial.

Canagliflozin has also been approved by NICE.[10] Canagliflozin inhibits both SGLT2 and SGLT1. The guidance is shown in Box 2.

Box 2. NICE guidance on canagliflozin for the treatment of type 2 diabetes mellitus

1. Guidance on canagliflozin

1.1 Canagliflozin in a dual therapy regimen in combination with metformin is recommended as an option for treating type 2 diabetes, only if:

- a sulfonylurea is contraindicated or not tolerated or
- the person is at significant risk of hypoglycaemia or its consequences

1.2 Canagliflozin in a triple therapy regimen is recommended as an option for treating type 2 diabetes in combination with:

- metformin and a sulfonylurea
- metformin and a thiazolidinedione

1.3 Canagliflozin in combination with insulin with or without other antidiabetic drugs is recommended as an option for treating type 2 diabetes.

The ERG prefer to use the terms “dual therapy” and “triple therapy” to “second-line” and “third-line” because the latter terms could cover substitution as well as addition.

The difference in the guidance on use as triple therapy is because at the time of the dapagliflozin appraisal, evidence on its use in triple therapy was not available. These drugs act through a mechanism that is not dependent on insulin secretion and so may be effective when other drugs that depend entirely sulphonylureas or in part (gliptins and GLP-1 analogues) on stimulating insulin release have lost effectiveness. In type 2 diabetes, the capacity of the pancreatic beta cells to produce insulin often falls over time.

The dapagliflozin and canagliflozin guidance differ also in use in moderate renal impairment. The guidance on dapagliflozin says that it should not be used in patients with GFRs below 60 ml/min, whereas the guidance on canagliflozin says that if it was started before renal function declined to a eGFR of 60 ml/min, it may be continued till eGFR falls below 45 ml/min.

Since there are existing drugs which are inexpensive and with a long safety record, it is unlikely that SGLT2 inhibitors would be used first line, and we therefore see their role as second or third drugs used in combination therapy in T2DM.

There are two main issues for this appraisal:

i) The first question is whether empagliflozin is clinically effective in improving glycaemic control in T2DM, with an acceptable adverse event profile;
ii) The second question is about whether it is cost-effective.

One issue that arises is where the SGLT2 inhibitors fit into the therapeutic pathway. Factors to be considered include:

- Effect on glycaemic control as reflected in HbA1c reductions
- Effect on weight, compared to other drugs, some of which cause marked weight gain
- Effect on cardiovascular risk, including on blood pressure and lipid levels
- Adverse effects, particularly increased genital and urinary infections
- Duration of diabetes. In long-standing T2DM, the efficacy of the flozins will not be affected by a fall in endogenous insulin production
- Interactions with other drugs, especially in patients on treatment for co-morbidities
- Ease of use, by oral administration rather than injection
- Cost

Figure 2 shows the costs of drug therapies for T2DM.

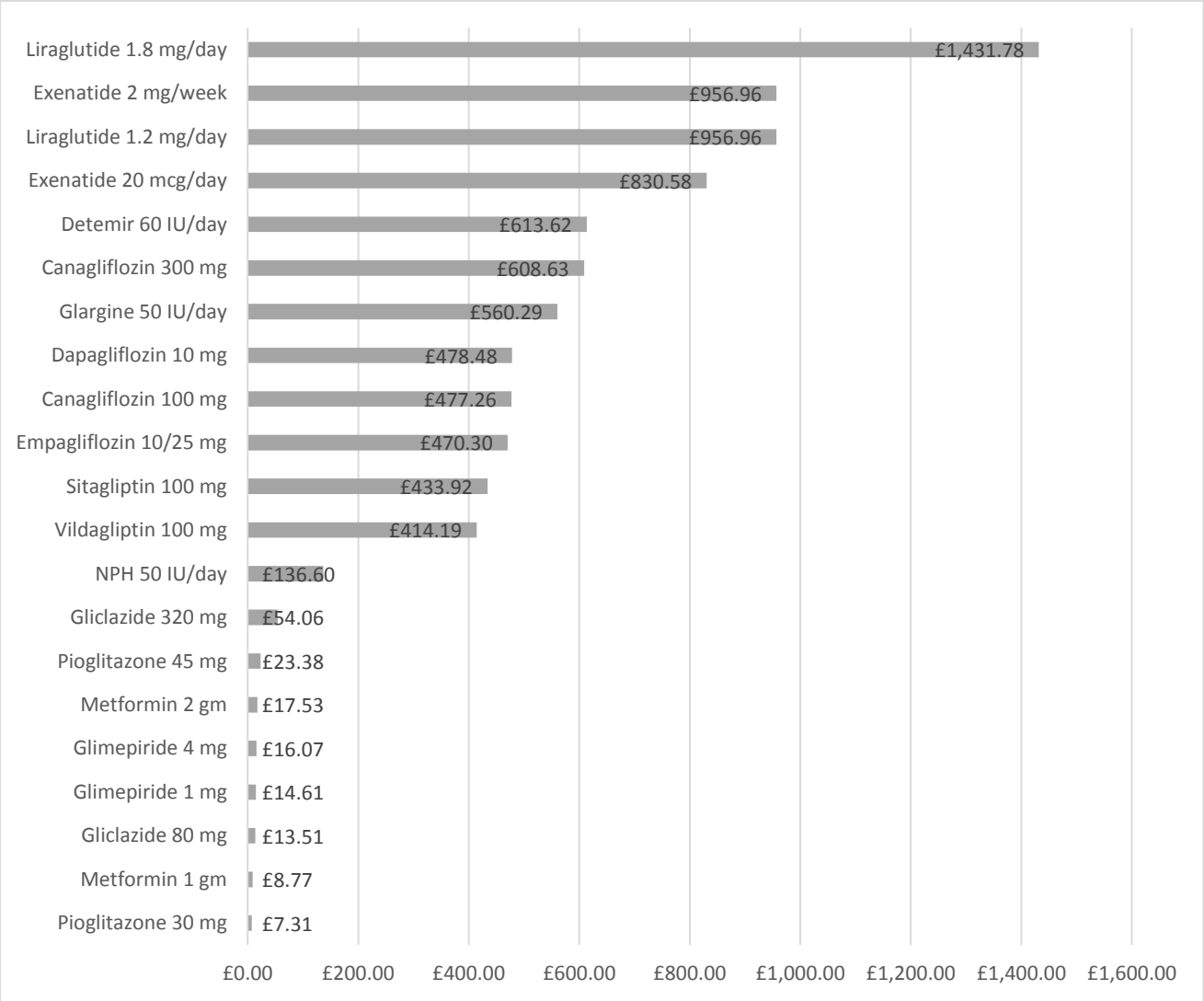


Figure 2. Costs of different pharmacological interventions for diabetes

Source: British National Formulary[11]; Manufacturer submission/ERG report of Canagliflozin[12]; Manufacturer submission of Empagliflozin

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer description of the underlying health problem (T2DM) in terms of prevalence, relevant symptoms, complications and required treatments is generally accurate. The quoted proportion of people with undiagnosed type 2 diabetes is over-estimated at 50%. It is now probably under 20% in most areas. Screening people at high risk of diabetes in the Cambridge centre of the ADDITION trial gave an added 0.64% of newly-diagnosed people to the previous 3.1% of the population with diagnosed diabetes, suggesting an undiagnosed proportion of 17%.[13]

2.2 Critique of manufacturer's overview of current service provision

The manufacturer correctly summarises variations in current care, quoting the two National Diabetes Audits, and noting that the proportions of people with diabetes meeting target levels for glycaemic control varies. (Section 2.6).

In Section 2.4, the Boehringer submission mentions relevant NICE guidances, including CG 87, clinical guidelines for type 2 diabetes.

Two key recommendations from that guideline are not mentioned in the manufacturer's introduction, but should be borne in mind. The first is that the target for glycaemic control was set at an HbA1c level of 6.5%. It may be that this will be relaxed in the next update of the guidelines, following controversy over intensified control in type 2 diabetes, in the wake of the ADVANCE and ACCORD trials. The target level might be raised, to perhaps an HbA1c of 7.0% or less overall, though targets need to be individualised.

The second recommendation was that the HbA1c level at which treatment should be intensified, known as the "switching point" was set at 7.5%, higher than the target for good glycaemic control. So if a patient does not achieve an HbA1c less than 7.5%, the NICE guideline recommends that intensification should be considered. One implication of that recommendation is that what matters in assessing the effectiveness of a drug in type 2 diabetes will not be just the mean reduction in HbA1c, but whether it reduces HbA1c to under 7.5%.

Intensification often involves insulin. However, there has been reluctance to start insulin treatment, documented in reports for past NICE appraisals, and summarised in the assessment report for the CG87 guidelines group.[14] In brief;

- Both patients and their doctors have been reluctant to start insulin, as documented in the DAWN study[15]
- This is partly because most patients with T2DM who are on insulin, do not achieve good control[16]
- Many patients therefore remain poorly controlled on combination oral agents for years before stating insulin[17, 18]

The manufacturer states (section 2.2) that 45% of the 1.6 million people with diabetes in England and Wales on drug treatment are on combination therapy. From this, they conclude that 720,000 patients

might be relevant to this appraisal. However this may be an over-estimate, because many will have adequate glycaemic control on existing treatment.

The manufacturer (section 2.3) correctly notes that people with type 2 diabetes have poorer life expectancy than people without diabetes. This is based on data from the UK Prospective Diabetes Study, and may now be a little out of date because of wider use of statins to reduce blood cholesterol, and tighter control of blood pressure, incentivised by QOF targets and payments.

When considering comparators (pages 36 and 43), the submission is too dismissive of the TZD drugs; *"Pioglitazone was considered not relevant since TZDs are currently very rarely used and their use is falling".*

This is not correct, and the manufacturer's own submission in Table 107 (page 337) gives figures for the use of TZDs in combination therapy (not including with insulin) of almost 60,000 people. Rosiglitazone is no longer used in the UK (though in the USA, the FDA is reconsidering its use). However pioglitazone is still used, despite awareness of adverse effects including oedema (including macular oedema), heart failure and fractures, and concern about an increase in bladder cancer. A review by the European Medicines Agency concluded that pioglitazone was associated with an increase in the small risk of bladder cancer, from 7 in 10,000 in people with diabetes not treated with pioglitazone to 15 per 10000. In view of the small absolute risk, the EMA therefore concluded that pioglitazone should retain its licence, but should be used with caution.[19] This advice was welcomed by the Association of British Clinical Diabetologists on the grounds that some patients respond well to pioglitazone, and that the alternative would be to use newer agents without long-term safety data.[20] Boehringer are correct to say that use of the TZDs has fallen, and is probably still falling, but in 2012/13 pioglitazone was still used almost as commonly as sitagliptin, so it is incorrect to say that TZDs are "very rarely " used.[21] Pioglitazone is now available in inexpensive generic form.

DECISION PROBLEM

The NICE guideline recommends starting with diet and lifestyle, adding metformin if control is inadequate, and next adding a sulphonylurea. There is an option in the current guideline to use pioglitazone as an alternative to a sulphonylurea.

Hence in dual therapy, if sulphonylureas or metformin cannot be tolerated, we would expect a gliptin as an oral alternative to be tried if patients could not tolerate either metformin or a sulphonylurea.

The gliptins therefore seem to be the key comparator for the flozins in dual therapy.

In triple therapy, comparators include the gliptins, a GLP-1 analogue (probably now once-weekly exenatide) or insulin. We would expect the gliptins to be tried before long-acting exenatide on grounds of cost and the need to inject exenatide. So in triple therapy, the main comparators are again the gliptins. It could be argued that insulin with once daily NPH would cost less, but as noted in the previous chapter, there tends to be resistance to starting insulin because of its adverse effects of weight gain and hypoglycaemia, and because insulin often fails to ensure good control unless intensified. Intensive life style interventions have been shown to be as good as insulin in one small Danish study[22] but that needs to be confirmed by further research.

The combination of insulin and a GLP-1 analogue was unlicensed but widely used, as a logical combination. Twice daily exenatide has now been licensed for use in combination with insulin.[23, 24]

The NICE scope did not mention acarbose, nor the meglitinide analogues, repaglinide and nateglinide. The latter are insulin secretagogues, shorter acting but less potent than the SUs.[25] None of these drugs are widely used in the UK, and their effectiveness in triple therapy is limited.[26]

In conclusion, the ERG regards the gliptins (DPP-4 inhibitors) as the key comparators for the flozins, and the place of flozins to be mainly in triple oral therapy, and as an add-on to insulin.

The ERG therefore agrees with the comment in the Boehringer submission (p 36), that;
"The only remaining treatments in the decision space are dapagliflozin, canagliflozin and the DPP-4 inhibitors".

3 CLINICAL EFFECTIVENESS

The ERG does not consider that any important trials of empagliflozin, dapagliflozin or canagliflozin in type 2 diabetes mellitus have been omitted.

The manufacturer included evidence on empagliflozin from eight trials, and one extension study that recruited patients from three trials. All but two trials were against placebo. Of the trials against active comparators, one was against sitagliptin in monotherapy (so not used in the submission, except as part of the extension study), and the other was against glimepiride in dual therapy with metformin, which we regard as being of less relevance because SUs would be a precursor not a comparator.
The trials are listed by numbers such as 1245.23, but also have names such as EMPA-REG MET. For convenience we will refer to them by abbreviations such as ER Met.

Four trials of empagliflozin are most relevant to the decision problem, taking the sub-studies of the 1245.23 randomised controlled trial (RCT) as two separate trials. One of the sub-studies of 1245.23, ER Met, assessed efficacy of empagliflozin in dual therapy with metformin and the other, ER MetSU examined the efficacy in triple therapy with metformin plus sulphonylurea (SU). The two other studies were in combination with insulin regimens, one on basal insulin alone (ER Basal), the other with multiple daily injections (ER MDI), with or without concomitant metformin and/or SU. For convenience, the relevant section of Table 8 has been reproduced below (Table 1).

The trial, 1245.23, assessing efficacy of empagliflozin 10 mg or 25 mg in patients inadequately controlled with metformin (ER Met) or metformin plus SU (ER MetSU) is published in full.[27, 28]. While preparing this report, the study assessing efficacy and safety of empagliflozin added to MDI was published.[29] The manufacturer has also submitted results from an extension study (l 1245.31 – ER EXTEND), which also included patients from trials not relevant to the decision problem. Hence, we have only summarised results of patients completing ER Met and ER MetSU who were followed up for another 52 weeks.

Those trials not listed in the table below will be described in brief.
The EMPA-REG BP trial is not described in detail here because it was in monotherapy in people with diabetes who had never had glucose-lowering agents. In this trial, 1830 people with hypertension (mean baseline BP 142/84) in 121 centres were randomised to empagliflozin 10 or 25 mg, or placebo. SBP fell by means of 2.95 mmHg on 10mg and 3.68 mmHg, but rose by 0.48 mmHg on placebo. HbA1c fell by 0.59 on 10mg and by 0.62 on 25mg.

The EMPA- REG Renal trial was in people with renal impairment, classed as having estimated GFRs as follow;

- Mild renal impairment – GRF 60-89
- Moderate GFR 30-59
- Severe GFR <30.

Empagliflozin was given in addition to background treatments which included metformin, pioglitazone and insulin, but no details are given of how many were on monotherapy, dual or insulin regimens. The mild group were given empagliflozin 10mg or 25mg, the moderate and severe groups only 25mg, all versus placebo. HbA1c fell in the moderate group by 0.46% on 10mg, by 0.63% on 25mg, and rose slightly on placebo. In the moderate renal impairment group, HbA1c fell by 0.37 on 25mg and rose by 0.05 on placebo. No results were provide for the severe group.
The EMA has said that empagliflozin should not be started once GFR drops below 60ml/min, but that if patients have started on it before reaching that threshold, they may continue down to 45, as follows;

EMA recommendations for patients with renal impairment
'Due to the mechanism of action, the efficacy of empagliflozin is dependent on renal function. No dose adjustment is required for patients with an eGFR ≥60 ml/min/1.73 m2 or CrCl ≥60 ml/min. Empagliflozin should not be initiated in patients with an eGFR <60 ml/min/1.73 m2 or CrCl <60 ml/min. In patients tolerating empagliflozin whose eGFR falls persistently below 60 ml/min/1.73 m2 or CrCl below 60 ml/min, the dose of empagliflozin should be adjusted to or maintained at 10 mg once daily. Empagliflozin should be discontinued when eGFR is persistently below 45 ml/min/1.73 m2 or CrCl persistently below 45 ml/min (see sections 4.4, 4.8, 5.1, and 5.2). Empagliflozin should not be used in patients with end stage renal disease (ESRD) or in patients on dialysis as it is not expected to be effective in these patients (see sections 4.4 and 5.2).'

This is similar to the NICE guidance on canagliflozin, whereas the guidance on dapagliflozin says there should be no use under 60ml/minute.

Table 1. List of relevant RCTs

Trial no. (acronym)	Intervention	Comparator	Population[†]	Primary study ref.
1245.23 (EMPA-REG MET, EMPA-REG METSU; NCT01159600) **ERG abbreviations** **ER Met and ER MetSU**	Empagliflozin 10mg or 25mg once daily add-on to metformin or metformin plus SU	Placebo	**All studies:** • T2DM • ≥18 years (and ≤65 years in India) • BMI ≤45kg/m^2 • Diet and exercise programme • Glucose level ≤13.3 mmol/L • eGFR ≥30 mL/min/1.73 m^2 **Metformin only sub-study:** • N=638 • HbA1C ≥7.0% to ≤10% **Open-label metformin only sub-study:** • N=69 • HbA1C >10.0% **Metformin plus SU sub-study:** • N=669 • HbA1C ≥7.0% to ≤10% **Open-label met + SU sub-study**: • N=103 • HbA1C > 10.0%	Haring *et al.* 2013[27] Haring et al 2014[28] Data on file (clinical study report 1245.23)
1245.33 (EMPA-REG BASAL; NCT01011868) **ER Basal**	Empagliflozin 10 mg or 25 mg once daily add-on to basal insulin (glargine or detemir insulin [≥20 IU/day] or NPH insulin [≥14 IU/day] with or without concomitant metformin and/or SU	Placebo	• N=494 • T2DM • ≥18years • HbA1C >7.0% to ≤10.0% • BMI ≤45 kg/m2 • Stable insulin dose ≥12 weeks prior to randomisation	Data on file (clinical trial report 1245.33)
1245.49 (EMPA-REG MDI; NCT01370005) **ER MDI**	Empagliflozin 10 mg or 25 mg once daily add-on to multiple daily injections of insulin (total insulin >60IU/day) either alone or with metformin	Placebo	• N=566 • T2DM • ≥18 years • HbA1C >7.5% to ≤10.0% • BMI 30-45kg/m^2 • Stable insulin dose ≥12 weeks prior to randomisation	Rosenstock et al 2014[29]

3.1 Summary and critique of submitted clinical effectiveness evidence

3.1.1 Quality of included RCTs

The manufacturer presented quality assessment results in Table 15 (pages 105 and 106) of the submission. The ERG has used the Cochrane risk of bias tool[30] to assess the quality of the included studies, and considers all trials to be of good quality (Appendix 1).

The primary analyses were undertaken on the full analysis set (FAS) which included all patients who were randomised, treated with ≥ 1 dose of trial medication, and who had a baseline assessment. Missing values were imputed using a last observation carried forward (LOCF) method.

3.1.2 Overview of included RCTs

The overview of the included trials is given in section 6 of the MS.

The four key trials were reported to be randomised, double-blind, parallel-group, placebo-controlled trials. From table 11, the section on the four most relevant trials have been reproduced below as Table 2. Note that the open-label empagliflozin sub-study (details in the right-hand column) was not an RCT, being in patients whose baseline HbA1c was > 10%, in whom it was presumably felt inappropriate to use placebo. We include it for interest to show the size of reduction in HbA1c in this group.

Table 2. Characteristics of participants in the RCTs across randomised groups (FAS)

Baseline characteristic	**Treatment group**			
ER Met (N=706)	**Placebo (n=207)**	**Empagliflozin 10mg (n=217)**	**Empagliflozin 25mg (n=213)**	**Open-label empagliflozin 25mg (n=69)**
Demographic data				
Age, mean (SD) [years]	56.0 (9.7)	55.5 (9.9)	55.6 (10.2)	49.8 (11.5)
Gender, N (%)				
Male	116 (56.0)	125 (57.6)	120 (56.3)	41 (59.4)
Female	91 (44.0)	92 (42.4)	93 (43.7)	28 (40.6)
Race, N (%)				
American Indian/Alaska native	0	2 (0.9)	2 (0.9)	2 (2.9)
Asian	92 (44.4)	99 (45.6)	98 (46.0)	14 (20.3)
Black/African American	2 (1.0)	4 (1.8)	0	4 (5.8)
White	113 (54.6)	112 (51.6)	113 (53.1)	49 (71.0)
eGFR (MDRD), mean (SD) [mL/min/1.73m^2]	89.7 (21.4)	89.5 (19.6)	87.7 (19.3)	95.5 (20.7)
Baseline variables				
HbA_{1C}, mean (SD) [%]	7.90 (0.88)	7.94 (0.79)	7.86 (0.87)	11.7 (1.29)
Time since diagnosis of T2DM, N (%)				
≤1year	19 (9.2)	20 (9.2)	19 (8.9)	5 (7.2)
>1 to 5years	83 (40.1)	78 (35.9)	69 (32.4)	28 (40.6)
>5 to 10 years	65 (31.4)	68 (31.3)	74 (34.7)	19 (27.5)
>10 years	40 (19.3)	51 (23.5)	51 (23.9)	17 (24.6)
Weight, mean (SD) [kg]	79.73 (18.57)	81.59 (18.51)	82.21 (19.29)	85.07 (21.96)
BMI, mean (SD), [kg/m^2]	28.70 (5.22)	29.12 (5.48)	29.72 (5.72)	30.37(5.51)
SBP, mean (SD) [mmHg]	128.6 (14.7)	129.6 (14.1)	130.0 (15.1)	126.2 (11.4)
ER MetSU (N=767)	**Placebo (n=225)**	**Empagliflozin 10mg (n=225)**	**Empagliflozin 25mg (n=216)**	**Open-label empagliflozin 25mg (n=101)**
Demographic data				
Age, mean (SD) [years]	56.9 (9.2)	57.0 (9.2)	57.4 (9.3)	53.4 (10.5)
Gender, N (%)				
Male	112 (49.8)	113 (50.2)	114 (52.8)	54 (53.5)
Female	113 (50.2)	112 (49.8)	102 (47.2)	47 (46.5)
Race, N (%)				
American Indian/Alaska native	3 (1.3)	4 (1.8)	3 (1.4)	2 (2.0)

Baseline characteristic	**Treatment group**			
Asian	127 (56.4)	129 (57.3)	125 (57.9)	48 (47.5)
Black/African American	7 (3.1)	3 (1.3)	3 (1.4)	1 (1.0)
White	88 (39.1)	89 (39.6)	85 (39.4)	50 (49.5)
eGFR (MDRD), mean (SD) [mL/min/1.73m^2]	86.9 (20.1)	86.5 (21.8)	88.3 (22.6)	93.1 (23.7)
Baseline variables				
HbA_{1C}, mean (SD) [%]	8.15 (0.83)	8.07 (0.81)	8.10 (0.83)	11.18 (1.25)
Time since diagnosis of T2DM, N (%)				
≤1 year	2 (0.9)	3 (1.3)	7 (3.2)	4 (4.0)
>1 to 5 years	36 (16.0)	59 (26.2)	43 (19.9)	26 (25.7)
>5 to 10 years	94 (41.8)	74 (32.9)	79 (36.6)	33 (32.7)
>10 years	93 (41.3)	89 (39.6)	87 (40.3)	38 (37.6)
Weight, mean (SD) [kg]	76.23 (16.88)	77.08 (18.34)	77.50 (18.81)	76.93 (18.00)
BMI, mean (SD), [kg/m^2]	27.90 (4.93)	28.32 (5.43)	28.32 (5.45)	28.70 (5.49)
SBP, mean (SD) [mmHg]	128.8 (14.3)	128.7 (13.9)	129.3 (14.2)	126.4 (12.4)
ER MDI (N=563)	**Placebo (n =188)**	**Empagliflozin 10mg (n=186)**	**Empagliflozin 25mg (n=189)**	
Demographic data				
Age, mean (SD) [years]	58.1 (9.4)	58.6 (9.8)	59.9 (10.5)	
Gender, N (%)				
Male	90 (52.9)	93 (55.0)	93 (60.0)	
Female	80 (47.1)	76 (45.0)	62 (40.0)	
Race, N (%)				
Asian	33 (19.4)	37 (21.9)	28 (18.1)	
Black/African American	21 (12.4)	12 (7.1)	15 (9.7)	
Other	3 (1.8)	1 (0.6)	1 (0.6)	
White	113 (66.5)	119 (70.4)	111 (71.6)	
eGFR (MDRD), mean (SD) [mL/min/1.73m^2]	83.89 (22.73)	85.01 (23.63)	82.88 (25.46)	
Baseline variables				
HbA_{1C}, mean (SD) [%]	8.18 (0.79)	8.27 (0.83)	8.27 (0.84)	
Time since diagnosis of diabetes, N (%)				
≤1 year	4 (2.4)	0	1 (0.6)	
>1 to 5 years	20 (11.8)	15 (8.9)	12 (7.7)	
>5 years	146 (85.9)	154 (91.1)	142 (91.6)	
Weight, mean (SD) [kg]	90.46 (22.47)	91.59 (20.05)	94.71 (20.70)	
BMI, mean (SD), [kg/m^2]	31.75 (5.98)	32.13 (5.77)	32.65 (5.90)	
SBP, mean (SD) [mmHg]	133.9 (16.3)	132.4 (15.5)	132.8 (15.1)	
ER Basal (N=494)	**Placebo (n =170)**	**Empagliflozin 10mg (n=169)**	**Empagliflozin 25mg (n=155)**	
Demographic data				
Age, mean (SD) [years]	55.3 (10.1)	56.7 (8.7)	58.0 (9.4)	
Gender, N (%)				
Male	75 (39.9)	97 (52.2)	84 (44.2)	
Female	113 (60.1)	89 (47.8)	105 (55.6)	
Race, N (%)				
American Indian/Alaska native	4 (2.1)	3 (1.6)	2 (1.1)	
Asian	2 (1.1)	0	2 (1.1)	
Black/African American	8 (4.3)	7 (3.8)	4 (2.1)	
Hawaiian/ Pacific islander	0	1 (0.5)	0	
White	174 (92.6)	175 (94.1)	182 (96.3)	
eGFR (MDRD), mean (SD) [ml/min/1.73m^2]	83.41 (15.40)	84.14 (17.76)	84.35 (16.59)	
Baseline variables				
HbA_{1C}, mean (SD) [%]	8.33 (0.72)	8.39 (0.74)	8.29 (0.72)	

Baseline characteristic	Treatment group		
Time since diagnosis of T2DM, N (%)			
≤1 year	1 (0.5)	0	0
>1 to 5years	17 (9.0)	22 (11.8)	11 (5.8)
>5 to 10 years	40 (21.3)	44 (23.7)	38 (20.1)
>10 years	130 (69.1)	120 (64.5)	140 (74.1)
Body weight, mean (SD) [kg]	95.5 (17.5)	96.7 (17.9)	95.9 (17.3)
BMI, mean (SD) [kg]	34.65 (4.30)	34.72 (3.83)	34.99 (4.04)
SBP, mean (SD) [mmHg]	132.6 (15.8)	134.2 (16.4)	132.9 (14.2)

BMI = body mass index; DBP = diastolic blood pressure; eGFR = estimated glomerular filtration rate; FPG = fasting plasma glucose; HbA_{1C} = glycated haemoglobin; IU = international units; MDG = mean daily glucose; MDI = multiple daily injections; Met = metformin = PIO = pioglitazone; SBP = systolic blood pressure; SD = standard deviation; T2DM = type 2 diabetes mellitus

Settings

All the included studies were multicentre trials. ER Met and ER MetSU (reported as one in the MS) were conducted at 148 trial sites in 12 countries across Asia, Europe and North America (details of countries given in published studies – Haring et al 2013 and 2014.[27, 28] Countries included Canada, China, France, Germany, India, Korea, Mexico, Slovakia, Slovenia, Taiwan, Turkey and United States, so "Asian" will be a mix of ethnicities). ER Basal was conducted at 148 and 97 trial sites in 12 and 7 countries across Asia, Europe and North America respectively (details of countries not given in the MS). ER MDI was conducted at 104 sites in 14 countries across Europe, Latin America and North America (details of countries not given in the MS or published study).

Background treatments

In ER MetSU, the stable dose of SU used had to be at least half of the maximum recommended dose.

ER Basal included patients treated with basal glargine or determir insulin (≥ 20 IU/day) or NPH insulin (≥ 14 IU/day) with or without concomitant met and/or SU. The total insulin dose was not to change by more than 10% of the baseline value within 12 weeks prior to randomisation.

ER MDI included patients treated with multiple daily injections of basal and meal-time insulin alone or in combination with metformin. Pre-mixed insulin preparations were not allowed. The total prescribed insulin dose was to be > 60 IU/day at visit 1 (week 3 screening visit) and was not to be changed within 12 weeks prior to randomisation by more than 10% from the baseline value at randomisation.

Exclusion criteria are shown in Box 3.

Box 3. Reasons for exclusion

Uncontrolled hyperglycaemia defined as glucose level of > 13.3 mmol/l after an overnight fast during a 2 week open-label run-in; any other antidiabetic medication taken within 12 weeks prior to randomisation, except those defined as the permitted background medication; history of acute coronary syndrome, stroke or TIA ≤ 3 months prior to consent; indication of liver disease; history of renal dysfunction; history of bariatric surgery or other gastrointestinal surgeries that induce chronic malabsorption; history of cancer (except basal cell carcinoma) or treatment for cancer within the last 5 years; history of blood dyscrasias or any disorders causing haemolysis or unstable blood cells; contraindication to metformin; treatment with anti-obesity drugs ≤ 3 months prior to consent or any other treatment at the time of screening leading to unstable body weight; treatment with systemic steroids at time of consent; change in dosage of thyroid hormones ≤ 6 weeks of consent; any uncontrolled endocrine disorder except T2DM; premenopausal women who were nursing or pregnant or were of child-bearing potential but, not practicing an acceptable method of birth control or did not plan to continue using this method throughout the trial and did not agree to submit to periodic pregnancy testing during the trial; alcohol or drug abuse ≤ 3 months prior to consent; intake of an investigational drug in another trial ≤ 30 days prior to intake of trial medication; any other clinical condition that would jeopardise patient safety while participating in this trial

The demographic characteristics were generally well balanced across treatment groups in all studies (Table 2) except in ER MDI. Here, the proportion of male patients was significantly lower in placebo arm compared to the two empagliflozin arms (39.9% vs. 52.2% in empa 10 mg vs. 44.2% in empa 25 mg).

The mean baseline HbA1c levels were well-matched across randomised arms, but varied amongst the trials, as expected from their background therapies;

- ER Met between 7.86 and 7.94%
- ER MetSU 8.07 to 8.15%
- ER Basal 8.18 to 8.27%
- ER MDI between 8.29 and 8.39.

Most participants were White or Asian (there is insufficient data within the MS to specify proportion of Asian groups. Haring et al 2013 and 2014, assessing efficacy of empagliflozin in patients inadequately controlled with metformin or metformin plus SU, gives details of countries – please see above under ‘Settings’)), but varied amongst the trials;

- ER Met 53% white, 45% Asian
- ER MetSU, 39% white, 57% Asian
- ER Basal 93% white
- ER MDI 68% white, 20% Asian

Most patients in all studies had had T2DM for more than 5 years.
Mean baseline weights in ER Met ranged from 79.73 to 82.21 kg; in ER MetSU 76.23 to 77.50 kg; in ER Basal 90.46 to 94.71 kg; in ER MDI 95.5 to 96.7 kg. BMIs ranged between 28.32 and 29.72 kg/m^2 in ER MET and ER MetSU but was higher in the two insulin studies ranging between 31.75 and 34.72 kg/m^2.

Participants from the ER Met (69) and ER MetSU (103) whose HbA1c was ≥ 10%, were invited to participate in an open-label non-randomised study. All this group received empagliflozin 25 mg for 24 weeks. Those who completed the 24 weeks ER Met ER MetSU studies were followed up for another 52 weeks. The objective of this study was to assess the long-term safety and efficacy of empagliflozin. Patients continued the same initial treatment. We have presented findings of this study below.

Interventions and comparators
In each trial, patients were randomised to placebo, empagliflozin 10 mg, or empagliflozin 25 mg as add-ons to background regimens. In the basal insulin study the dose of basal insulin was not changed during the first 18 weeks but after that, it was adjusted if FPG level was >110 mg/dL (ERG conversion ~ 6.1mmol/l), at the discretion of the treating clinician.

Outcomes
The primary outcome measure was HbA1c, at 24 weeks in ER Met and ER MetSU and at 18 weeks in the insulin trials. Secondary outcome measures included change in body weight and mean daily glucose (MDG). The secondary outcome measure in ER Basal was change from baseline in HbA1c at 78 weeks. The secondary outcomes of the MDI study included change from baseline at 52 weeks in total insulin dose, body weight and HbA1c.

No data on lipid levels are given in the clinical effectiveness submission. Given some evidence of a rise in low density lipoprotein-cholesterol (LDL-C) with canagliflozin, the ERG sought data on lipid changes, and found them in the published version of the ER MET study by Haring et al 2014[28], and these are shown in Table 3 below.

Table 3. Lipid changes (source Haring et al 2014 - supplementary data table 4)

	Placebo		Empagliflozin 10 mg		Empagliflozin 25 mg	
	Baseline	**Change from baseline**	**Baseline**	**Change from baseline**	**Baseline**	**Change from baseline**
Total cholesterol (TC)(mmol/l)	4.55 (0.07)	0.09 (0.05)	4.50 (0.06)	0.23 (0.05)	4.59 (0.07)	0.21 (0.05)
Difference vs. placebo				0.14 (0.07)		0.13 (0.07)
p-value				0.043		0.071
HDL cholesterol (mmol/l)‡	1.22 (0.02)	0.00 (0.01)	1.28 (0.02)	0.08 (0.01)	1.28 (0.02)	0.06 (0.01)
Difference vs. placebo				0.08 (0.02)		0.06 (0.02)
p-value				<0.001		0.001
LDL cholesterol (mmol/l)‡	2.46 (0.06)	0.03 (0.04)	2.40 (0.06)	0.15 (0.04)	2.48 (0.06)	0.15 (0.04)
Difference vs. placebo				0.12 (0.06)		0.12 (0.06)
p-value				0.043		0.032
Triglycerides (mmol/l)‡	1.96 (0.09)	0.11 (0.08)	1.95 (0.09)	0.00 (0.08)	1.84 (0.08)	−0.04 (0.08
Difference vs. placebo				−0.11 (0.11)		−0.14 (0.11)
p-value				0.327		0.204

The TC/HD ratio at baseline is 3.6 and at end 3.6 for the 25 mg dose. This supports the manufacturer's assumption for modelling on page 227, where it states that the assumption was made because of the lack of information from the systematic review – which seems a little strange when the data were available.

3.2 Results

In the Boehringer submission, the results of primary and secondary outcomes are reported in section 6.5 (pages 107 to 158).

The manufacturer did not report proportions of patients achieving HbA1c level targets of ≤6.5%, ≤7% and ≤7.5% in the submission but provided the data in response to a clarification question from the ERG. We asked for these three thresholds because;

- 6.5% is the target in the NICE clinical guideline (CG87) for type 2 diabetes
- It may be that in the current review of this guideline, the target might be relaxed with 7.0% being a possibility
- The current switching point, at which treatment should be intensified, is 7.5%.

For convenience, results from all the relevant studies have been combined in a table and reported below (section 3.2.1). The ERG has also constructed bar charts to compare the findings across different treatment arms and different baseline HbA1c levels.

3.2.1 Proportion of patients achieving HbA1c levels of ≤6.5%, ≤7% and ≤7.5%

At 24 weeks, the proportion of patients achieving HbA1c levels targets of ≤6.5%, ≤7.0% and ≤7.5% was greater in the empagliflozin 25 mg group (ER Met: 23%, 48.8%, 71.4%; ER MetSU: 15.7%, 36.6%, 58.3%) than in the lower dose of empagliflozin (ER Met: 15.7%, 45.2%, 71%; ER MetSU: 14.2%, 35.6%, 57.3%) and placebo groups in both ER Met and ER MetSU studies (Table 4 and Figure 3). Similar findings were seen in the ER MDI insulin study at 18 weeks follow-up period (empa 25 mg: 15.3%, 36%, 55.6%; empa 10 mg: 8.1%, 25.8%, 51.1%; placebo: 6.9%, 17%, 35.6%) (Table 5 and Figure 3). The findings in the ER Basal study was mixed (Table 5 and Figure 3). At 18 weeks, the proportion of patients achieving the target of ≤6.5% was similar with two doses of empagliflozin (6.5% with both doses), while more patients in the empagliflozin 10 mg group met the target of ≤7.5% than in the empagliflozin 25 mg group (45% vs. 40%) and placebo (45% vs. 24.7%) (Table 5).

Table 4. Proportion of achieving HbA1c targets at 24 weeks (ER Met and ER MetSU)

	Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**
ER Met (24 weeks)			
≤ 6.5%			
All HbA1c	15.7%	23%	6.8%
Baseline <8.0%	21.3%	31.5%	6.8%
Baseline 8.0-8.9%	9%	15.2%	0
Baseline 9.0% and over	7.1%	0	0
≤ 7.0%			
All HbA1c	45.2%	48.8%	19.8%
Baseline <8.0%	59.8%	58.1%	30.6%
Baseline 8.0-8.9%	26.9%	42.4%	3.3%
Baseline 9.0% and over	25%	17.4%	7.7%
≤ 7.5%			
All HbA1c	71%	71.4%	44%
Baseline <8.0%	85.2%	85.5%	64.5%
Baseline 8.0-8.9%	58.2%	59.1%	16.7%
Baseline 9.0% and over	39.3%	30.4%	11.5%
	Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**
ER MetSU (24 weeks)			
≤ 6.5%			
All HbA1c	14.2%	15.7%	4.9%
Baseline <8.0%	21.8%	23.8%	8.9%
Baseline 8.0-8.9%	6.2%	10.3%	1.4%
Baseline 9.0% and over	8.8%	3%	0
≤ 7.0%			
All HbA1c	35.6%	36.6%	13.8%
Baseline <8.0%	50.9%	49.5%	23.2%
Baseline 8.0-8.9%	24.7%	30.8%	5.6%
Baseline 9.0% and over	11.8%	9.1%	2.4%
≤ 7.5%			

All HbA1c	57.3%	58.3%	32.4%
Baseline <8.0%	70%	69.5%	49.1%
8.0-8.9%	48.1%	57.7%	23.9%
9.0% and over	38.2%	24.2%	2.4%

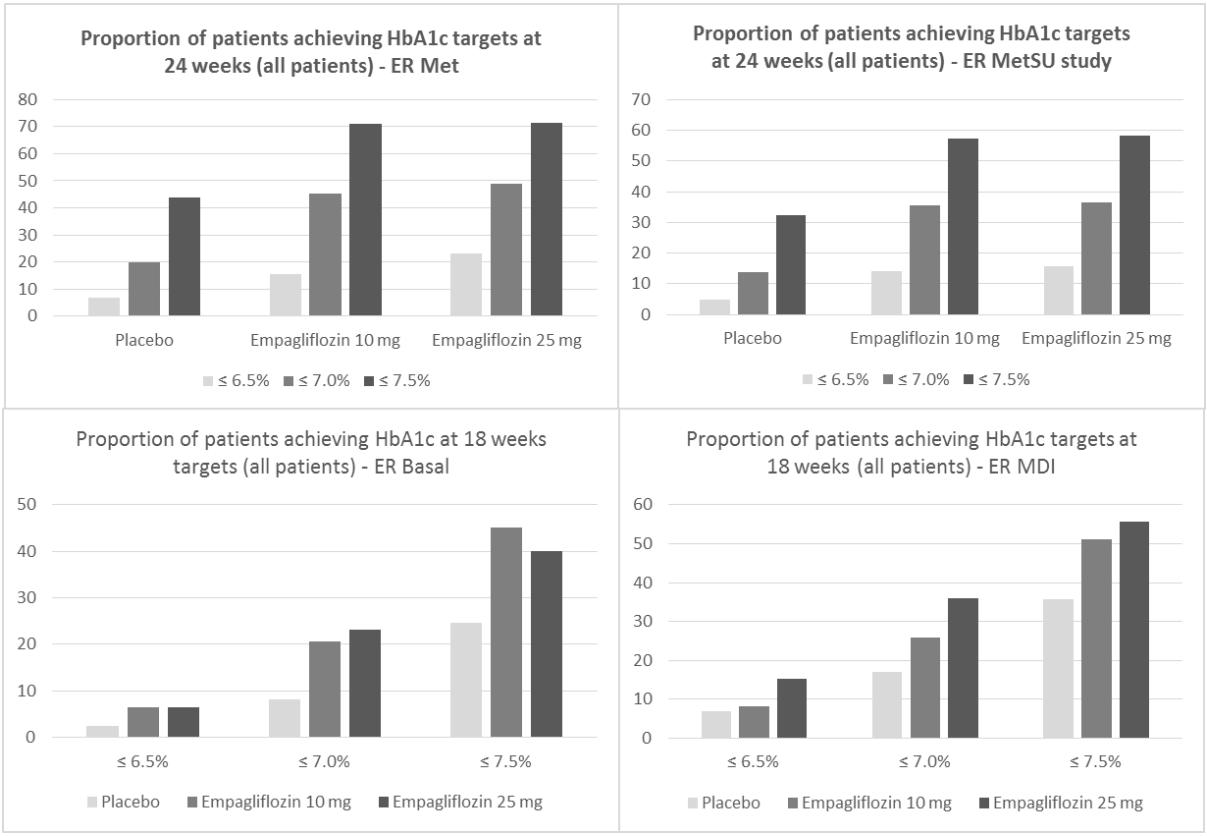


Figure 3. Proportion of patients achieving HbA1c targets (ER Met; ER MetSU; ER Basal; ER MDI)

Table 5. Proportion of achieving HbA1c targets at 18 weeks (ER Basal and ER MDI)

	Empagliflozin 10 mg	Empagliflozin 25 mg	Placebo
ER Met Basal			
≤ 6.5%			
All HbA1c	6.5%	6.5%	2.4%
Baseline <8.0%	11.1%	13%	4.2%
Baseline 8.0-8.9%	5.2%	2.1%	1.5%
Baseline 9.0% and over	0	0	0
≤ 7.0%			
All HbA1c	20.7%	23.2%	8.2%
Baseline <8.0%	34.7%	31.9%	4.1%
Baseline 8.0-8.9%	13.8%	19.1%	6.0%
Baseline 9.0% and over	5.1%	12.8%	0
≤ 7.5%			
All HbA1c	45%	40%	24.7%
Baseline <8.0%	59.7%	52.2%	43.7%
Baseline 8.0-8.9%	43.1%	34%	11.9%
Baseline 9.0% and over	20.5%	25.6%	9.4%
ER Met MDI			
≤ 6.5%			
All HbA1c	8.1%	15.3%	6.9%
Baseline <8.0%	11.1%	22.4%	10.8%
Baseline 8.0-8.9%	5.9%	14.3%	3.7%
Baseline 9.0% and over	8.5%	5.3%	7.1%
≤ 7.0%			
All HbA1c	25.8%	68/189 (36%)	32/188 (17%)
Baseline <8.0%	37%	32 (47.8%)	19 (29.2%)
Baseline 8.0-8.9%	20%	30 (35.7%)	10 (12.3%)
Baseline 9.0% and over	23.4%	6 (15.8%)	3 (7.1%)
≤ 7.5%			
All HbA1c	51.1%	55.6%	35.6%
Baseline <8.0%	64.8%	62.7%	52.3%
Baseline 8.0-8.9%	44.7%	60.7%	37%
Baseline 9.0% and over	46.8%	31.6%	7.1%

The manufacturer also provided the results according to baseline HbA1c levels. In ER Met study, the proportion of patients achieving HbA1c target of ≤ 6.5% was greater in the empagliflozin 25 mg group (31.5%; 15.2%) than in the empagliflozin 10 mg (15.7%; 9%) and placebo group (6.8%; 0%) in patients with baseline HbA1c levels of <8% and 8 to 8.9% (Table 4 and Figure 4). Those with baseline HbA1c level of ≥9%, the HbA1c target of ≤ 6.5% was only met with empagliflozin 10 mg (Figure 4). The results were mixed for other HbA1c targets. In patients with baseline HbA1c levels of <8%, similar proportion of patients met the targets of ≤7.0% and ≤7.5% (~58% and ~85%). The number of patients meeting the target of ≤7.5% with a baseline HbA1c level of 8-8.9% was also similar with two doses of empagliflozin (~58%). Empagliflozin 25 mg led to more patients meeting the target of ≤7.0% than empagliflozin 10 mg and placebo if their baseline HbA1c level was 8-8.9% (42.4% vs. 26.9% vs. 3.3%). More patients met the target of ≤7.0% (25% vs. 17.4% vs. 7.7%) and ≤7.5% (39.3% vs. 30.4% vs. 11.5%) with empagliflozin 10 mg than with empagliflozin 25 mg and placebo if their baseline HbA1c level was ≥9%.

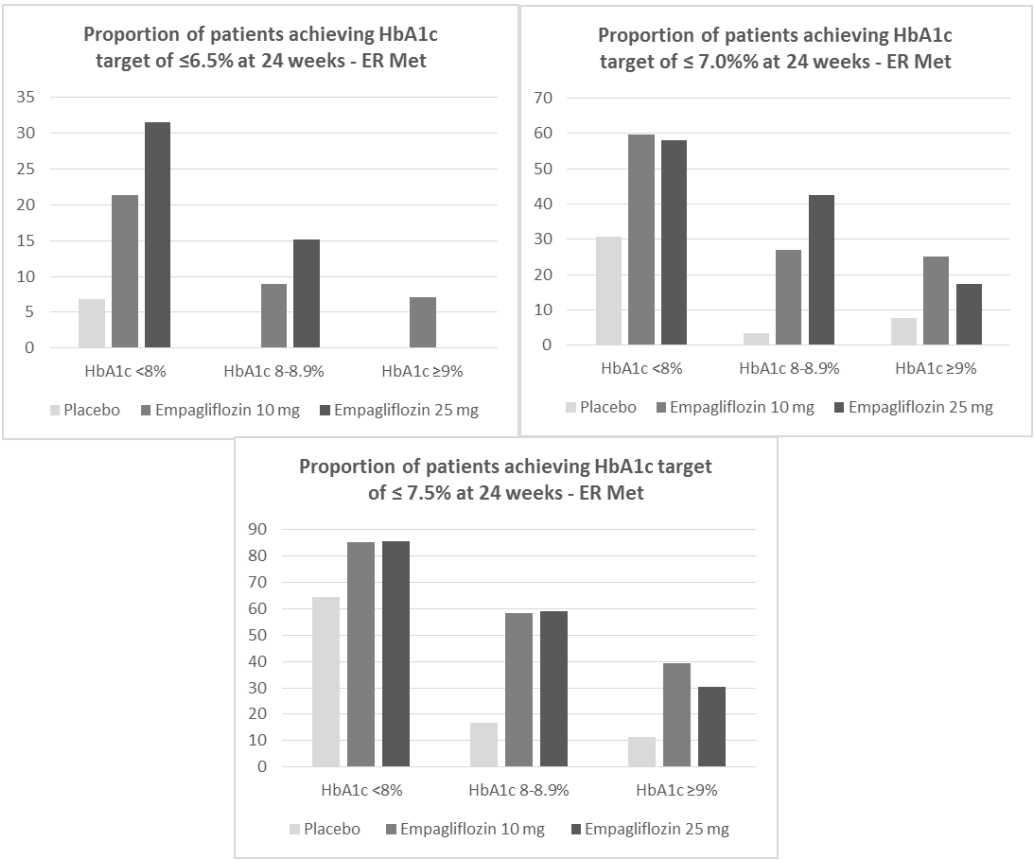


Figure 4. Proportion of patients achieving HbA1c targets according to their baseline HbA1c (ER Met study)

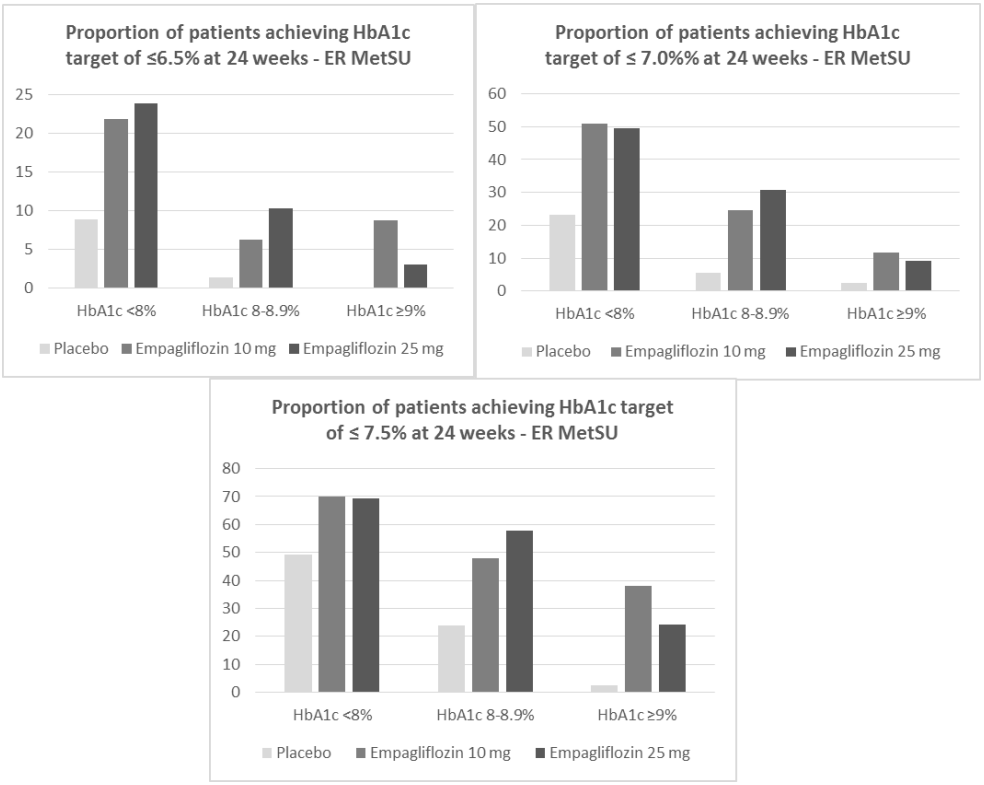


Figure 5. Proportion of patients achieving HbA1c targets according to their baseline HbA1c (ER MetSU study)

In the MetSU study, the findings according to baseline HbA1c levels were mixed. The proportion of patients achieving the target of ≤ 6.5% was more with empagliflozin 25 mg in patients with baseline HbA1c levels of <8% and 8-8.9% compared to empagliflozin 10 mg and placebo. In contrast, more patients with HbA1c levels of ≥9% met the target of ≤6.5% with empagliflozin 10 mg than with empagliflozin 25 mg and placebo. Empagliflozin 10 mg led to slightly more people with baseline HbA1c levels of 8% or ≥9% meet HbA1c level target of ≤7.0% and ≤7.5%, whereas comparatively more patients with baseline HbA1c level of 8 to 8.9% met the target of ≤7% and ≤7.5% with empagliflozin 25 mg (Figure 5).

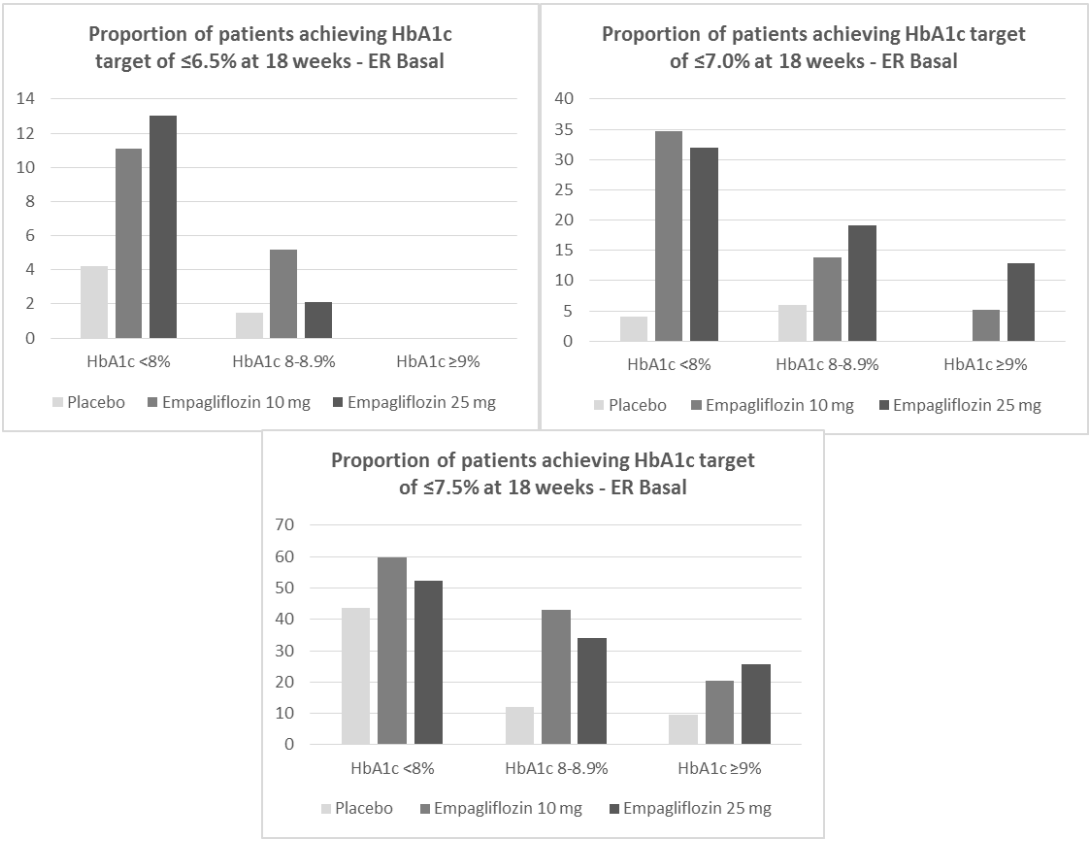


Figure 6. Proportion of patients achieving HbA1c targets according to their baseline HbA1c (ER Basal study)

In ER Basal study, mixed results were observed (Figure 6). In patients with baseline HbA1c level of 8%, more patients receiving empagliflozin 25 mg met the target of ≤6.5% while more patients receiving empagliflozin 10 mg met the targets of ≤7.0% and ≤7.5%. Those with baseline HbA1c level of 8-8.9%, empagliflozin 10 mg led more patients to meet the targets of ≤6.5% and ≤7.5% while, more patients met the target of ≤7.0% with empagliflozin 25 mg. None of the patients with baseline HbA1c level of ≥9% met the target of ≤6.5%. Empagliflozin 25 mg led to more patients meeting the HbA1c target of ≤7.0% and ≤7.5% if their baseline HbA1c was ≥9%.

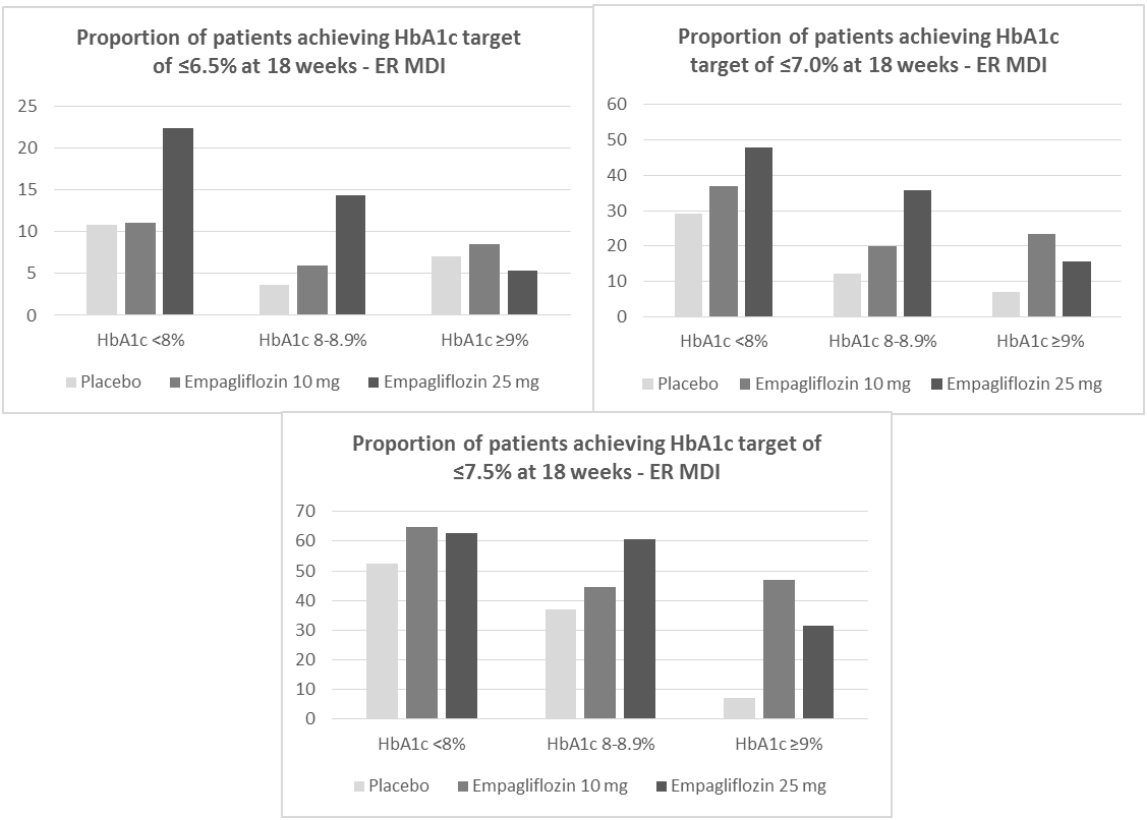


Figure 7. Proportion of patients achieving HbA1c targets according to their baseline HbA1c (ER MDI study)

In ER MDI study, it was observed that in patients with baseline HbA1c level of <8%, more patients receiving the higher dose of empagliflozin met the targets of ≤6.5% and ≤7.0% while, slightly more patients receiving empagliflozin 10 mg met the target of ≤7.5% (Figure 7). In patients with baseline HbA1c of 8-8.9%, more patients receiving empagliflozin 25 mg met the targets of ≤6.5%, ≤7.0% and ≤7.5%. In contrast, empagliflozin 10 mg led to more patients meeting all the HbA1c if their baseline HbA1c level was ≥9%.

3.2.2 Mean change in HbA1c

The manufacturer has reported results for mean change in HbA1c (%) in section 6.5 (tables 18, 21 25 and figures 15, 16, 22, 26). The findings from the extension study are reported in table 20 (but are now in the public domain as a conference abstract) and figures 20 and 21. For convenience, the ERG has combined findings from all the relevant studies in the following table (Table 6).

Table 6. Mean change in HbA1c (%)

ER Met (24 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.70% SE 0.05	-0.77% SE 0.05	-0.13% SE 0.05	**Empa 10 mg vs. placebo:** -0.57% SE 0.07 (95% CI -0.72 to -0.42; p<0.0001); **Empa 25 mg vs. placebo:** -0.64% SE 0.07 (95% CI -0.79 to -0.48; p<0.0001)
Met only roll-over study (76 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.62 SE 0.05	-0.74 SE 0.05	-0.01 SE 0.05	**Empa 10 mg vs. placebo:** -0.61% SE 0.07 (95% CI -0.75 to -0.46; p<0.0001); **Empa 25 mg vs. placebo:** -0.73% SE 0.07 (95% CI -0.88 to -0.58; p<0.0001)
ER MetSU (24 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.82% SE 0.05	-0.77% SE 0.05	-0.17% SE 0.05	**Empa 10 mg vs. placebo:** -0.64% SE 0.07 (97.5% CI -0.79 to -0.49; p<0.0001); **Empa 25 mg vs. placebo:** -0.59% SE 0.07 (97.5% CI -0.74 to -0.44; p<0.0001)
Met plus SU roll-over study (76 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.74% SE 0.06	-0.72% SE 0.06	-0.03% SE 0.06	**Empa 10 mg vs. placebo:** -0.72% SE 0.08 (95% CI -0.87 to -0.56; p<0.0001); **Empa 25 mg vs. placebo:** -0.69% SE 0.08 (95% CI -0.85 to -0.53; p<0.0001)
Basal insulin study (18 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.57% SE 0.07	-0.71% SE 0.07	-0.01% SE 0.07	**Empa 10 mg vs. placebo:** -0.56% SE 0.10 (97.5% CI -0.78 to -0.33; p<0.0001); **Empa 25 mg vs. placebo:** -0.70% SE 0.10 (97.5% CI -0.93 to -0.47; p<0.0001)
Basal insulin study (78 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.48% SE 0.08	-0.64% SE 0.09	-0.02% SE 0.09	**Empa 10 mg vs. placebo:** -0.46% SE 0.12 (97.5% CI -0.73 to -0.19; p=0.0001); **Empa 25 mg vs. placebo:** -0.62% SE 0.12 (97.5% CI -0.90 to -0.34; p<0.0001)
MDI insulin study (18 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-0.94% SE 0.05	-1.02% SE 0.05	-0.50% SE 0.05	**Empa 10 mg vs. placebo:** -0.44% SE 0.08 (97.5% CI -0.61 to -0.27; p<0.0001); **Empa 25 mg vs. placebo:** -0.52% SE 0.07 (97.5% CI -0.52 to -0.07; p<0.0001)
MDI insulin study (52 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-1.18% SE 0.08	-1.27% SE 0.08	-0.81% SE 0.08	**Empa 10 mg vs. placebo:** -0.38% SE 0.11 (97.5% CI -0.62 to -0.13; p<0.0001); **Empa 25 mg vs. placebo:** -0.46% SE 0.11 (97.5% CI -0.70 to -0.22; p<0.0001)

In ER Met, the adjusted mean differences compared to placebo were -0.57% SE 0.07 (95% CI -0.72 to -0.42; p<0.0001) and -0.64% SE 0.07 (95% CI -0.79 to -0.48; p<0.0001) in empagliflozin 10 mg and 25 mg groups respectively. The larger dose reduced HbA1c by only 0.07% more.
In the long term extension study, the mean reduction in HbA1c at 76 weeks was greater (-0.73% SE 0.07 (95% CI -0.88 to -0.58; p<0.0001) with empagliflozin 25 mg than 10 mg. For convenience, figures 15 and 20 have been reproduced below (Figure 8).

In ER MetSU, the adjusted mean decrement in HbA1c from baseline to week 24 was slightly (0.05%) greater with empagliflozin 10 mg than with 25 mg, with both larger than placebo (-0.82% SE 0.05 vs. -0.77% SE 0.05 vs. -0.17% SE 0.05). At week 76, the reduction in HbA1c was slightly greater with 10 mg empagliflozin than with 25 mg, with reductions in both empagliflozin groups significantly greater than placebo (-0.74% vs. -0.72% vs. -0.03%). Figures 16 and 21 have been reproduced below (Figure 8).

In both insulin studies, changes in HbA1c from baseline at 18 weeks were higher in the empagliflozin 25 mg group than in the empagliflozin 10 mg or placebo groups [basal study: -0.71% vs. -0.57% vs. -0.01%; MDI study: -1.02% vs. -0.94% vs. -0.50%]. In the basal study, at 78 weeks follow-up and at 52 weeks in the MDI insulin study, mean change in HbA1c from baseline was again found to be greater in the higher dose of empagliflozin than in the lower dose of empagliflozin [basal study: -0.64% vs. -0.48% vs. -0.02%; MDI study: -1.27% vs. -1.18% vs. -0.81%].

The differences between the two doses of empagliflozin are therefore modest – 0.07, 0.05, 0.14 and 0.08% in the primary outcomes - with the second figure favouring the lower dose.

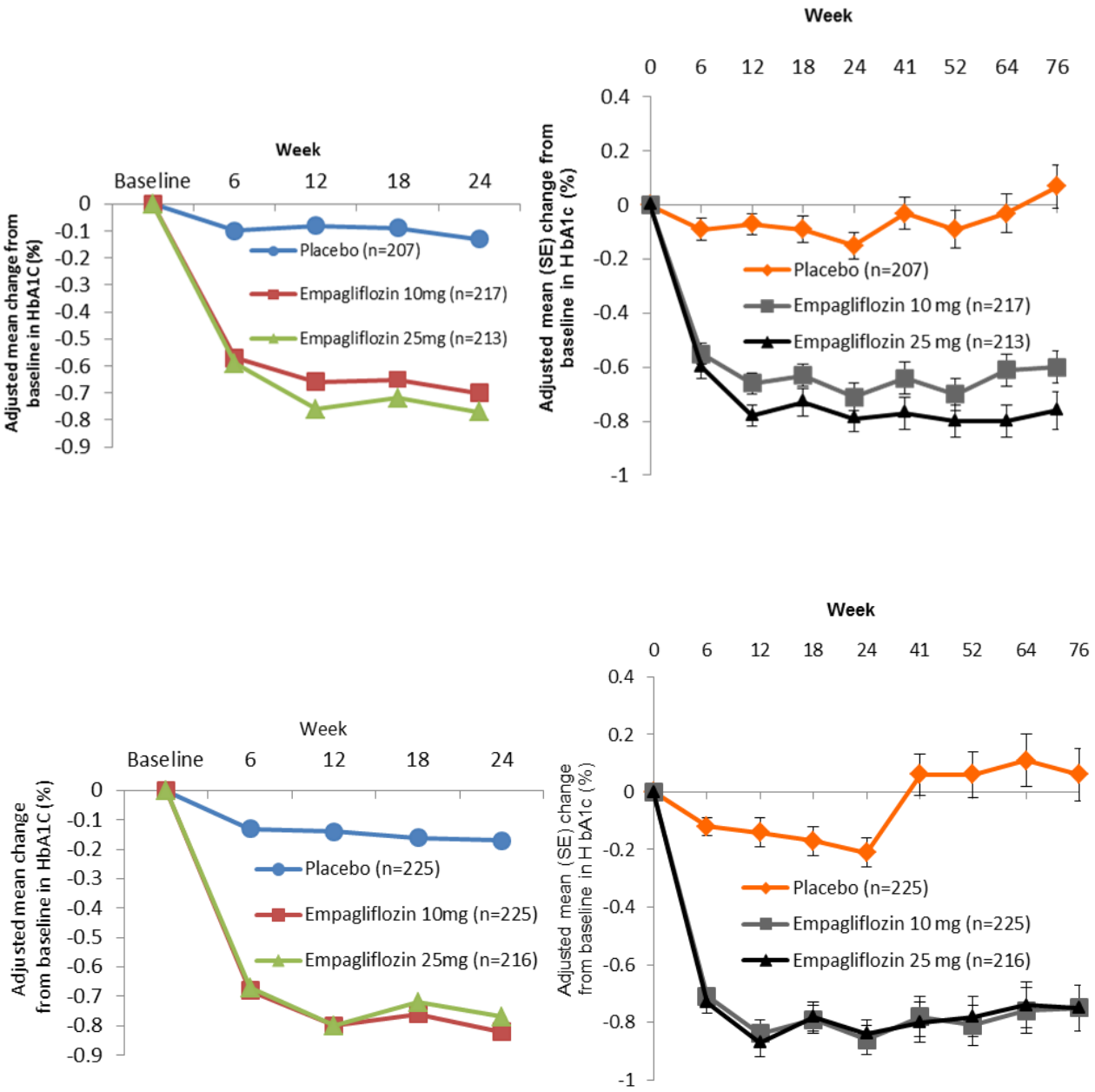


Figure 8. Adjusted mean change from baseline in HbA1c (%) results overtime (metformin only – top two figures)(metformin plus SU study – two bottom figures)

3.2.3 Mean change in body weight

The results for mean change in body weight (kg) are reported in MS section 6.5 (tables 18, 20, 21 and 25) including that from the long-term extension study (1245.31). For convenience, the ERG has combined findings from all the relevant studies in the following table (Table 7).

In ER Met mean weight reduction was slightly (0.38kg) greater on empagliflozin 25 mg than on 10 mg, with both greater than placebo at both 24 weeks (-2.46 kg vs. -2.08 kg vs. -0.45) and 76 weeks (-2.65 kg vs. -2.39 kg vs. -0.46 kg). In contrast, the mean reduction in weight in the ER MetSU was slightly more in the empagliflozin 10 mg group than the 25 mg group at 76 weeks (-2.44 kg vs. -2.28

kg vs. -0.63 kg on placebo), but, at 24 weeks, the reduction was greater with the larger dose (-2.39 kg in 25 mg group vs. -2.16 kg in 10 mg group vs. -0.39 kg placebo).

In ER Basal, mean weight reduction was higher with empagliflozin 10 mg at both 18 weeks (-2.09 kg vs. -0.92 kg vs. -0.05 kg) and 78 weeks (-2.47 kg vs. -1.96 kg vs. +1.16 kg) than with empagliflozin 25 mg and placebo. In ER MDI at 52 weeks, mean weight reduction was greater with the higher dose of empagliflozin (-2.04 kg vs. -1.95 kg). Overall, differences in weight loss between the two doses were inconsistent and small.

Table 7. Mean change in bodyweight (kg)

Met only study (24 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-2.08 kg SE 0.17	-2.46 kg SE 0.17	-0.45 kg SE 0.17	Empa 10 mg vs. placebo: -1.63 kg SE 0.24 (97.5% CI -2.17 to -1.08; p<0.0001) Empa 25 mg vs. placebo: -2.01 kg SE 0.24 (97.5% CI -2.56 to -1.46; p<0.0001)
Met only roll-over study (76 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-2.39 kg SE 0.21	-2.65 kg SE 0.21	-0.46 kg SE 0.22	Empa 10 mg vs. placebo: -1.93 kg SE 0.30 (95% CI -2.52 to -1.34; p<0.0001) Empa 25 mg vs. placebo: -2.19 kg SE 0.30 (95% CI -2.79 to -1.60; p<0.0001)
Met plus SU study (24 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-2.16 kg SE 0.15	-2.39 kg SE 0.16	-0.39 kg SE 0.15	Empa 10 mg vs. placebo: -1.76 kg SE 0.22 (97.5% CI -2.25 to -1.28; p<0.0001) Empa 25 mg vs. placebo: -1.99 kg SE 0.22 (97.5% CI -2.48 to -1.50; p<0.0001)
Met plus SU roll-over study (76 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-2.44 kg SE 0.19	-2.28 kg SE 0.20	-0.63 kg SE 0.19	Empa 10 mg vs. placebo: -1.81 kg SE 0.27 (95% CI -2.34 to -1.27; p<0.0001) Empa 25 mg vs. placebo: -1.64 kg SE 0.27 (95% CI -2.18 to -1.11; p<0.0001)
Basal insulin study (18 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-2.09 kg SE 0.66	-0.92 kg SE 0.72	-0.05 kg SE 0.68	Empa 10 mg vs. placebo: -2.04 kg SE 0.95 (95% CI -3.90 to 0.18; p=0.0320) Empa 25 mg vs. placebo: -0.87 kg SE 0.99 (95% CI -2.81 to 1.08; p=0.3818)
Basal insulin study (78 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-2.47 kg SE 0.76	-1.96 kg SE 0.82	1.16 kg SE 0.80	Empa 10 mg vs. placebo: -3.63 kg SE 1.10 (95% CI -5.81 to -1.45; p=0.0012) Empa 25 mg vs. placebo: -3.12 kg SE 1.15 (95% CI -5.39 to -0.85; p=0.0073)
MDI insulin study (52 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-1.95 kg SE 0.36	-2.04 kg SE 0.36	0.44 kg SE 0.36	Empa 10 mg vs. placebo: -2.39 kg SE 0.51 (95% CI -3.54 to -1.24; p<0.0001) Empa 25 mg vs. placebo: -2.48 kg SE 0.51 (95% CI -3.63 to -1.33; p<0.0001)

The ERG for convenience and brevity, because of importance in the economic model, have represented the key weight changes in the empagliflozin trials graphically (Figure 9 and Figure 10).

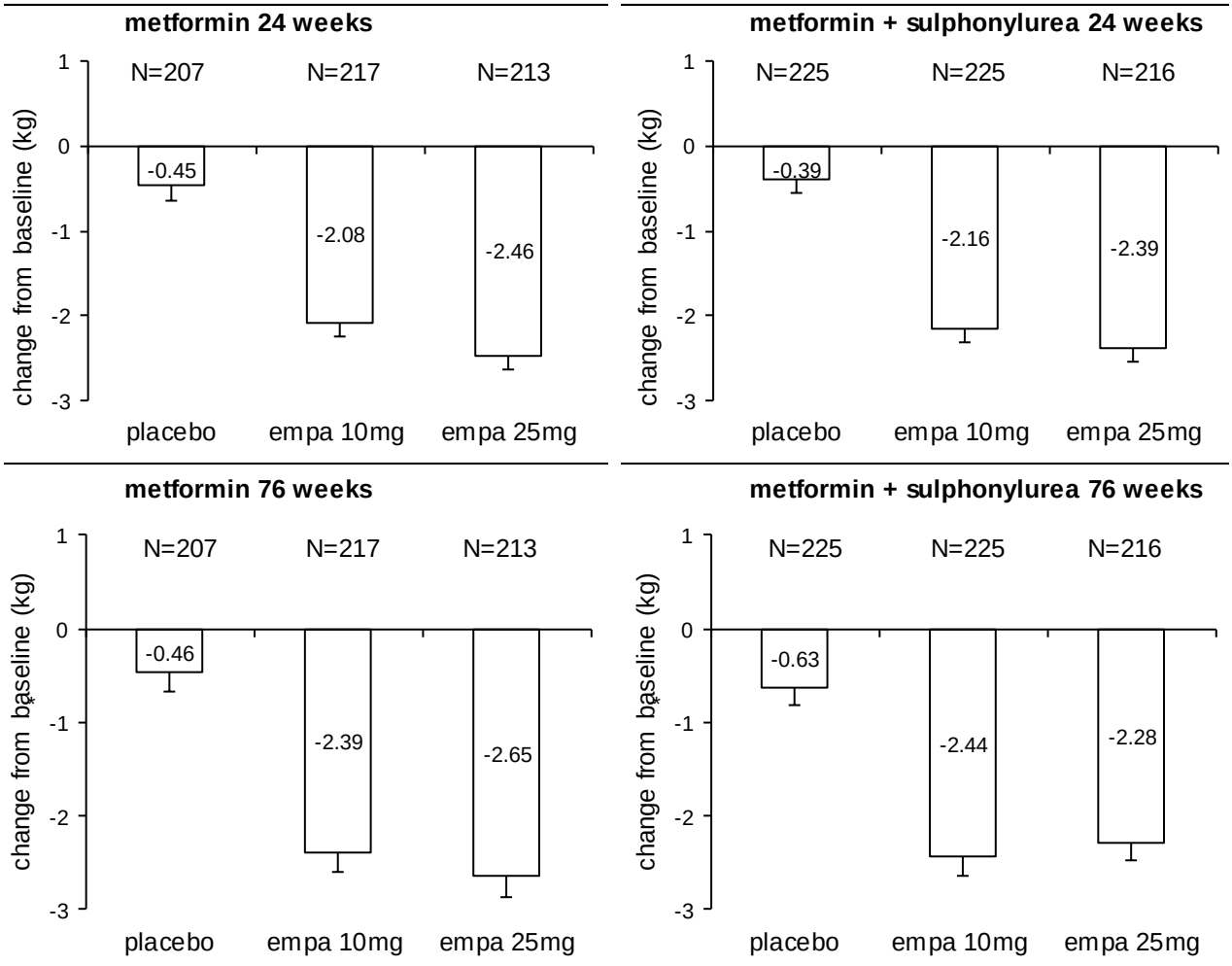


Figure 9. Mean change in weight (kg) from baseline in empagliflozin trials - dual and triple therapy

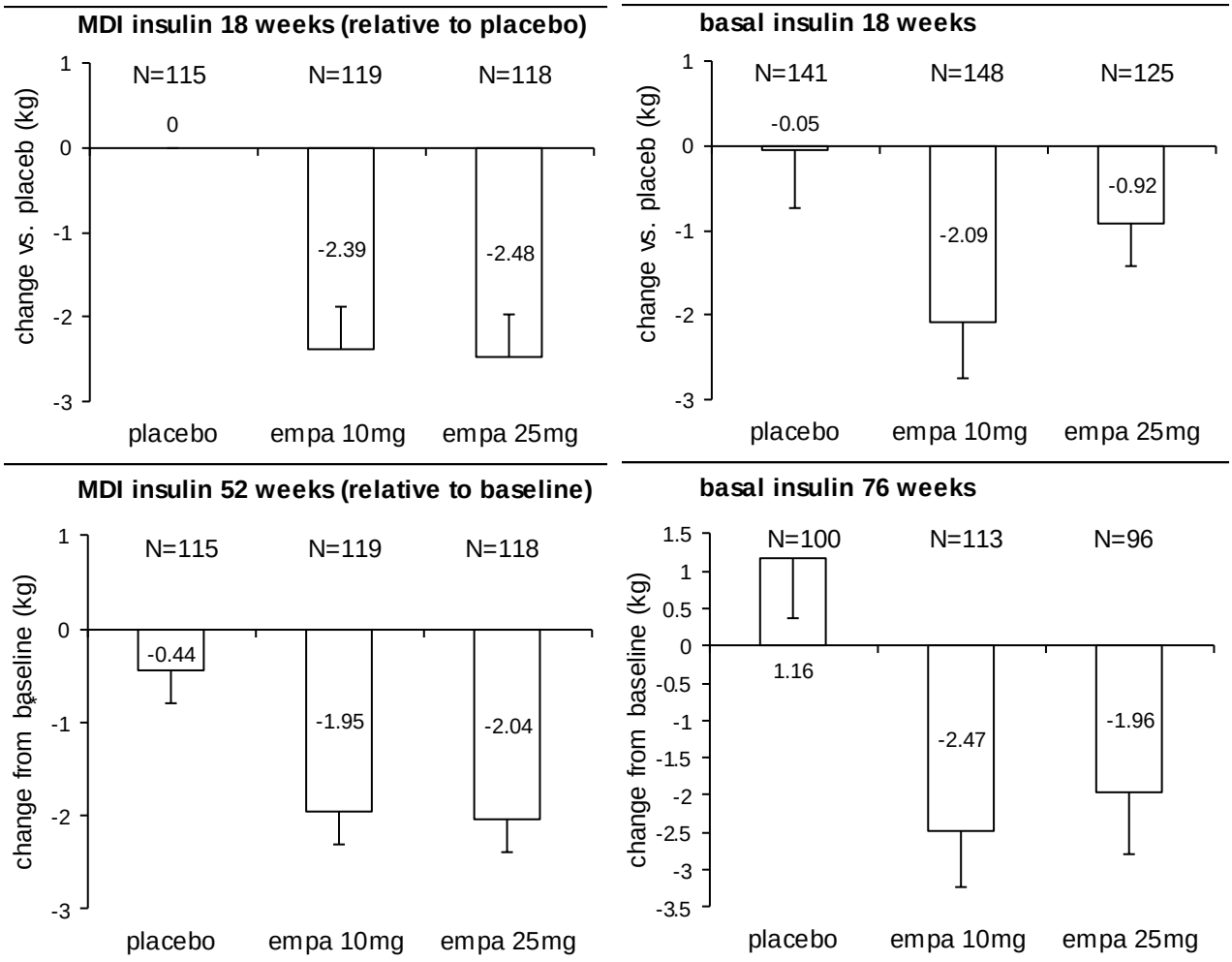


Figure 10. Mean change in weight (kg) from baseline in empagliflozin trials- add on to insulin

3.2.4 Mean change in systolic blood pressure (SBP)

In the MS, findings related to mean change in SBP (mmHg) are reported in section 6.5 (tables 18, 20, 21 and 25). Results from all the relevant studies have been combined and presented below (Table 8).

The ERMet results suggest that at 24 weeks, mean reduction in SBP was higher with empagliflozin 25 mg than with empagliflozin 10 mg and placebo (-5.2 mmHg vs. -4.5 mmHg vs. -0.4 mmHg) but the difference between doses (0.7 mmHg) is trivial. In contrast, at 76 weeks, mean reduction in SBP was greater with the lower dose of empagliflozin than with the higher dose of empagliflozin and placebo (-5.2 mmHg vs. -4.5 mmHg vs. -0.8 mmHg). In the ER MetSU study, the mean reduction in SBP at 24 weeks was slightly more with empagliflozin 10 mg compared to empagliflozin 25 mg and placebo (-4.1 mmHg vs. -3.5 mmHg vs. -1.4 mmHg). However, at 76 weeks, there was no difference between the two doses of empagliflozin (-3.8 with empagliflozin 10 mg vs. -3.7 mmHg with empagliflozin 25 mg).

In ER Basal, the mean reduction in SBP was greater with empagliflozin 10 mg than with empagliflozin 25 mg and placebo at both 18 weeks (-3.7 mmHg vs. -3.3 mmHg vs. -0.3 mmHg) and 78 weeks (-4.1 mmHg vs. -2.4 mmHg vs. 0.1 mmHg). In ER MDI, the mean reduction in SBP was slightly more with the lower dose empagliflozin at 18 weeks than with the higher dose empagliflozin (-3.6 mmHg vs. -2.9 mmHg) but, at 52 weeks, there was no difference between the two (-3.9 mmHg vs. -4.0 mmHg).

So as with HbA1c and weight, there were no clinically significant differences in SBP between the two doses of empagliflozin.

Table 8. Mean change in SBP (mmHg) from baseline

Met only study (24 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-4.5 mmHg SE 0.7	-5.2 mmHg SE 0.7	-0.4 mmHg SE 0.7	Empa 10 mg vs. placebo: -4.1 mmHg SE 1.0 (95% CI -6.2 to -2.1; p<0.0001) Empa 25 mg vs. placebo: -4.8 mmHg SE 1.0 (95% CI -6.9 to -2.7; p<0.0001)
Met only roll-over study (76 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-5.2 mmHg SE 0.8	-4.5 mmHg SE 0.8	-0.8 mmHg SE 0.8	Empa 10 mg vs. placebo: -4.4 mmHg SE 1.1 (95% CI -6.6 to -2.3; p<0.0001) Empa 25 mg vs. placebo: -3.7 mmHg SE 1.1 (95% CI -5.9 to -1.5; p=0.0008)
Met plus SU study (24 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-4.1 mmHg SE 0.7	-3.5 mmHg SE 0.7	-1.4 mmHg SE 0.7	Empa 10 mg vs. placebo: -2.7 mmHg (95% CI -4.6 to -0.8; p=0.0049) Empa 25 mg vs. placebo: -2.1 mmHg SE 1.0 (95% CI -4.0 to -0.2; p=0.0321)
Met plus SU roll-over study (76 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-3.8 mmHg SE 0.7	-3.7 mmHg SE 0.7	-1.6 mmHg SE 0.7	Empa 10 mg vs. placebo: -2.2 mmHg SE 1.0 (95% CI -4.1 to -0.3; p=0.0213) Empa 25 mg vs. placebo: -2.1 mmHg SE 1.0 (95% CI -4.1 to -0.2; p=0.0288)
Basal insulin study (18 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-3.7 mmHg SE 0.9	-3.3 mmHg SE 1.0	-0.3 mmHg SE 0.9	Empa 10 mg vs. placebo: -3.4 mmHg SE 1.3 (95% CI 6.0 to -0.8; p=0.0111) Empa 25 mg vs. placebo: -3.0 mmHg SE 1.4 (95% CI -5.7 to -0.4; p=0.0267)
Basal insulin study (78 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-4.1 mmHg SE 1.0	-2.4 mmHg SE 1.1	0.1 mmHg SE 1.0	Empa 10 mg vs. placebo: -4.2 mmHg SE 1.4 (95% CI -7.0 to -1.3; p=0.0040) Empa 25 mg vs. placebo: -2.4 mmHg SE 1.5 (95% CI -5.4 to 0.5; p=0.0987)
MDI insulin study (18 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-3.6 mmHg SE 0.8	-2.9 mmHg SE 0.8	-1.2 mmHg SE 0.8	Empa 10 mg vs. placebo: -2.4 mmHg SE 1.2 (95% CI -4.7 to -0.2; p=0.0366) Empa 25 mg vs. placebo: -1.7 mmHg SE 0.9

			(95% CI -3.9 to 0.6; p=0.1409)
MDI insulin study (52 weeks)			
Empagliflozin 10 mg	**Empagliflozin 25 mg**	**Placebo**	**Difference between groups**
-3.9 mmHg SE 0.8	-4.0 mmHg SE 0.8	-2.6 mmHg SE 0.8	Empa 10 mg vs. placebo: -1.4 mmHg SE 1.1 (95% CI -3.6 to 0.9; p=0.2337) Empa 25 mg vs. placebo: -1.4 mmHg SE 1.1 (95% CI -3.7 to 0.8; p=0.2097)

The manufacturer also provided data from the EMPA-REG BP trial, which is not described in detail here because it was in monotherapy in people with diabetes who had never had glucose-lowering agents. In this trial, 1830 people with hypertension (mean baseline BP 142/84) in 121 centres were randomised to empagliflozin 10 or 25 mg, or placebo. SBP fell by means of 2.95 mmHg on 10 mg and 3.68 mmHg, but rose by 0.48 mmHg on placebo. HbA1c fell by 0.59 on 10 mg and by 0.62 on 25 mg.

3.2.5 Health-related quality of life

The health-related quality of life was reported as changes in EQ-5D and EQ-5D VAS score from baseline. The findings are reported on pages 155 and 156 of the MS. The evidence for this comes from six trials (1245.19, 20, 23, 28, 36 and 49). Background treatments in these trials differed. In 1245.19, patients were receiving pioglitazone, in 1245.20 metformin, in 1245.23 either metformin or metformin plus SU, in 1245.36 any existing anti-diabetic treatment however, patients also had renal impairment and in 1245.49 patients were receiving multiple daily injections of insulin with or without metformin. The manufacturer performed the analysis in two phases. In the first phase, the manufacturer analysed the data from four trials – 1245.19, 20, 23 and 36. In the latter phase, data from the two remaining trials i.e. 1245.28 and 1245.49 were analysed and then all the results from both phases were pooled together to compare empagliflozin and placebo.

The baseline mean EQ-5D utility index score ranged between 0.791 and 0.813 across all studies.
The pooled data have been reported in Table 26 of the MS, which has been reproduced below (Table 9).

Table 9. Change from baseline EQ-5D utility index score by treatment and visit (pooled data)

Study week	Change from baseline EQ-5D utility index score; mean (95% CI)			
	Placebo (*N*=1286)	Empagliflozin 10mg (*N*=1082)	Empagliflozin 25mg (*N*=1289)	Empagliflozin 10mg and 25mg (*N*=2371)
Week 4	-0.017 (-0.044, 0.010)	-0.011 (-0.035, 0.013)	-0.004 (-0.029, 0.021)	-0.008 (-0.025, 0.009)
Week 6	0.008 (-0.000, 0.017)	0.013 (0.004, 0.021)	0.011 (0.003, 0.019)	0.012 (0.006, 0.018)
Week 12	0.007 (-0.002, 0.015)	0.007 (-0.002, 0.017)	0.013 (0.004, 0.021)	0.010 (0.004, 0.016)
Week 18	0.011 (-0.017, 0.038)	0.020 (-0.002, 0.043)	-0.008 (-0.035, 0.018)	0.006 (-0.012, 0.023)
Week 24	0.0006 (-0.004, 0.015)	0.008 (-0.001, 0.018)	0.008 (-0.001, 0.018)	0.008 (0.002, 0.015)
Week 40	-0.010 (-0.044, 0.025)	-0.003 (-0.028, 0.021)	-0.016 (-0.045, 0.013)	-0.010 (-0.029, 0.009)
Week 52	-0.022 (-0.039, -0.004)	0.014 (-0.010, 0.037)	-0.007 (-0.025, 0.011)	0.000 (-0.014, 0.015)

Source data: Project 0303126: BI Empagliflozin QOL and HCRU Table 5.1

EQ-5D = EuroQoL 5 Dimensions; CI = Confidence Interval

The results above suggest that at 24 weeks, there was no difference between treatments. Similarly, at week 40 and week 52, change in EQ5D from baseline was negligible. In addition to above findings, the manufacturer reported that patients in the trials did not report problems with self-care or usual activities (96% to 97%). Pain/discomfort was the most commonly reported problems. In each trial, almost 31 to 35% of patients had at least moderate pain or discomfort. The proportion of patients with pain/discomfort was slightly higher in the trial which included patients with renal impairment (data not given in the MS). The manufacturer also undertook analysis based on different subgroups and found no difference in EQ-5D utility index scores and EQ-5D VAS between treatment groups.

3.3 Network meta-analysis (NMA)

In absence of head to head comparisons of different flozins, the manufacturer undertook NMA. All the findings related to this have been reported in MS section 6.7 (pages 160 to 199).

The manufacturer undertook this analysis to assess the effectiveness of empagliflozin as dual therapy with metformin, as triple therapy with metformin and SU, as triple therapy with metformin and TZD and finally, as an add-on to insulin. The main comparators were dapagliflozin 10 mg, canagliflozin 100 mg, canagliflozin 300 mg and sitagliptin 100 mg. However, in their network diagrams and tables, the manufacturer also reported studies related to other gliptins.

The manufacturer reported doing systematic searches to identify all the relevant studies. The outcomes analysed were change in HbA1c from baseline, change in SBP from baseline, change in body weight from baseline and safety which included hypoglycaemia (non-severe), hypoglycaemia

(severe), urinary tract infections (UTIs), and genital infections. The outcomes were compared at two time periods 24 ± 4 weeks and 52 ± 4 weeks.

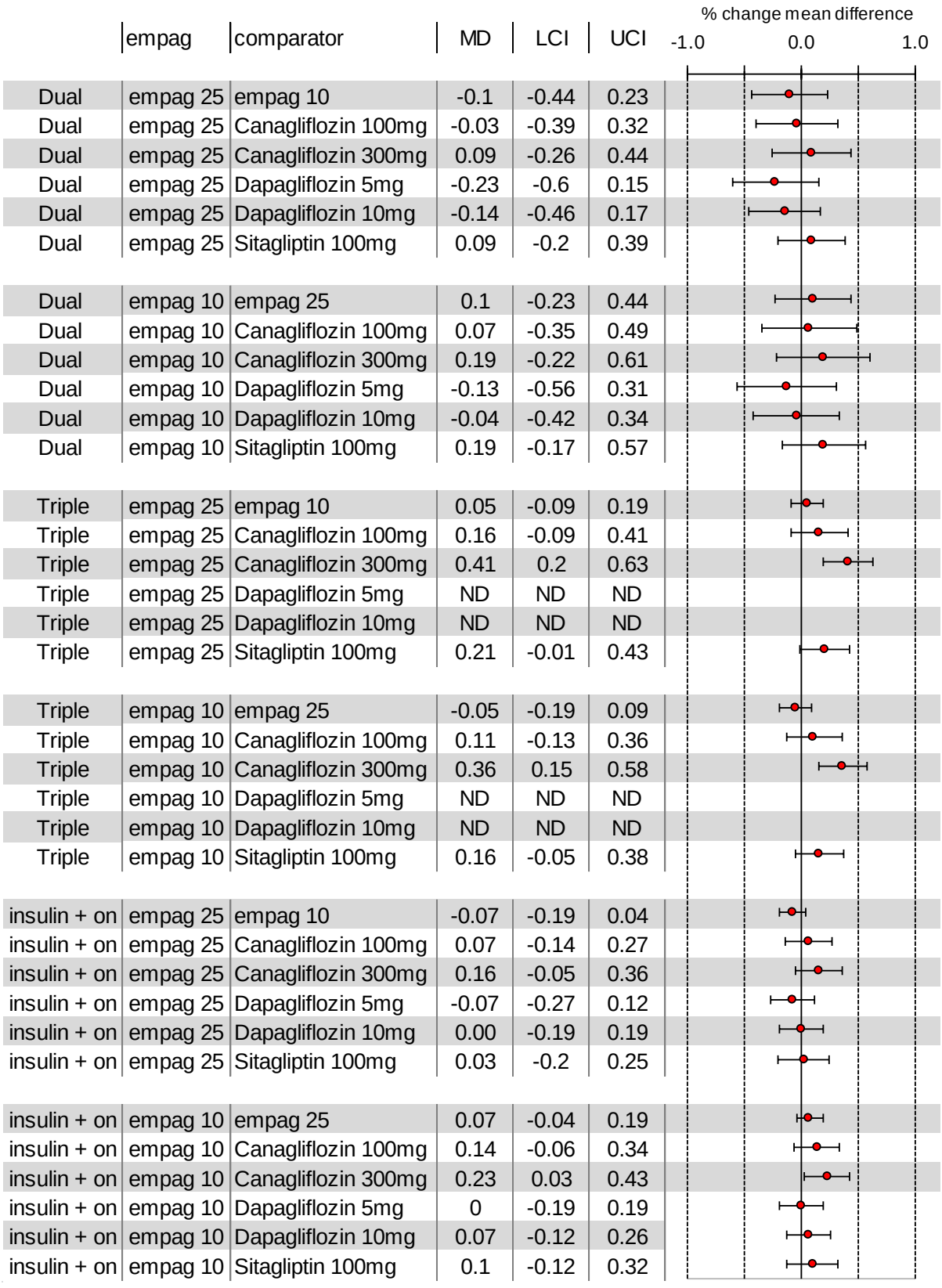
Summary of the manufacturer NMA results

The MS contained 25 pages of NMA results (Tables 31 to 39). For convenience the relevant data from these for key outcomes (change in HbA1c, body weight and SBP) in the comparisons of interest (empagliflozin versus canagliflozin, dapagliflozin and sitagliptin) in dual, triple and insulin add on therapies have been summarised in Figure 11, Figure 12 and Figure 13. These refer to NMA results from data for 24 ± 4 weeks of treatment. Results at 52 ± 4 weeks were based on less evidence and were similar to those for 24 ± 4 weeks.

The percentage change in HbA1c across all NMA comparisons indicated trivial differences between drugs, with 95% credible intervals spanning zero difference; the exception to this generalisation was in the comparison canagliflozin 300 mg versus empagliflozin in which empagliflozin appeared inferior at 1 0mg in both triple and insulin + therapies and at 25 mg in triple therapy. This difference is based on one trial for each comparison. No head to head data was available for empagliflozin versus sitagliptin.

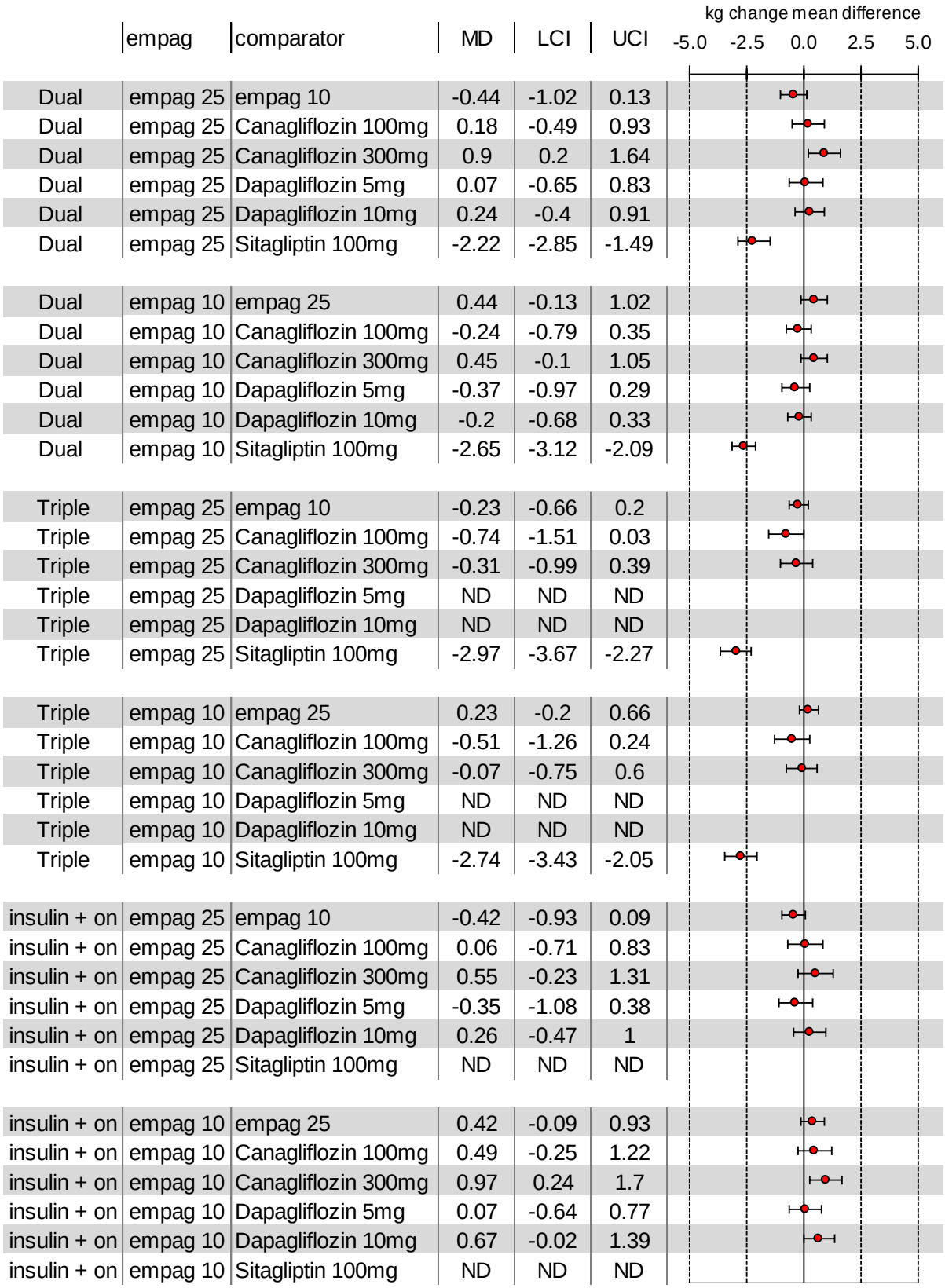
For body weight change with dual and triple therapies NMA indicated empagliflozin was superior to sitagliptin; for this comparison no data was presented for add on to insulin regimens. For all three therapies NMA results were similar for all comparisons between flozins with 95% credible intervals spanned zero difference.

For SBP, where data was available, again the results from the NMAs indicated trivial differences between compared flozins in dual, triple and insulin add on therapies. For these comparisons the 95% credible intervals spanned zero difference. According to NMA results empagliflozin at 10 mg or at 25 mg reduced SBP to a greater extent than did sitagliptin (100 mg) in both dual and triple therapy (no data was available for insulin add on therapy); there was no head to head evidence for this comparison in the network and only a single sitagliptin study contributed evidence.



	empag	comparator	MD	LCI	UCI
Dual	empag 25	empag 10	-0.1	-0.44	0.23
Dual	empag 25	Canagliflozin 100mg	-0.03	-0.39	0.32
Dual	empag 25	Canagliflozin 300mg	0.09	-0.26	0.44
Dual	empag 25	Dapagliflozin 5mg	-0.23	-0.6	0.15
Dual	empag 25	Dapagliflozin 10mg	-0.14	-0.46	0.17
Dual	empag 25	Sitagliptin 100mg	0.09	-0.2	0.39
Dual	empag 10	empag 25	0.1	-0.23	0.44
Dual	empag 10	Canagliflozin 100mg	0.07	-0.35	0.49
Dual	empag 10	Canagliflozin 300mg	0.19	-0.22	0.61
Dual	empag 10	Dapagliflozin 5mg	-0.13	-0.56	0.31
Dual	empag 10	Dapagliflozin 10mg	-0.04	-0.42	0.34
Dual	empag 10	Sitagliptin 100mg	0.19	-0.17	0.57
Triple	empag 25	empag 10	0.05	-0.09	0.19
Triple	empag 25	Canagliflozin 100mg	0.16	-0.09	0.41
Triple	empag 25	Canagliflozin 300mg	0.41	0.2	0.63
Triple	empag 25	Dapagliflozin 5mg	ND	ND	ND
Triple	empag 25	Dapagliflozin 10mg	ND	ND	ND
Triple	empag 25	Sitagliptin 100mg	0.21	-0.01	0.43
Triple	empag 10	empag 25	-0.05	-0.19	0.09
Triple	empag 10	Canagliflozin 100mg	0.11	-0.13	0.36
Triple	empag 10	Canagliflozin 300mg	0.36	0.15	0.58
Triple	empag 10	Dapagliflozin 5mg	ND	ND	ND
Triple	empag 10	Dapagliflozin 10mg	ND	ND	ND
Triple	empag 10	Sitagliptin 100mg	0.16	-0.05	0.38
insulin + on	empag 25	empag 10	-0.07	-0.19	0.04
insulin + on	empag 25	Canagliflozin 100mg	0.07	-0.14	0.27
insulin + on	empag 25	Canagliflozin 300mg	0.16	-0.05	0.36
insulin + on	empag 25	Dapagliflozin 5mg	-0.07	-0.27	0.12
insulin + on	empag 25	Dapagliflozin 10mg	0.00	-0.19	0.19
insulin + on	empag 25	Sitagliptin 100mg	0.03	-0.2	0.25
insulin + on	empag 10	empag 25	0.07	-0.04	0.19
insulin + on	empag 10	Canagliflozin 100mg	0.14	-0.06	0.34
insulin + on	empag 10	Canagliflozin 300mg	0.23	0.03	0.43
insulin + on	empag 10	Dapagliflozin 5mg	0	-0.19	0.19
insulin + on	empag 10	Dapagliflozin 10mg	0.07	-0.12	0.26
insulin + on	empag 10	Sitagliptin 100mg	0.1	-0.12	0.32

Figure 11. Summary of NMA results for mean change (%) in HbA1c at 24 ± 4 weeks (values less than 0 indicates empagliflozin superior)



	empag	comparator	MD	LCI	UCI
Dual	empag 25	empag 10	-0.44	-1.02	0.13
Dual	empag 25	Canagliflozin 100mg	0.18	-0.49	0.93
Dual	empag 25	Canagliflozin 300mg	0.9	0.2	1.64
Dual	empag 25	Dapagliflozin 5mg	0.07	-0.65	0.83
Dual	empag 25	Dapagliflozin 10mg	0.24	-0.4	0.91
Dual	empag 25	Sitagliptin 100mg	-2.22	-2.85	-1.49
Dual	empag 10	empag 25	0.44	-0.13	1.02
Dual	empag 10	Canagliflozin 100mg	-0.24	-0.79	0.35
Dual	empag 10	Canagliflozin 300mg	0.45	-0.1	1.05
Dual	empag 10	Dapagliflozin 5mg	-0.37	-0.97	0.29
Dual	empag 10	Dapagliflozin 10mg	-0.2	-0.68	0.33
Dual	empag 10	Sitagliptin 100mg	-2.65	-3.12	-2.09
Triple	empag 25	empag 10	-0.23	-0.66	0.2
Triple	empag 25	Canagliflozin 100mg	-0.74	-1.51	0.03
Triple	empag 25	Canagliflozin 300mg	-0.31	-0.99	0.39
Triple	empag 25	Dapagliflozin 5mg	ND	ND	ND
Triple	empag 25	Dapagliflozin 10mg	ND	ND	ND
Triple	empag 25	Sitagliptin 100mg	-2.97	-3.67	-2.27
Triple	empag 10	empag 25	0.23	-0.2	0.66
Triple	empag 10	Canagliflozin 100mg	-0.51	-1.26	0.24
Triple	empag 10	Canagliflozin 300mg	-0.07	-0.75	0.6
Triple	empag 10	Dapagliflozin 5mg	ND	ND	ND
Triple	empag 10	Dapagliflozin 10mg	ND	ND	ND
Triple	empag 10	Sitagliptin 100mg	-2.74	-3.43	-2.05
insulin + on	empag 25	empag 10	-0.42	-0.93	0.09
insulin + on	empag 25	Canagliflozin 100mg	0.06	-0.71	0.83
insulin + on	empag 25	Canagliflozin 300mg	0.55	-0.23	1.31
insulin + on	empag 25	Dapagliflozin 5mg	-0.35	-1.08	0.38
insulin + on	empag 25	Dapagliflozin 10mg	0.26	-0.47	1
insulin + on	empag 25	Sitagliptin 100mg	ND	ND	ND
insulin + on	empag 10	empag 25	0.42	-0.09	0.93
insulin + on	empag 10	Canagliflozin 100mg	0.49	-0.25	1.22
insulin + on	empag 10	Canagliflozin 300mg	0.97	0.24	1.7
insulin + on	empag 10	Dapagliflozin 5mg	0.07	-0.64	0.77
insulin + on	empag 10	Dapagliflozin 10mg	0.67	-0.02	1.39
insulin + on	empag 10	Sitagliptin 100mg	ND	ND	ND

Figure 12. Summary of NMA results for mean change (kg) in body weight at 24 ± 4 weeks. (values less than 0 indicates empagliflozin superior)

	empag	comparator	MD	LCI	UCI
Dual	empag 25	empag 10	-0.82	-3.86	2.11
Dual	empag 25	Canagliflozin 100mg	-0.96	-3.88	2.24
Dual	empag 25	Canagliflozin 300mg	0.34	-2.54	3.51
Dual	empag 25	Dapagliflozin 5mg	-2.87	-6.21	0.71
Dual	empag 25	Dapagliflozin 10mg	-1.27	-4.16	1.49
Dual	empag 25	Sitagliptin 100mg	-2.97	-7	1.1
Dual	empag 10	empag 25	0.82	-2.11	3.86
Dual	empag 10	Canagliflozin 100mg	-0.12	-3.74	3.73
Dual	empag 10	Canagliflozin 300mg	1.16	-2.37	5
Dual	empag 10	Dapagliflozin 5mg	-2.06	-5.8	2.09
Dual	empag 10	Dapagliflozin 10mg	-0.44	-3.96	2.98
Dual	empag 10	Sitagliptin 100mg	-2.19	-6.56	2.33
Triple	empag 25	empag 10	0.6	-1.36	2.55
Triple	empag 25	Canagliflozin 100mg	0.24	-3.07	3.46
Triple	empag 25	Canagliflozin 300mg	-0.35	-3.64	2.84
Triple	empag 25	Dapagliflozin 5mg	ND	ND	ND
Triple	empag 25	Dapagliflozin 10mg	ND	ND	ND
Triple	empag 25	Sitagliptin 100mg	-6.31	-10.1	-2.7
Triple	empag 10	empag 25	-0.6	-2.55	1.36
Triple	empag 10	Canagliflozin 100mg	-0.37	-3.63	2.88
Triple	empag 10	Canagliflozin 300mg	-0.94	-4.24	2.29
Triple	empag 10	Dapagliflozin 5mg	ND	ND	ND
Triple	empag 10	Dapagliflozin 10mg	ND	ND	ND
Triple	empag 10	Sitagliptin 100mg	-6.93	-10.61	-3.26
insulin + on	empag 25	empag 10	0.59	-1.12	2.31
insulin + on	empag 25	Canagliflozin 100mg	0.36	-3.08	3.81
insulin + on	empag 25	Canagliflozin 300mg	2.14	-1.34	5.6
insulin + on	empag 25	Dapagliflozin 5mg	0.13	-3.02	3.23
insulin + on	empag 25	Dapagliflozin 10mg	0.88	-2.33	3.98
insulin + on	empag 25	Sitagliptin 100mg	ND	ND	ND
insulin + on	empag 10	empag 25	-0.59	-2.31	1.12
insulin + on	empag 10	Canagliflozin 100mg	-0.26	-3.64	3.16
insulin + on	empag 10	Canagliflozin 300mg	1.55	-1.94	5.01
insulin + on	empag 10	Dapagliflozin 5mg	-0.45	-3.57	2.6
insulin + on	empag 10	Dapagliflozin 10mg	0.27	-2.87	3.38
insulin + on	empag 10	Sitagliptin 100mg	ND	ND	ND

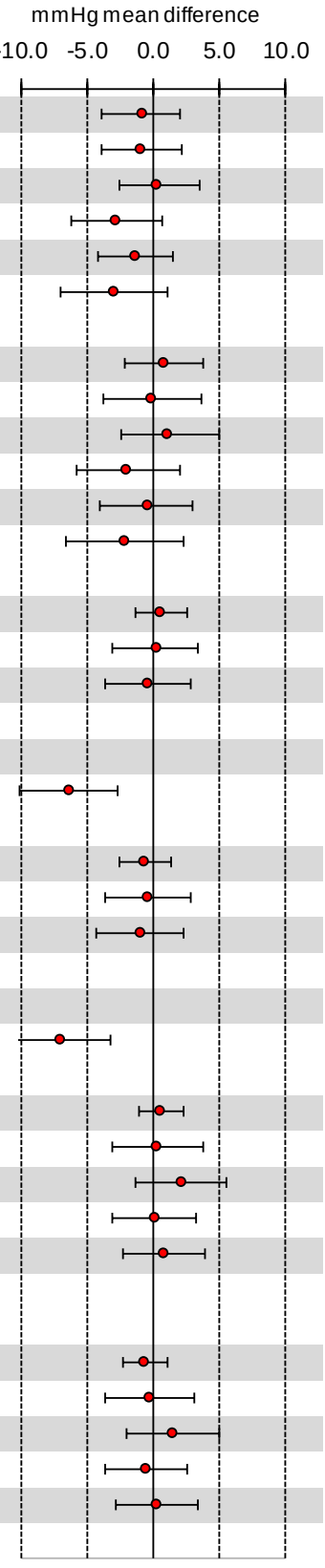


Figure 13. Summary of NMA results for mean change (mmHg) in SBP at 24 ± 4 weeks (values less than 0 indicates empagliflozin superior)

Description and critique of manufacturers approach to evidence synthesis

The MS contained a brief narrative summary of evidence flowing from the empagliflozin RCT programme (MS section 6.10). Assessment of empagliflozin study quality was provided in Appendix 3. ERG considers this material to be balanced; ERG concerns about claims on incidence of genital tract infections (GTIs) were satisfactorily addressed in the manufacturer's clarification response.

The core of the manufacturer's synthesis and assessment of evidence rested heavily on NMAs that deployed 24 ± 4 or 52 ± 4 week outcome data. Five types of NMA were undertaken (represented by MS figures 28 to 32). The three that ERG consider relevant to the decision problem (shown in MS figures 28, 29 and 32) analysed: [a] Metformin based dual therapies in which patients received: Met + a gliptin or Met + a flozin or Met + a SU or Met + placebo; [b] Met + SU based triple therapies in which patients received: Met and SU + a gliptin or + a flozin or + placebo; [c] Insulin based therapies in which patients received add on therapy with a gliptin or a flozin. Separate network diagrams were not provided for 24 and 52 week analyses, however tables for 24 and 52 weeks gave information on the studies used. These would appear appropriate for making clinical effectiveness estimates to be used in an economic model.

The outcomes analysed in NMAs were: a] changes from start of treatment in HbA_{1C}, in SBP and in body weight; b] risk of hypoglycaemia (severe and non-severe), of UTIs and of GTIs. Separate data extraction tables for each outcome according to type of NMA and period of analysis (24 or 52 weeks) were assembled in MS Appendix 7. NMA model codes for WinBUGS software were provided in additional documents rather than MS. The ERG used checklists by Ades et al 2013[31] and Donegan et al 2010[32] to critique the manufacturer's NMA. Please see Table 14 and Table 15 for more details. (Note to NICE; we used both quality assessment tools because we have not yet decided which to use. We would welcome the appraisal committee's view(s) as to which they prefer.)

Potential weaknesses in the implementation of the NMA include:

- Although the search strategy was comprehensive the flow from recovered studies to included studies lacked clarity; e.g. the review process involved for applying study eligibility criteria was unspecified, eligibility criteria for comparators was somewhat ambiguous and did not include SU, the number of specified excluded studies was surprisingly small and the information on reasons for study exclusions lacked detail.

- Details of the data extraction process were not reported. For such a large amount of extracted data some assurance should be provided regarding steps taken to avoid human error.

- There was no assessment or mention of the quality of studies included in the NMA (Appendix 5 was blank). The MS may have been assumed that all RCTs were of acceptable quality but ERG could find no explicit statement to this effect.

- No sensitivity analyses or statistical tests were conducted for any of NMAs undertaken.

- There is some concern regarding the inclusion of studies in the major network for metformin based dual therapies; this is considered in more detail below.

- For the dual therapy NMA there were worrying discrepancies between input data, identification of treatments, and WINBUGS codes. These human errors probably result from insufficient checking during data extraction. Discrepancies encompassed: the exclusion from the WINBUGS code of studies listed as providing relevant data, the inclusion in the WINBUGS code of data from an unlisted and unidentified study, incorrect identification of treatment as saxagliptin rather than sitagliptin, inconsistent use of data from studies listed as providing zero event results for safety outcomes. ERG had insufficient time to check more of the coding, more errors might have been found. Note that not all the WinBUGS coding has been provided.

Network for metformin based dual therapies.

The MS NMA diagram for metformin based dual therapies is shown in Figure 14 (Figure 28 of the MS). There are twenty six studies, including two with 3 arms and two with 4 arms and one head to head trial comparing sitagliptin with canagliflozin (100 mg and 300 mg)

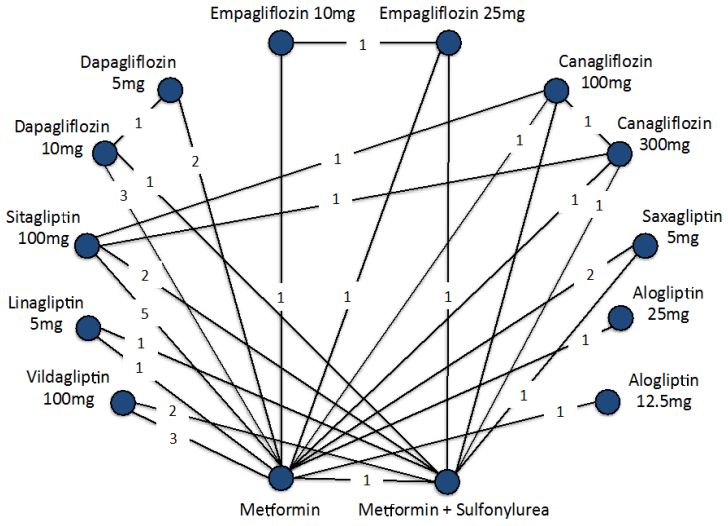


Figure 14. MS dual therapy NMA diagram (metformin background therapy)

In this diagram the Metformin node is “Metformin + placebo” and a line annotated with a single study links this node to the node labelled “Metformin + Sulfonylurea”. ERG found the inclusion of only a single study for this comparison surprising. In clarification the manufacturer stated that this study had been included inappropriately and supplied a new network diagram with named studies annotated as requested by ERG. This is shown below with ERG modifications to increase transparency (Figure 15). As can be seen the study for the link Metformin + placebo to Metformin + Sulfonylurea has been omitted. In clarification the manufacturer stated that this omission had very little effect upon the model output and supplied tables of results using the new network. ERG can confirm by inspection of old and new results tables that the reported results have been minimally influenced by removal of the link study (see above).

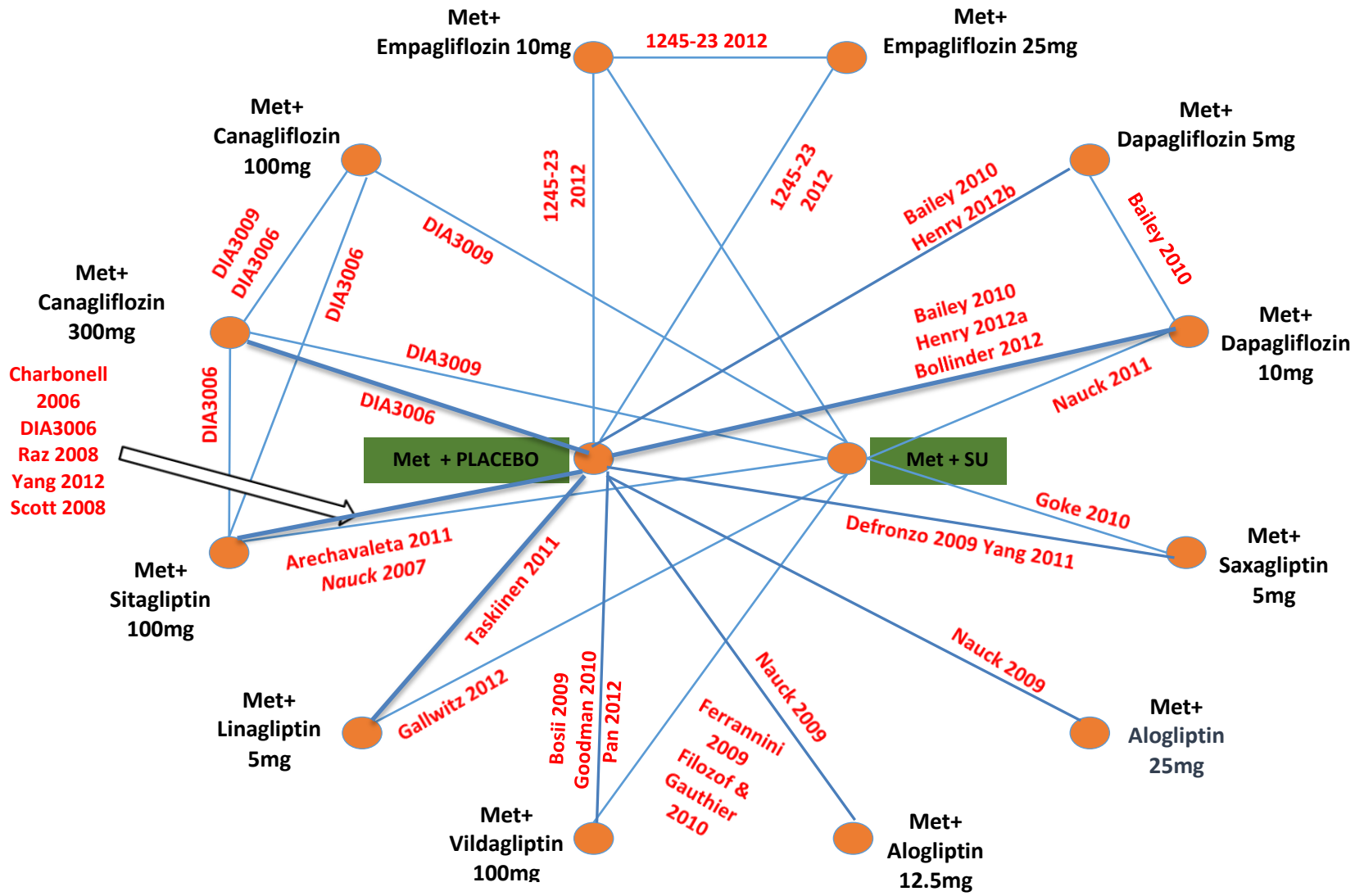


Figure 15. Revised NMA diagram (metformin background therapy)

Because the network for dual therapy was changed in clarification the ERG sought to test the validity of the new results reported in the clarification document. Unfortunately the clarification document did not include new tables showing NMA input data for the new analyses (that is information equivalent to that provided in the original MS appendix 7 tables 115 to 160); furthermore the new WINBUGS codes were not supplied in clarification.

The network shown in Figure 2 includes the study by Scott et al. 2008 (reference 54 in the MS) for the comparison sitagliptin 100mg + metformin versus metformin + placebo. This study was not shown in the manufacturer's clarification diagram but was listed as an included study in both the clarification and MS Appendix 7 data extraction tables. According to these this study provided evidence for the following safety outcomes (Table 127 MS): overall hypoglycaemia, non-severe hypoglycaemia, and severe hypoglycaemia. Unfortunately in this table the sitagliptin arm is coded as a saxagliptin arm; however within the WINBUGS code supplied the outcome data from this trial has been omitted altogether. Further problems encountered in checking safety outcomes are described below.

The ERG checked WinBUGS codes for three outcomes i.e. overall hypoglycaemia (file named as 'Second Hypo RE'), non-severe hypoglycaemia (file named as 'Second Non Sev Hypo RE') and urinary tract infection (UTI) (file named as 'Second UTI RE'). The ERG found that some studies were excluded in the WinBUGS analysis (Please see Table 10, Table 11 and Table 12 below). The ERG also noticed that the following data were included in the analysis, however, this has not reported in the MS or clarification document.

Studies	**In original MS**	**Arm1**	**Arm2**	**Arm 3**	**Arm 4**	**WinBUGS**
Unknown	Not listed	3.5/104	1.5/213	0.5/210	-	Included

On checking the codes assigned to these treatment arms, the ERG can confirm that the study compared met + placebo vs. met + alogliptin 12.5 mg vs. met + alogliptin 25 mg.

Table 10. Overall hypoglycaemia (dual therapy at 24 weeks) – file named 'Second Hypo RE'

Studies	In original MS	Arm1	Arm2	Arm 3	Arm 4	WinBUGS
Forst et al 2010	Listed	0/71	3/65	0/66	-	Not included
Charpentier et al. 2001	Listed	11/75	22/147	-	-	Included
Feinglos et al 2005	Listed	2/61	9/61	-	-	Included
BI GMBH (1245-23) 2012	Listed	1/225	4/224	3/217		Included
Scott et al. 2008	Listed	2/92	1/94	-	-	Not Included
Charbonnel et al. 2006	Listed	5//237	6/464	-		Included
Defronzo et al. 2009	Listed	9/179	10/191	-		Included
Raz et al. 2008	Listed	0/94	1/96	-		Not included
Taskinen et al. 2011	Listed	5/177	3/523			Included
Yang et al. 2011	Listed	4/287	4/283			Included
Yang et al. 2012	Listed	3/198	1/197	-		Included
BI GMBH (1245-28) 2013	Listed	165/780	15/765	-		-
Arechavaleta et al. 2011	Listed	114/518	36/516	-		-
Gallwitz et al. 2012	Listed	280/775	58/776	-		-
Goke et al. 2010	Listed	156/430	13/428	-		-
Nauck et al. 2007	Listed	187/584	29/588	-		-
Bosi et al. 2007	Listed	0/182 [used as 0.5]	1/185 [used as 1.5]			Included
Ferrannini et al 2009	Listed	224/1393	23/1396	-		-
Filozof et al 2010	Listed	11/494	6/513	-		-
Goodman et al 2009	Listed	0/122 **[used as 0.5]**	2/248 **[used as 2.5]**	-		Included
Nauck et al 2006	Listed	0/36 **[used as 0.5]**	1/36 **[used as 1.5]**	-		Included
Pan et al 2012	Listed	0/144	0/148	-	-	Not included
Janssen (DIA3009) 2013	Listed	165/482	27/483	24/485		-
Janssen (DIA3006) 2013	Listed	-	-	-		
Bolinder et al 2012	Listed	3/91	2/91	-	-	Included
Henry et al 2012a	Listed	6/208	7/211	-	-	Included
Henry et al 2012b	Listed	0/201 **[used as 0.5]**	5/194 **[used as 5.5]**	-	-	Included
Bailey et al 2010	Listed	4/137	5/137	5/135	-	Included
Nauck et al 2011	Listed	162/408	14/406	-	-	-

**Highlighted studies compares Met + SU against Met + active drug. No placebo arm, therefore these studies were not included in the analysis*

Table 11. Non severe hypoglycaemia (dual therapy at 24 weeks) – file named 'Second Non Sev Hypo RE'

Studies	In original MS	Arm1	Arm2	Arm 3	WinBUGS
Forst et al 2010	Listed	0/71	3/65	0/66	Not included
Charpentier et al. 2001	Listed	11/75	20/147	-	Included
Feinglos et al 2005	Listed	2/61	9/61	-	Not included
BI GMBH (1245-23) 2012	Listed	1/206	4/217	3/214	Included
Scott et al. 2008	Listed	2/92	1/94	-	Not included
Charbonnel et al. 2006	Listed	5/237	6/464	-	Included
Defronzo et al. 2009	Listed	9/179	10/191	-	Included
Raz et al. 2008	Listed	0/94	1/96	-	Not included
Taskinen et al. 2011	Listed	5/177	3/523		Included
Yang et al. 2011	Listed	4/287	4/283		Included
Yang et al. 2012	Listed	3/198	1/197	-	Included
BI GMBH (1245-28) 2013	Listed	164/780	15/765	-	-
Arechavaleta et al. 2011	Listed	111/518	35/516	-	-
Gallwitz et al. 2012	Listed	268/755	57/764	-	-
Goke et al. 2010	Listed	149/430	13/428	-	-
Nauck et al. 2007	Listed	180/584	28/588	-	-
Bosi et al. 2007	Listed	0/182 [used as 0.5]	1/185 [used as 1.5]		Included
Ferrannini et al 2009	Listed	214/1393	23/1396	-	-
Filozof et al 2010	Listed	-	-	-	-
Goodman et al 2009	Listed	0/122 [used as 0.5]	2/248 [used as 2.5]	-	Included
Nauck et al 2006	Listed	0/36 [used as 0.5]	1/36 [used as 1.5]	-	Included
Pan et al 2012	Listed	0/144	0/148		Not included
Janssen (DIA3009) 2013	Listed	-	-	-	-
Janssen (DIA3006) 2013	Listed	-	-	-	-
Bolinder et al 2012	Listed	2/91	2/91		Included
Henry et al 2012a	Listed	6/208	7/211		Included
Henry et al 2012b	Listed	0/201 [used a 0.5]	5/194 [used as 5.5]		Included
Bailey et al 2010	Listed	4/137	5/137	5/135	Included
Nauck et al 2011	Listed	147/408	7/406	-	-

**Highlighted studies compares Met + SU against Met + active drug. No placebo arm*

Table 12. UTI (dual therapy at 24 weeks) – file named 'Second UTI RE'

Studies	In original MS	Arm1	Arm2	Arm 3	WinBUGS
Forst et al 2010	Listed	1/65	0/71	0/66	Not included
Charpentier et al. 2001	Listed	-	-	-	
Feinglos et al 2005	Listed	-	-	-	
BI GMBH (1245-23) 2012	Listed	8/214	9/206	9/217	Included
Scott et al. 2008	Listed	-	-	-	
Charbonnel et al. 2006	Listed	13/237	22/464	-	Included
Defronzo et al. 2009	Listed	8/179	10/191	-	Included
Raz et al. 2008	Listed	3/94	4/96	-	Not included
Taskinen et al. 2011	Listed	7/177	16/523		Included
Yang et al. 2011	Listed	8/287	13/283		Included
Yang et al. 2012	Listed	-	-	-	
BI GMBH (1245-28) 2013	Listed	67/780	81/765	-	-
Arechavaleta et al. 2011	Listed	-	-	-	-
Gallwitz et al. 2012	Listed	1/755	1/764	-	-
Goke et al. 2010	Listed	-	-	-	-
Nauck et al. 2007	Listed	-	-	-	-
Bosi et al. 2007	Listed	1/185	0/185		Included
Ferrannini et al 2009	Listed	-	-	-	-
Filozof et al 2010	Listed	-	-	-	-
Goodman et al 2009	Listed	-	-	-	-
Nauck et al 2006	Listed	-	-	-	-
Pan et al 2012	Listed	0/148	1/148		Included
Janssen (DIA3009) 2013	Listed	-	-	-	-
Janssen (DIA3006) 2013	Listed	-	-	-	
Bolinder et al 2012	Listed	0/91	2		Included
Henry et al 2012a	Listed	4/211	6		Included
Henry et al 2012b	Listed	10/194	10/194		Included
Bailey et al 2010	Listed	-	-		
Nauck et al 2011	Listed	-	-	-	-

**Highlighted studies compares Met + SU against Met + active drug. No placebo arm*

The ERG re-ran WinBUGS models for the three outcomes by including data of all the excluded studies (except those highlighted in the table above). The table below compares relative risk (RR) (95% CrI) reported by the manufacturer and that obtained by the ERG (Table 13).

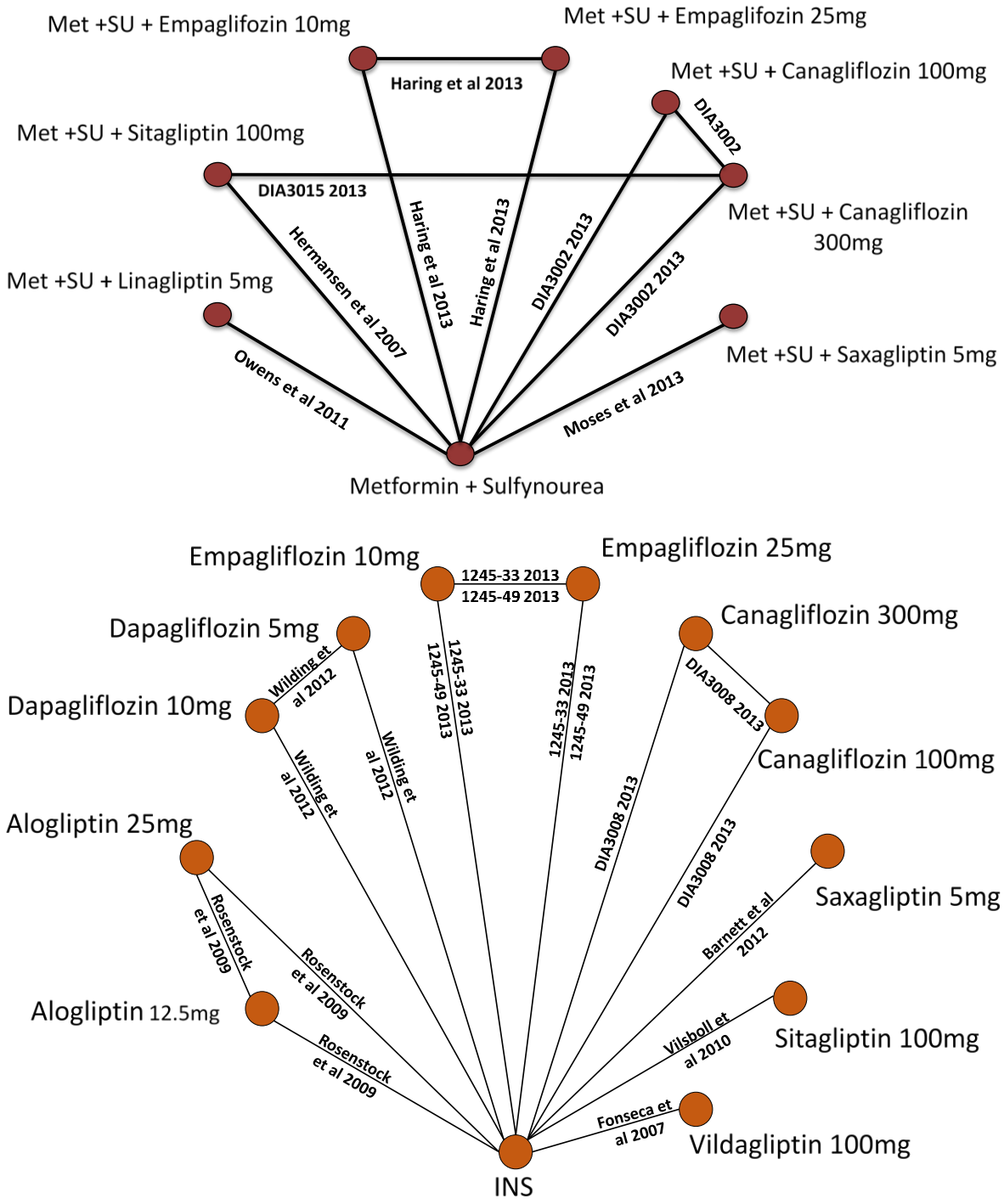
Table 13. Empagliflozin vs. Sitagliptin (RR; 95% CrI)

Outcomes	MS (Appendix 7) (Empagliflozin 10 mg)	MS (Appendix 7) (Empagliflozin 25 mg)	Clarification (Table 20) (Empagliflozin 10 mg)	Clarification (Table 20) (Empagliflozin 25 mg)	ERG (Empagliflozin 10 mg)	ERG (Empagliflozin 25 mg)
Overall hypoglycaemia	0.12 (0.01 to 1.61)	0.16 (0.01 to 2.51)	0.1 (0.01 to 1.44)	0.15 (0.01 to 2.38)	0.14 (0.01 to 1.94)	0.17 (0.01 to 2.8)
Non Severe hypoglycaemia	0.12 (0.01 to 1.64)	0.17 (0.01 to 2.55)	0.1 (0.01 to 1.77)	0.14 (0.01 to 2.67)	0.13 (0.01 to 2.18)	0.18 (0.01 to 3.71)
UTI	0.76 (0.02 to 37.8)	0.79 (0.02 to 37.6)	0.76 (0.02 to 37.8)	0.79 (0.02 to 37.6)	0.81 (0.05 to 21.8)	0.90 (0.05 to 23.03)

The dual therapy network contains many treatments (i.e. gliptins) outside of the comparisons defined for the decision problem (i.e. sitagliptin versus flozins in dual therapy with metformin for patients no longer responding adequately to metformin). In clarification the manufacturer offered two justifications for the inclusion of extra gliptin treatments; firstly they could act as a validity check since, as a class, they would be expected to generate similar clinical results to the comparator of interest (i.e. sitagliptin) and secondly the inclusion of additional trials of similar designs investigating various gliptins adds data points for the estimation of between-trial heterogeneity. While ERG agrees these are reasonable aims the ERG was unable to find further reference in the submission or clarification documents that addressed either of these issues.

The ERG does not think it appropriate to exclude studies linking Met + placebo to Met + SU; furthermore the original inclusion of only a single study for this link may have been inappropriate since the ERG has found several relevant systematic reviews that list several studies of Metformin versus SU, (e.g. McIntosh et al 2011[33] list 3 studies only one of which was that identified by the manufacturer). Inclusion of such studies may influence model output values for Metformin + placebo. This value is fed into the economic model for dual therapy under the heading "Baseline" (MS Table 56 and MS Table 59). As a criticism of the manufacturer's presentation it should be noted that the baseline values (i.e. Metformin + placebo) were not included in the NMA results tables.

Figure 16. shows MS networks for triple therapy (upper) and insulin add-on therapy (lower).



The triple therapy network encompassed six studies including one head to head dual arm trial (Met + SU + 100 mg sitagliptin versus Met + SU + canagliflozin 300 mg), and two triple arm studies a] Met + SU + placebo versus Met + SU + canagliflozin 100 mg or 300 mg, and b] Met + SU + placebo versus Met + SU + empagliflozin at 10 mg or 25 mg (Figure 16). The three dual arm studies compared Met + SU + placebo with Met + SU + saxagliptn or + linagliptin or sitagliptin.

The NMA assumes equivalence of Met + SU and Met + SU + placebo; in the economic model (Table 57) the value for "HbA1c Baseline (Met + SU) -0.135" comes from a NMA in which Met + SU is assumed equal to Met + SU + placebo.

Following clarification responses, the ERG used the NICE DSU RE model for multi-arm trials to analyse mean change in HbA1c in dual therapy at 24 weeks while omitting data for Charpentier et al. 2001 to see if the results were consistent with those included in the response document. The ERG got very similar results to that included in the clarification document.

Eight studies were included in the insulin add on network. All compared insulin + placebo with insulin + a gliptin or insulin + a flozin. There were no head to head trials. Three trials were dual arm, four trials had three arms, and a fourth had four arms but only three of these were included in the network. The "baseline risk (insulin)" value entered into the economic model (MS Table 58) is derived from Insulin + placebo arms only.

Table 14 and Table 15 reports the ERG's critical appraisal of the manufacturer's NMA. In general the overall methods in the NMA appeared to be of reasonable quality, but in those sections of results that the ERG had sufficient time to check there were errors that may be symptomatic of more general deficiencies. The clarity of NMA reporting was less good.

Table 14. Critique of the manufacturer NMA using checklist by Ades et al 2013[31]

		Item Satisf-actory?	Comments
A.	**DEFINITION OF THE DECISION PROBLEM**		
A1.	**Target Population for Decision**		
A1.1	Has the target patient population for decision been clearly defined?	✓	Yes. Patients with T2DM who require additional therapy to background therapy of Met or Met + SU or, or insulin
A2.	**Comparators**		
A2.1	Decision comparator set: Have all the appropriate treatments in the decision been identified?	✓	Yes; sitagliptins and other flozins
A2.2	Synthesis comparator set: Are there additional treatments in the synthesis comparator set that are not in the decision comparator set? If so, is this adequately justified?	✓	Yes, there are other comparators added to the networks that are not relevant to the decision problem. Comparison of flozins against sitagliptin would have been adequate. However, the manufacturer has included other gliptins and for dual therapy also SU. The manufacturer offered justifications which, in ERGs judgement, were reasonable.
A3.	**Trial Inclusion/Exclusion**		
A3.1	Is the search strategy technically adequate and appropriately reported?	✓	Yes. Section 6.1; appendix 2 gives details of their search strategy
A3.2	Have all trials involving at least 2 of the treatments in the synthesis comparator set been included?	✓	Yes It appears probable that all the relevant trials pertaining to the decision problem have been included in the NMAs. However the selection process applied was not described.
A3.3	Have all trials reporting relevant outcomes been included?	✓	Yes
A3.4	Have additional trials been included? If so, is this adequately justified?		Yes, trials not relevant to decision problem have been included and justified
A4.	**Treatment Definition**		
A4.1	Are all the treatment options restricted to specific doses and co-treatments, or have different doses and co-treatments been "lumped" together? If the latter, is it adequately justified?	✓	Yes. Different doses of the treatment have been compared. The manufacturer has not lumped together different doses of treatment
A4.2	Are there any additional modelling assumptions?	✓	Yes. To accommodate the inclusion of data from multiple studies, the manufacturer has assumed that data within a window of 24 ± 4 weeks were equivalent; similar data within the range 52 ± 4 weeks were assumed equivalent. The following rationale for a further assumption was stated in the clarification document: *The rationale for assuming that Met = Met + placebo and Met + Su = Met + SU + Placebo was that it has been accepted in previous reviews by both CADTH and NICE (Canagliflozin STA) and therefore, was*

			considered appropriate here as well.
A5.	**Trial Outcomes and Scale of Measurement Chosen for the Synthesis**		
A5.1	Where alternative outcomes are available, has the choice of outcome measure used in the synthesis been justified?	x	NA
A5.2	Have the assumptions behind the choice of scale been justified?	X	NA
A6.	**Patient Population: Trials with Patients outside the Target Population**		
A6.1	Do some trials include patients outside the target population? If so, is this adequately justified?	x	Not clear. The manufacturer has been explicit in terms of patient population they want to include in their NMA i.e. background therapy of different oral hypoglycaemic drugs or, insulin however, table 29 and also tables included in the clarification response document do not clearly indicate patients background treatments across all included studies
A6.2	What assumptions are made about the impact or lack of impact this may have on the relative treatment effects? Are they adequately justified?	x	NA No description
A6.3	Has an adjustment been made to account for these differences? If so, comment on the adequacy of the evidence presented in support of this adjustment and on the need for a sensitivity analysis.	x	NA No description
A7.	**Patient Population: Heterogeneity within the Target Population**		
A7.1	Have potential modifiers of treatment effect been considered?	x	No. The NMA has been performed in patients receiving different background therapy
A7.2	Are there apparent or potential differences between trials in their patient populations, albeit within the target population? If so, has this been adequately taken into account?	x	The NMA has been performed in patients receiving different background therapy but a full assessment of potential differences in trial populations has not been considered
A8.	**Risk of Bias**		
A8.1	Is there a discussion of the biases to which these trials, or this ensemble of trials, are vulnerable?	x	No. Appendix 5 only shows a blank template for the quality assessment of RCTs
A8.2	If a bias risk was identified, was any adjustment made to the analysis and was this adequately justified?	x	No. Appendix 5 only shows a blank template for the quality assessment of RCTs
A9.	**Presentation of the Data**		
A9.1	Is there a clear table or diagram showing which data have been included in the base-case analysis?	x	The tables depicting NMA results and the background treatment lacked clarity but were improved in response to clarification queries.
A9.2	Is there a clear table or diagram showing which data have been excluded and why?	x	No. There is no table or diagram showing which data have been excluded. However, Appendix 6 shows list of studies excluded, however in this the reasons for exclusion lack clarity
B.	**METHODS OF ANALYSIS AND PRESENTATION OF RESULTS**		
B1.	**Meta-Analytic Methods**		
B1.1	Is the statistical model clearly described?	✓	In section 6.7.5, the manufacturer has briefly explained the methods used to undertake their

			NMA.
B1.2	Has the software implementation been documented?	x	No The manufacturer has not explicitly mentioned in their MS that WinBUGS has been used to undertake their NMA.
B2.	**Heterogeneity in the Relative Treatment Effects**		
B2.1	Have numerical estimates been provided of the degree of heterogeneity in the relative treatment effects?	x	No The manufacturer mentioned that a fixed effect model was used for analysis where only one study contributed data for a particular intervention while, random effect model to analyse data if there were more than one studies. The ERG thinks the manufacturer could have selected the appropriate model by running some goodness-of-fit test for e.g. comparing DIC results (the model with the lower DIC value is considered to be a parsimonious model; the difference of less than 5 between the two models is insignificant).
B2.2	Has a justification been given for choice of random or fixed effect models? Should sensitivity analyses be considered?	x	Please see above
B2.3	Has there been adequate response to heterogeneity?	x	Please see above
B2.4	Does the extent of unexplained variation in relative treatment effects threaten the robustness of conclusions?	x	Please see above
B2.5	Has the statistical heterogeneity between baseline arms been discussed?	x	Please see above
B3.	**Baseline Model for Trial Outcomes**		
B3.1	Are baseline effects and relative effects estimated in the same model? If so, has this been justified?	x	Some lack of clarity, see A4.2.
B3.2	Has the choice of studies to inform the baseline model been explained?	x	No No adequate description
B4.	**Presentation of Results of Analyses of Trial Data**		
B4.1	Are the relative treatment effects (relative to a placebo or ‘‘standard’’ comparator) tabulated, alongside measures of between study heterogeneity if an RE model is used?	x	No The manufacturer has not presented the findings as suggested in the DSU document.
B4.2	Are the absolute effects on each treatment, as they are used in the CEA, reported?	x	No Not reported for all outcomes. The manufacturer provided absolute effects on each treatment, except placebo, in the clarification document
B5.	**Synthesis in Other Parts of the Natural History Model**		
B5.1	Is the choice of data sources to inform the other parameters in the natural history model adequately described and justified?	x	NA
B5.2	In the natural history model, can the longer-term differences between treatments be explained by their differences on randomized trial outcomes?	x	NA

C.	**ISSUES SPECIFIC TO NETWORK SYNTHESIS**		
C1.	**Adequacy of Information on Model Specification and Software Implementation**	x	No As mentioned previously, the manufacturer has not explicitly reported that WinBUGS was used to undertake their NMA. The manufacturer mentioned the choice of their model i.e. FE or RE model based on the number of studies. However, the ERG feel that the choice of model should have also been determined by running good of fit tests for e.g. DIC
C2.	**Multiarm Trials**		
C2.1	If there are multiarm trials, have the correlations between the relative treatment effects been taken into account?	✓	There are multi arm trials in all NMAs reported. The manufacturer had provided all the WinBUGS codes for their original NMA. The ERG checked the codes to see if cumulative adjustments for multi-arm trials were done. The manufacturer has appropriately done this for random effect models.
C3.	**Connected and Disconnected Networks**		
C3.1	Is the network of evidence based on randomized trials connected?	✓	Yes All the included are RCTs
C4.	**Inconsistency**		
C4.1	How many inconsistencies could there be in the network?	x	No In Figure 28, the manufacturer should have included a study comparing met + placebo and met + SU. In the original submission, Charpentier et al 2001 was included that compared these two arms however, in their clarification responses this study was excluded. The manufacturer explained that this study was erroneously included in the main submission therefore it was excluded from their revised NMA. The ERG found several studies including systematic reviews comparing these two treatments. The ERG thinks studies comparing these two treatment arms should haven included in their analysis including in the network diagram for consistency. Other network diagrams seems to be consistent.
C4.2	Are there any a priori reasons for concern that inconsistency might exist, due to systematic clinical differences between the patients in trials comparing treatments A and B, the patients in trials comparing treatments A and C, and so on?	x	No There is insufficient information about trial populations to address question.
C4.2	Have adequate checks for inconsistency been made?	x	No
C4.4	If inconsistency was detected, what adjustments were made to the analysis, and how was this justified?		NA
D.	**EMBEDDING THE SYNTHESIS IN A PROBABILISTIC COST-EFFECTIVENESS ANALYSIS**		
D1.	**Uncertainty Propagation**		
D1.1	Has the uncertainty in parameter estimates	x	The CE model makes multiple two way

	been propagated through the CEA model?		comparisons between treatments and the results were made probabilistic using random sampling procedure which appears may not to have been conducted correctly. (see CE section of the ERG report)
D2.	**Correlations**		
D2.1	Are there correlations between parameters? If so, have the correlations been propagated through the CEA model?	x	Yes
Mark ✓ to indicate that the issue has been addressed satisfactorily and if there is any cause for concern on the item. The Comments column should be used to answer the question (YES, NO, NA: not applicable) and/or to spell out the reasons for any concerns, the need for sensitivity analyses and so on.			

Table 15. Critique of the manufacturer NMA using checklist by Donegan et al 2010[32]

	Yes/No/Unclear/Not applicable
Indirect comparison method	
Is the method applied to undertake the indirect comparison adequate?	Yes
If an adequate method is used, is a treatment effect estimate and measure of precision reported?	Yes; mean and 95% CrI are reported
Similarity	
Is the assumption of similarity stated?	No
Is a method described to assess the similarity assumption within the review methods section?	No
Is a reasonable approach used to assess the assumption of similarity?	Unclear
Are patient or trial characteristics reported for all trials in the indirect comparison?	Unclear [to some extent; baseline body weight and HbA1c values would have been appropriate]
Are patient or trial characteristics compared across the two trial sets involved in the indirect comparison?	No
Are patient or trial characteristics reported to be comparable for the two trial sets involved in the indirect comparison?	Unclear [Implied not explicit]
Homogeneity across trials within each of the two trials set involved in the indirect comparison	
Is the method used to determine the presence of statistical homogeneity adequate?	Unclear [Unknown]
Is the homogeneity assumption satisfied or is statistical heterogeneity accounted for if present?	Unclear [Unknown]
If the homogeneity assumption is not satisfied, is clinical or methodological homogeneity across trials in each trial set involved in the indirect comparison investigated by an adequate method?	Unclear [Unknown]
Consistency	
Is consistency of effects assessed?	No [Implied in results from class flozins and gliptins]
If the direct and indirect evidence is reported to be consistent, is the evidence combined and the result presented?	No [Not reported]
If consistency is reported, is this accounted for by not combining the direct and indirect evidence?	Not applicable
Are patient or trial characteristics compared between direct and indirect evidence trials?	No
Are any included 3-arm trials correctly analysed?	Yes
Is justification given for using indirect evidence and direct evidence?	Yes
Does the review present results from all trials providing direct evidence?	Yes
Interpretation	
Is a distinction made between direct and indirect comparisons?	No
Does the review state that more trials providing direct evidence as needed?	No [could not find the statement]
Reporting	
Does the review present both of the meta-analysis results from each of the two trial sets involved in the indirect comparison?	No
Was it highlighted which results were from indirect evidence?	No
Are the individual trials treatment effect estimates reported?	Yes

As previously mentioned, it has been reported by Polidori and colleagues that canagliflozin has a dual action on both SGLT1 and SGLT2, and it has been suggested that this might make its glucose-lowering action greater than flozins without SGLT1 activity. This suggestion was raised at the NICE scoping meeting on empagliflozin. Prior to this appraisal, a group at Warwick Medical School carried out an independent NMA of dapagliflozin, canagliflozin and empagliflozin in dual therapy. This

NMA (academic in confidence till publication) found no clinically significant differences between empagliflozin and canagliflozin.

Proportion of patients completing the trials

In ER Met 92.8% of patients completed the 24 weeks, with slightly more discontinuations in the placebo arm (10%) than in the empagliflozin 10 mg (4%) and 25 mg (8%) arms. The reasons for discontinuation included adverse events (AEs – 2.2%) and refusal to continue treatment (2%).

In ER MetSU 91.3% of patients completed the 24 week period with slightly more not completing the trial in the placebo group than in the empagliflozin 10 mg and 25 mg group (10.7% vs. 7.6% vs. 7.9%) The main reason for discontinuation was AEs (3.6% in placebo vs. 2.7% in empagliflozin 10 mg vs. 3.2% in empagliflozin 25 mg). There were two more patients (0.9%) in the placebo group who discontinued due to lack of efficacy.

A total of 360/494 patients completed the 76 weeks period in the basal insulin study (1245.33). The proportion of patients discontinuing from the study was greater in the placebo group (30.6%) than in the empagliflozin group (22.5% in the empagliflozin 10 mg vs. 28.4% in the empagliflozin 25 mg). Patients discontinued because of AEs (10% overall), other reasons (5.3% overall), lost to follow-up (4.3% overall) and refusal to continue treatment (3.6% overall; the manufacturer has stated that this is not due to AEs). The proportion of patients withdrawing from the study due to AEs and other reasons was slightly greater in the empagliflozin 25 mg group than in the empagliflozin 10 mg or the placebo group (12.9% vs. 11.2% vs. 7.6%). In contrast, patients lost to follow-up or refusing treatment were fewer in the empagliflozin 25 group than in the empagliflozin 10 mg or the placebo group (n=6 vs. 4 for lost to follow-up; n=8 vs. 1 for refusal of treatment) (details in figure 9 of the MS).

In the MDI insulin study (1245.49), a total of 475/563 (84.4%) patients completed the 52 week treatment period. Patients discontinued the trial mainly because of AEs (5%) and refusal to treatment (4.6%). Other reasons included protocol non-compliance, and lost to follow-up.

Adverse events (AEs)

The manufacturer presented information about adverse events in section 6.9 (pages 202 to 217) of the MS. The evidence for this comes from eight trials, the results from these studies have been combined and presented below (Table 16). Table 44 of the MS reports adverse events of the extension study (ER EXTEND). The relevant results have been added to the table below. The ERG also checked published studies to see if any additional information was available.

Table 16. Adverse events from the relevant trials

	Empagliflozin 10 mg	Empagliflozin 25 mg	Placebo
One or more drug-related AEs			
ER Pio	24/165 (14.5)	31/168 (18.5)	31/165 (18.8)
ER Met 24 weeks	35/217 (16.1%)	27/214 (12.6%)	25/121 (12.1%)
ER Met EXTEND	66/217 (30.4%)	43/214 (20.1%)	46/206 (22.3%)
ER MetSU 24 weeks	54/224 (24.1%)	43/217 (19.8%)	34/225 (15.1%)
ER MetSU EXTEND	80/224 (35.7%)	69/217 (31.8%)	59/225 (26.2%)
ER SU	-	190/765 (24.8)	252/780 (32.3)
ER Basal	65/169 (38.5%)	68/155 (43.9%)	52/170 (30.6%)
ER MDI	56/186 (30.1%)	76/189 (40.2%)	64/188 (34.0%)
ER Renal	37/98 (37.8)	101/321 (31.5)	87/319 (27.3)
ER BP	55/276 (19.9)	54/276 (19.6)	21/272 (7.7)
AEs leading to discontinuation			
ER Pio	2/165 (1.2)	5/168 (3.0)	4/165 (2.4)
ER Met 24 weeks	2/217 (0.9%)	5/214 (2.3%)	7/206 (3.4%)
ER Met EXTEND	7/217 (3.2%)	12/214 (5.6%)	10/206 (4.9%)
ER MetSU 24 weeks	6/224 (2.7%)	7/217 (3.2%)	8/225 (3.6%)
ER MetSU EXTEND	10/224 (4.5%)	15/217 (6.9%)	16/225 (7.1%)
ER SU	-	39/765 (5.1)	34/780 (4.4)
ER Basal	19/169 (11.2%)	20/155 (12.9%)	13/170 (7.6%)
ER MDI	10/186 (5.4%)	9/189 (4.8%)	9/188 (4.8%)
ER Renal	4/98 (4.1)	21/321 (6.5)	1/3197 (5.3)
ER BP	4/276 (1.4)	6/276 (2.2)	5/272 (1.8)
One or more serious AEs			
ER Pio	7/165 (4.2)	6/168 (3.6)	7/165 (4.2)
ER Met 24 weeks	7/217 (3.2%)	5/214 (2.3%)	7/206 (3.4%)
ER Met EXTEND	19/217 (8.8%)	17/214 (7.9%)	24/206 (11.7%)
ER MetSU 24 weeks	11/224 (4.9%)	1/217 (0.5%)	14/225 (6.2%)
ER MetSU EXTEND	29/224 (12.9%)	24/217 (11.1%)	31/225 (13.8%)
ER SU	-	119/765 (15.6)	89/780 (11.4)
ER Basal	28/169 (16.6%)	28/155 (18.1%)	28/170 (16.5%)
ER MDI	20/186 (10.8%)	22/189 (11.6%)	22/188 (11.7%)
ER Renal	4/98	21/321	44/319

	(4.1)	(6.5)	(13.8)
ER BP	3/276 (1.1)	4/276 (1.4)	7/272 (2.6)
UTI (belonging to BICMQ*)			
ER Pio	28/165 (17.0)	20/168 (11.9)	27/165 (16.4)
ER Met 24 weeks	11/217 (5.1%)	12/214 (5.6%)	10/206 (4.9%)
ER Met EXTEND	31/217 (14.3%)	22/214 (10.3%)	28/206 (13.6%)
ER MetSU 24 weeks	23/217 (10.3%)	18/214 (8.3%)	18/206 (8.0%)
ER MetSU EXTEND	38/224 (17%)	35/217 (16.1%)	36/225 (16%)
ER SU	-	105/765 (13.7)	102/780 (13.1)
ER Basal	25/169 (14.8%)	18/155 (11.6%)	15/170 (8.8%)
ER MDI	29/186 (15.6%)	29/189 (15.3%)	29/188 (15.4%)
ER Renal	14/98 (14.3)	47/321 (14.6)	47/319 (14.7)
ER BP	11/276 (4.0)	13/276 (4.7)	10/272 (3.7)
Genital infection (belonging to BICMQ)			
ER Pio	14/165 (8.5)	6/168 (3.6)	4/165 (2.4)
ER Met 24 weeks	8/217 (3.7%)	10/214 (4.7%)	0
ER Met EXTEND	18/217 (8.3%)	20/214 (9.3%)	1/206 (0.5%)
ER MetSU 24 weeks	6/217 (2.7%)	5/214 (2.3%)	2/206 (0.9%)
ER MetSU EXTEND	10/224 (4.5%)	13/217 (6.0%)	2/225 (0.9%)
ER SU	-	90/765 (11.8)	17/780 (2.2)
ER Basal	13/169 (7.7%)	8/155 (5.2%)	3/170 (1.8%)
ER MDI	8/186 (4.3%)	18/189 (9.5%)	3/188 (1.6%)
ER Renal	7/98 (7.1)	11/321 (3.4)	8/319 (2.5)
ER BP	14/276 (5.1)	15/276 (5.4)	1/272 (0.4)

AEs = adverse events; BICMQ = BI-customised medical dictionary for drug regulatory activities (MedDRA) query; UTI = urinary tract infection

The proportion of patients reporting one or more serious AEs (there was no description within the MS and published studies to confirm which events were considered to be serious) in the ER Met and ER MetSU studies was higher in the placebo group than in the two empagliflozin groups (ER Met: 3.4% vs. 3.2% with 10 mg vs. 2.3% with 25 mg; ER MetSU: 6.2% vs. 4.9% with 10 mg vs. 0.5% with 25 mg). The findings were similar in the extension studies (ER Met EXTEND: 11.7% with placebo vs. 8.8% with 10 mg vs. 7.9% with 25 mg; ER MetSU EXTEND: 13.8% with placebo; 12.9% with 10 mg; 11.1% with 25 mg). In the ER Basal study, there were slightly more patients reporting serious AEs in the 25 mg empagliflozin group compared to other treatment groups (18.1% vs. 16.6% with 10 mg vs. 16.5% with placebo). There was no difference across the groups in the ER MDI study (11 to 12%). There were no deaths in any study.

In the ER Met and ER MetSU studies, at 24 weeks the proportion of patients reporting one or more drug-related AEs was slightly higher in the 10 mg empagliflozin group than in the 25 mg empagliflozin group [ER Met: 16.1% vs. 12.6%; ER MetSU: 24.1% vs. 19.8%]. Details of events related to the study treatment have not been provided in the MS. The findings were similar in the extension study (ER Met EXTEND: 30.4% vs. 20.1%; ER MetSU EXTEND: 35.7% vs. 31.8%). In contrast in the two insulin studies, the proportion of patients reporting one or more drug-related AEs was higher in the 25 mg empagliflozin group than in the 10 mg empagliflozin group [ER Basal: 43.9% vs. 38.5%; ER MDI: 40.2% vs. 30.1%].

The proportion of patients discontinuing treatment due to AE was higher in the placebo group (3.4%) than in the two empagliflozin groups (0.9% with 10 mg; 2.3% with 25 mg) in the ER Met study and in the ER MetSU study (3.6% with placebo; 2.7% with 10 mg; 3.2% with 25 mg). There were no details in the MS as to what events led to discontinuation. However, Haring et al (2014) reported that two patients (one each in empagliflozin treatment groups) discontinued treatment due to genital infection.[28] None of the patients discontinued treatment due to UTI. On comparing two doses of empagliflozin in the extension study, more patients in the empagliflozin 25 mg group discontinued treatment due to AEs (ER Met EXTEND: 5.6% vs. 3.2%; ER MetSU EXTEND: 6.9% vs. 4.5%). The findings in the two insulin studies were mixed. In the ER Basal study, more patients in the empagliflozin 25 mg group (12.9%) discontinued study treatment than in the empagliflozin 10 mg group (11.2%) and placebo (7.6%). The number of patients discontinuing study treatment in the ER MDI study was slightly more in the lower dose empagliflozin group (5.4% vs. 4.8% with 25 mg and placebo).

The most commonly reported adverse events were urinary tract infection (UTI) and genital infection. At 24 weeks, the proportion of patients reporting UTI (mostly mild in intensity)(Haring et al 2014)[28] was similar across all groups (4.9% in placebo; 5.1% in empa 10 mg; 5.6% in empa 25 mg)) in the ER Met study. UTIs were more common in females, with lower rates in placebo group (7.7%) and very similar rates in the two empagliflozin groups (12% in empa 10 mg and 11.8% in empa 25 mg). UTIs were less frequent in males (2.6% in placebo; 0% in empa 10 mg and 0.8% in empa 25 mg). The authors report that most of the UTIs were mild in intensity and none were severe in intensity causing patients to discontinue study treatment. There were also no cases of urosepsis or pyelonephritis.[28] In the ER Met EXTEND study, more patients in the lower dose of empagliflozin group had UTI (14.3% vs. 10.3%, p value not reported). In the ER MetSU study and the extension study (ER MetSU EXTEND), the proportion of patients with UTI was slightly greater with the lower dose of empagliflozin (ER MetSU: 10.3% vs. 8.3%; ER MetSU EXTEND: 17% vs. 16.1%). Similarly, in the ER Basal study, the proportion of patients with UTI was higher in the 10 mg empagliflozin group

(14.8% vs. 11.6%). The proportion of patients affected by UTI was similar across all treatment groups in the ER MDI study (15.3% to 15.6%).

The proportions of patients complaining of genital infection were greater, but usually only slightly with the higher dose of empagliflozin the ER Met (4.7% on 25mg vs. 3.7% on 10mg); ER Met EXTEND studies (9.3% vs. 8.3%); ER MDI study (9.5% with 25 mg vs. 4.3% with 10 mg); ER MetSU EXTEND study (6.0% vs. 4.5%). In contrast, the proportion of patients with genital infection was slightly greater in the lower dose of empagliflozin in the two remaining studies [ER MetSU: 2.7% vs. 2.3%; ER Basal: 7.7% vs. 5.2%]. P values were not reported.

Other adverse events reported in the trials were hypoglycaemia, nasopharyngitis, upper respiratory tract infection (URTI), hyperglycaemia, diarrhoea and volume depletion. The incidence of hypoglycaemia in the ER Met study ranged between 0.5% and 1.8. Not surprisingly, numbers of patients reporting were slightly higher in the ER MetSU study [8.4% in placebo; 16.1% in 10 mg; 11.5% in 25 mg] but this would be due to the sulphonylurea. In the ER Met EXTEND study, hypoglycaemia was similar in the two empagliflozin groups (4.1% in 10 mg; 4.2% in 25 mg) while, slightly higher in the low dose empagliflozin group in the ER MetSU EXTEND study (23.7% vs. 19.4%). The incidence of hypoglycaemia was higher in the insulin studies but there was no difference between the two empagliflozin groups in the ER Basal study (36.1% in both) and only slight difference in the 25 mg empagliflozin group in the ER MDI study (57.7% vs. 51.1%).

Adverse events compared to other flozins

In absence of head to head comparisons of different flozins, the manufacturer has undertaken a network meta-analysis to compare the safety data. The outcome measures compared included hypoglycaemia (non-severe), hypoglycaemia (severe), UTIs and genital infections. The critique of the manufacturer's network meta-analysis by the ERG has been done in section 3.3 (page 47) of this report.

The results of NMA for use of empagliflozin in dual therapy at 24 weeks ± weeks and as triple therapy at 24 ± 4 weeks have been given in table 31 (pages 175 to 179) and 33 (pages 185 to 188) of the MS. The ERG has reproduced safety data below (Table 17 and Table 18).

Table 17. NMA results for the comparative safety of empagliflozin versus other flozins in dual therapy (24 ± 4 weeks)

Overall hypoglycaemia, RR (95% credible interval)					
intervention/ comparator	**Empagliflozin 10mg**	**Empagliflozin 25mg**	**Canagliflozin 100mg**	**Canagliflozin 300mg**	**Dapagliflozin 10 mg**
Empagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***	***NA***
Empagliflozin 25mg	0.77 (0.11 to 4.60)	***NA***	***NA***	***NA***	***NA***
Canagliflozin 100mg	***NA***	***NA***	***NA***	***NA***	***NA***
Canagliflozin 300mg	***NA***	***NA***	***NA***	***NA***	***NA***
Dapagliflozin 10 mg	0.28 (0.03 to 3.04)	0.34 (0.04 to 4.88)	***NA***	***NA***	***NA***
Hypoglycaemia (non-severe), RR (95% credible interval)					
intervention/ comparator	**Empagliflozin 10mg**	**Empagliflozin 25mg**	**Canagliflozin 100mg**	**Canagliflozin 300mg**	**Dapagliflozin 10 mg**
Empagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***	***NA***
Empagliflozin 25mg	0.71 (0.11 to 4.55)	***NA***	***NA***	***NA***	***NA***
Canagliflozin 100mg	***NA***	***NA***	***NA***	***NA***	***NA***
Canagliflozin 300mg	***NA***	***NA***	***NA***	***NA***	***NA***
Dapagliflozin 10 mg	0.33 (0.03 to 4.21)	0.48 (0.04 to 6.63)	***NA***	***NA***	***NA***
UTIs, RR (95% credible interval)					
intervention/ comparator	**Empagliflozin 10mg**	**Empagliflozin 25mg**	**Canagliflozin 100mg**	**Canagliflozin 300mg**	**Dapagliflozin 10 mg**
Empagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***	***NA***
Empagliflozin 25mg	0.96 (0.06 to 15.2)	***NA***	***NA***	***NA***	***NA***
Canagliflozin 100mg	***NA***	***NA***	***NA***	***NA***	***NA***
Canagliflozin 300mg	***NA***	***NA***	***NA***	***NA***	***NA***
Dapagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***	***NA***
Genital infection, RR (95% credible interval)					
NA					

NA=Not available

From Table 17, it can be seen that the probability of overall hypoglycaemia was lower with the 25 mg empagliflozin than with the 10 mg empagliflozin but not significantly so (RR 0.77 95% CrI 0.11 to 4.60). The comparison of dapagliflozin 10 mg against two doses of empagliflozin suggested that the probability of hypoglycaemia was lower with dapagliflozin but CrI were wide). The findings were similar for non-severe cases of hypoglycaemia. There was no difference between the two doses of empagliflozin in terms of UTIs (RR 0.96 95% CrI 0.06 to 15.2).

Table 18. NMA results for the comparative safety of empagliflozin versus other flozins in triple therapy (24 ± 4 weeks)

Hypoglycaemia, RR (95% credible interval)				
intervention/ comparator	**Empagliflozin 10mg**	**Empagliflozin 25mg**	**Canagliflozin 100mg**	**Canagliflozin 300mg**
Empagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***
Empagliflozin 25mg	0.77 (0.47 to 1.22)	***NA***	***NA***	***NA***
Canagliflozin 100mg	1.35 (0.70 to 2.63)	1.76 (0.87 to 3.57)	***NA***	***NA***
Canagliflozin 300mg	1.64 (0.88 to 3.12)	**2.15 (1.11 to 4.22)**	1.22 (0.84 to 1.79)	***NA***
Hypoglycaemia (non-severe), RR (95% credible interval)				
intervention/ comparator	**Empagliflozin 10mg**	**Empagliflozin 25mg**	**Canagliflozin 100mg**	**Canagliflozin 300mg**
Empagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***
Empagliflozin 25mg	0.77 (0.47 to 1.22)	***NA***	***NA***	***NA***
Canagliflozin 100mg	1.35 (0.70 to 2.63)	1.76 (0.87 to 3.57)	***NA***	***NA***
Canagliflozin 300mg	1.64 (0.88 to 3.12)	**2.15 (1.11 to 4.22)**	1.22 (0.84 to 1.79)	***NA***
UTIs, RR (95% credible interval)				
intervention/ comparator	**Empagliflozin 10mg**	**Empagliflozin 25mg**	**Canagliflozin 100mg**	**Canagliflozin 300mg**
Empagliflozin 10 mg	***NA***	***NA***	***NA***	***NA***
Empagliflozin 25mg	0.80 (0.43 to 1.47)	***NA***	***NA***	***NA***
Canagliflozin 100mg	***NA***	***NA***	***NA***	***NA***
Canagliflozin 300mg	***NA***	***NA***	***NA***	***NA***
Genital infection, RR (95% credible interval)				
NA				

In triple therapy, the probability of hypoglycaemia including non-severe hypoglycaemia was non-significantly lower with 25 mg than with 10 mg empagliflozin (RR 0.77 95% CrI 0.47 to 1.22). The probability of hypoglycaemia (including non-severe) was higher with canagliflozin 100 mg (RR 1.35 vs. 10 mg empa; RR 1.76 vs. 25 mg empa) and 300 mg (RR 1.64 against 10 mg empa; RR 2.15 vs. 25 mg empa) compared against the two doses of empagliflozin, but credible intervals were again wide. The probability of UTI was non-significantly lower with 25mg than with 10 mg empagliflozin (RR 0.80, 95% CrI 0.43 to 1.47).

Conclusion

The findings suggest that patients taking empagliflozin are at increased risk of UTI and genital infections, with similar incidences in the two doses. The proportion of patients discontinuing treatment due to AEs was similar in the two empagliflozin groups (empa 10 mg ranged between 0.9% and 11.2%; empa 25 mg 2.2% and 12.9%), but higher than placebo (between 1.8% and 7.6%). There were no reports of deaths in the trials. There was no difference in the incidence of hypoglycaemia between the two doses of empagliflozin. However, the incidence of hypoglycaemia was higher in those on SU.

4 MODEL IMPLEMENTATION CROSS-CHECK

4.1 Introduction

The Empagliflozin Cost Effectiveness Model (ECEM) consists of;

- an Excel front end that acts as a store of parameter values
- visual basic (VB) code that forms the actual model
- an Excel back end for the outputting of the visual basic results.

This chapter briefly describes the ECEM. It then reviews the model implementation within the VB. The ERG has tried to provide sufficient examples of the VB to support the arguments raised and to aid the manufacturer in assessing whether the arguments raised are valid. But this does lead to a rather technical read that will not be accessible to all. We therefore start with a summary of the main ERG concerns, which provides the essential details. Most readers will move on the next chapter at this point, but full details of the ERG concerns follow.

Due to the extent and complexity of the VB code, the ERG has only cross checked some elements of it. This has focussed upon the areas that appear most problematic. As a consequence, the ERG does not warrant that the rest of the ECEM works as it should.

4.2 Summary: major issues with the ECEM implementation

While the ERG has a number of concerns with the ECEM implementation, which if confirmed by the manufacturer could be quite serious, the ERG is of the opinion that all are genuine errors and are not by design.

- The ECEM appears to sample one set of random numbers at the start of each probabilistic sensitivity analysis (PSA) iteration. These are used to determine whether events occur for each patient simulated within a PSA iteration. The same set of random numbers appears to be used for each patient. As a consequence, within a PSA iteration two patients who are identical at baseline will be simulated as having the same set of events at the same time and as a consequence the same net costs and net quality adjusted life years (QALYs). This sampling is a key consideration within an individual patient model. If the ERG identified error is confirmed by the manufacturer, this would seem to largely invalidate the modelling of the submission.
- The ECEM modelling of HBA1c does not appear to be in line with the UKPDS 68 and may be too aggressive, also resulting in it converging at too high a level. Since treatment changes, the complications of diabetes and some deaths are dependent upon HbA1c, if there is an error here it would be quite serious and could again largely invalidate the submitted results.

- The ECEM modelling of SBP may also be incorrect. Without any change of treatment it appears to model another step during the third cycle that is not warranted. It also appears that the calculation thereafter may not be in line with the UKPDS 68.
- The ECEM modelling of the lipids ratio may also be incorrect. The value for the third cycle may be incorrect, with this flowing through to the calculation of subsequent values. Similar errors to those identified in the modelling of systolic blood pressure (SBP) may apply to the modelling of the lipids ratio.
- Only 100 patients are simulated, and only 300 PSA iterations are run. These are unusually low numbers. The manufacturer accepts that the results for a single PSA iteration will not have converged over 100 patients. Given the concerns around random sampling of the first bullet, it is also not clear that the central estimate of the PSA will have converged after 300 patients. Even within the current modelling, variability of results appears to remain when moving from 250 to 300 PSA iterations.
- The results of the ECEM are sensitive to the seeds chosen for the random numbers. This sensitivity appears to mainly relate to the random number seed that is used for the 2nd order sampling and populating the matrix of random numbers that determines whether an event occurs or not.
- It appears that the quality of life detriment associated with weight gains may only be applied to the weight change during a cycle rather than to the difference between the weight and the baseline weight. This may tend to underestimate the importance of weight changes for a given disutility per body mass index (BMI) point. Given the centrality of the direct impact of weight changes upon quality of life in previous NICE assessments of treatments for T2DM, if this is an error it could bias the results against the treatment with the better weight profile. It would disadvantage the flozins.
- It also appears that weight gains above the baseline weight are asymmetrically handled. If a treatment is associated with an initial weight gain, further weight gains due to natural history appear to have a disutility applied. If a treatment is associated with an initial weight loss, further weight gains due to natural history that increase the patient weight to be above their baseline weight do not appear to have a disutility applied even though weight is above the initial baseline. This might tend to bias the analysis against treatments that are associated with an initial weight gain, such as insulin and pioglitazone, and favour those with an initial weight loss such as the flozins.
- The ECEM calculation of the total QALYs may have only applied half the overall QALY decrements associated with adverse events and the complications of diabetes. This possible error was only identified shortly before the deadline for the submission of the ERG report,

and as a consequence has not been fully parsed by the ERG. But if it is confirmed as an error by the manufacturer it would largely invalidate the ECEM modelling.

- Given the major issues and some of the minor issues, providing that the manufacturer does not refute all the main ERG concerns, it appears that the ECEM development has not involved sufficient validation and stress testing. The ERG has not cross checked the implementation of all the ECEM visual basic code. As a consequence, even if the manufacturer agrees with some or all of the major issues identified by the ERG and fixes them, there must remain some doubt that the remainder of the ECEM works as intended, as it should and provides robust cost effectiveness estimates.

Note that the issues around the handling of weight changes have only been checked by the ERG by an inspection of the visual basic code. They have not been subject to a detailed ERG rebuild and cross check of this with the interim outputs of the ECEM, because this is not possible within the time constraints of an STA.

Minor issues with the ECEM implementation

- There is uncertainty as to how to handle the adjustment to age that is required for implementing some of the equations of the UKPDS 68; e.g. the Gompertz of the mortality equation 9. How this is handled can have a large impact upon the modelled probability of mortality. Other adjustments such as that applied in the dapagliflozin assessment could have a major impact upon results. The ERG is of the opinion that the adjustment made by the manufacturer is the most reasonable given current information.
- For reasons that are unclear, the ECEM does not use the UKPDS 68 calculation of an annual probability of an event and then adjust this to a six-monthly probability. It bases most of its calculations upon a period spanning three cycles. The two are not equivalent and there is no obvious reason for the approach of the manufacturer. But the discrepancies between the six-monthly probabilities that result appear likely to be quite small.
- The sampling of 2nd order uncertainty around treatment effects adds the sampled relative value for e.g. sitagliptin to the mean value for empagliflozin. It may be more correct to add the sampled relative value for sitagliptin to the sampled value for empagliflozin. The ECEM may systematically underestimate 2nd order uncertainty as a result, though this may reflect a lack of understanding of the NMA on the part of the ERG.
- Despite an ERG clarification question it remains unclear how the values for the covariances associated with the UKPDS 68 parameters have been derived.

- The ECEM simulates some patients as dying twice. While this may be a relatively rare occurrence, it does highlight a possible lack of validation and stress testing during the ECEM development.
- Similar to the random sampling of numbers for the assessment of whether events happen, it appears that the ransom sequencing of the assessment of events is dependent upon a matrix of sampled values that is established at the start of each PSA iteration. As a consequence, it appears that the sequence in which events are assessed may be the same for each patient within a PSA iteration.
- The *sampled_value_beta_adj* procedure outlines that there may be an issue with the beta sampling of the model, which in unusual circumstances may simulate a negative α value. But the ERG has not identified any examples of this occurring, or for the *sampled_value_beta_adj* procedure to be required, within the ECEM as submitted. The *sampled_value_beta_adj* procedure would require correction to avoid bias.

Full details

4.3 Model structure

The main model structure is drawn from the equations of the UKPDS 68. Figure 33 of the submission presents a graphical representation of the ECEM, and is reproduced below for ease of reference.

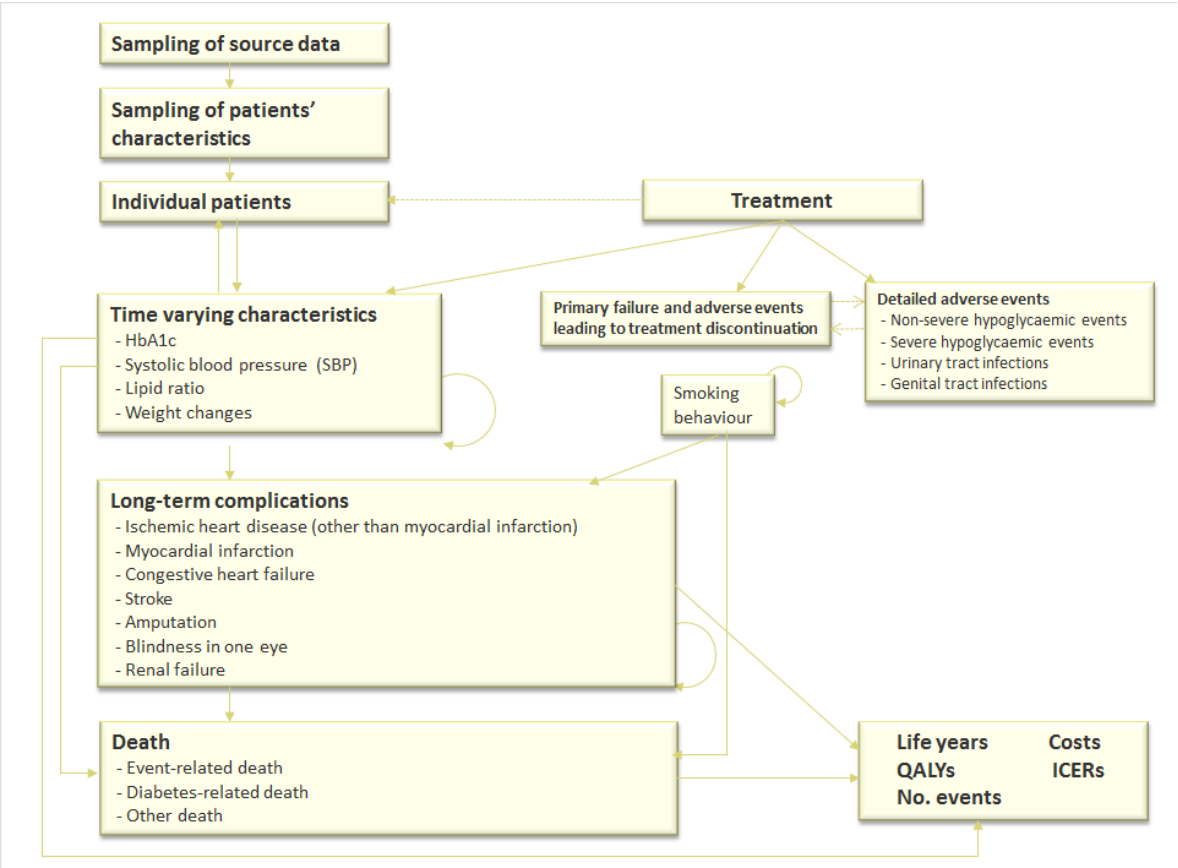


Figure 17. ECEM structure

Within the taxonomy of the NICE DSU technical support document 15,[34] the ECEM is a patient level state transition model. That is to say it models individual patients' transitions between health states using a fixed cycle length.

The cycle length of the model is six months. This is apparently mainly to enable patients to switch treatments at the six month point. Note that the UKPDS 68 estimates annual probabilities of transitions. The calculation of the six monthly probabilities of the ECEM is not entirely in line with the calculation of the annual probabilities of the UKPDS 68, but the two are closely aligned.

The time horizon of the model is 40 years, the perspective is that of the NHS and PSS for costs and patients for benefits, and costs and benefits are discounted at an annual 3.5%.

The sampling of the source data relates to 2nd order uncertainty. If the ECEM is being run probabilistically it samples the input parameters from their distributions; e.g. the UKPDS 68 regression coefficients. These values are used for one PSA iteration of the ECEM. The ECEM then samples another set of input parameters for the next PSA iteration. If the model is being run deterministically these parameters are not sampled and their mean values are used. Note that the

ECEM has the facility to turn off this sampling for subsets of parameters; e.g. utilities, while still running through a number of PSA iterations.

The sampling of patient characteristics relates to 1st order uncertainty, and establishes the patient characteristics of those modelled by the ECEM. These patient baseline characteristics cover both what the ECEM refers to as patient baseline characteristics and the patient disease history.

The detailed adverse events are external to the UKPDS 68: urinary tract infections (UTIs) and genital tract infections (GTIs), and non-severe and severe hypos. The impacts of UTIs and GTIs are assumed to only be felt during the first cycle. The impacts of the hypos are felt for the duration of treatment.

Given the parameterisations of the submitted model, a patient may stop their initial treatment for two reasons: discontinuations and treatment failures.

- Discontinuations occur during the first cycle and lead to the patient switching to another therapy of the same line. When this occurs the treatment efficacy of the initial treatment is retained.
- Treatment failures relate to when the patient's HbA1c has risen above 7.5%. At this point the patient is assumed to switch to insulin therapy, and to experience the clinical effects associated with it in terms of changes to HbA1c, SBP, lipids and weight. Note that this does not include the option of a long-acting glucagon like peptide-1 (GLP-1) analogue such as exenatide which results in weight loss, as well the need to inject once a week rather than every day.

After an initial treatment effect over six months and a short period of stability, typically one year, the evolutions of the main risk factors; HbA1c, SBP, lipids and smoking status, are modelled based upon the equations of table 4 of the UKPDS 68.

The evolution of BMI is more involved. Treatments modelled as causing an initial weight loss are assumed to reach that weight loss after 2 six month cycles. The weight loss is then maintained for one six month cycle, and then lost over another six month cycle.(Note that this does not match weight loss duration in the extension studies.) Thereafter an annual gain of 0.1 kg is assumed to occur for the time the patient remains on treatment. Treatments modelled as causing an initial weight gain are assumed to reach that weight gain after 2 six month cycles. Thereafter an annual natural history gain of 0.1 kg is assumed to occur for the time the patient remains on treatment. At switch of therapy due to treatment failure the intention of the ECEM is to model patients as reverting to the weight that natural history would have placed them at.

During each cycle of the ECEM, given the patient characteristics and risk factors the probability of experiencing one of the seven complications of diabetes; ischaemic heart disease (IHD), myocardial infarction (MI), congestive heart failure (CHF), stroke, amputation, blindness in one eye and/or renal failure is modelled based upon the equations of table 2 of the UKPDS 68. Due to the ECEM being a state transition model rather than a discrete event simulation, the ordering of the assessment of whether an event occurs is in part randomised.

During each cycle of the ECEM, given the patient characteristics and risk factors the probability of dying is similarly calculated from the equations of table 3 of the UKPDS 68.

Costs and quality of life values are attached to the health states of the model, these also being largely drawn from UKPDS publications. The exceptions to this are those arising from detailed adverse events and those arising from changes in patient weight. The direct quality of life impacts from changes in weight are applied to treatments which cause an initial weight gain, but they are not applied to treatments which cause an initial weight loss, which may disadvantage the flozins. The submitted models do not associate weight changes with any direct costs. Weight gain is associated with an increase in daily insulin dose so if weight gain was considerable, insulin costs would rise.

The ECEM has been constructed to only permit a total of 30,000 patients to be simulated. Within a deterministic analysis this permits a sufficiently large number of patients to be simulated. But within a probabilistic analysis only 30,000 patients can be simulated in total, not per PSA iteration. As a consequence, if 300 PSA iterations are performed only 100 patients can be run within each PSA iteration. The values reported within the submission relate to 100 patients being run through 300 PSA iterations.

4.4 ERG cross check of the ECEM implementation

The ERG has reviewed the visual basic of the model. There is a great deal of this. Some of it is redundant code. The ERG has not been able to comprehensively assess the visual basic code, but has traced a number of visual basic elements through the model, some in detail, some less so. To date a number of issues have arisen. Unless refuted by the manufacturer these could largely invalidate the cost effectiveness estimates of the submission.

As already noted, the ERG has not had time to cross check in detail the workings of all the visual basic code that makes up the ECEM. The ERG focus has also been on the elements that may be problematic.

A detailed cross check of the code involves not only an inspection of the logic of the visual basic code, but also outputting the input values and output values for a section of code to excel, rebuilding the logic of the visual basic code within excel and cross checking that the values in the excel rebuild correspond with the values outputted from the visual basic code.

Elements of the code that appear to work as intended have not all been subject to a detailed cross check. Some have only had their visual basic inspected. The following outlines those elements of the model that the ERG thinks do not work as intended, and those that the ERG thinks do work as intended. Those that the ERG thinks do not work as intended are further subdivided into those that appear to be major issues which if confirmed as errors by the manufacturer may invalidate the ECEM results, and those that appear to be minor issues.

4.5 Major Issues with the ECEM implementation

4.5.1 Random sampling for evaluation of events within a PSA iteration

Most economic models submitted for the STA process are patient cohort models rather than individual patient models, and some brief background about the methodological differences between the two modelling approaches may be helpful before reviewing the random sampling in the model.

To simplify matters, suppose that the condition under consideration was prostate cancer and that in the first cycle of the model there was a 20% chance of dying. This estimate of 20% is subject to 2nd order uncertainty, and probabilistic modelling would sample this for each PSA iteration. But suppose that we are only concerned with the deterministic modelling, and so want to apply the 20% central estimate.

Within a patient cohort model a cohort of perhaps 1,000 would be simulated. This would typically be 1,000 identical patients such as 1,000 men age 73 at baseline. In the first cycle the 20% would be applied to this 1,000, resulting in 200 patients being simulated as dying during the first cycle and 800 patients being simulated as surviving the first cycle.

An individual patient model could also be used to simulate the same 1,000 identical patients age 73 at baseline. But the method is different. The first patient is simulated by drawing a random number for him. If this random number is less than or equal to 20% he is simulated as dying during the first cycle. If not, he is simulated as surviving the first cycle. The model then moves on to the second patient, repeating the process for this patient. Another random number is drawn for the second patient and this determines if he is simulated as dying or as surviving during the first cycle. The model works through each patient individually in like fashion, keeping a tally of the number that have died and the number

that have survived. Provided that enough patients are simulated the model will converge around 20% of the individual patients being simulated as dying during the first cycle and 80% of patients being simulated as surviving the first cycle.

The key point in the above is that a different random number is drawn for each patient when determining whether they are simulated as dying or surviving the first cycle. If a single random number was drawn and then applied to each of the 1,000 patients, if it was less than or equal to 20% all 1,000 patients would be simulated as dying during the first cycle. If it was greater than 20% all 1,000 patients would be simulated as surviving the first cycle. The method requires that a different random number is drawn for each patient.

Unfortunately, within the ECEM it appears that for a given PSA iteration the same set of random numbers is applied to each patient. Given the centrality of this to any individual patient modelling the ERG has cross checked this in three ways.

- Examination of the visual basic code
- Simulation of 100 identical patients with no sampling of 2nd order uncertainty
- Simulations involving sampling of 1st order and 2nd order uncertainty with additional code inserted into the visual basic to output the calculated probabilities of events and random numbers these are evaluated against by PSA iteration, patient number and cycle number.

These cross checks all support the view that for a given PSA iteration the same set of random numbers is applied to each patient. These cross checks are outlined in more detail below.

There are two procedures that generate the random number sequences used in the model:

- *GenerateRandomValuesSample*
- *GenerateRandomValuesPat*.

GenerateRandomValuesSample is run once and only once for each PSA iteration and takes the index of PSA iterations as its parameter value. The matrices defined and populated in this procedure will remain constant for a PSA iteration; i.e. they will not change per patient.

The following outlines how the matrix *RandMatComplicat* is defined and populated in the procedure *GenerateRandomValuesSample* .

1. The procedure redefines the matrix row / column lengths for all random number matrices.
    a. Note: we are assuming that *CEM_Planning* = 1 (as hard coded in the VB code) therefore the matrix *RandomMatrix1* will have 61 columns.
    b. The number of rows in *RandomMatrix1* is defined as *RandomSeed1* (user defined and constant throughout the model) + *PSA Iteration* + 181

RandomMatrixRows1 = RandomSeed1 + sample_run + RandMatNoVbles

2. The seed for the random number generator is the *RandomSeed1 + sample_run*. *Sample_run* will change every PSA iteration. As this procedure is only run once for each PSA iteration a random set of numbers will be generated each time it is run; i.e. once per PSA iteration.
3. The procedure fills the matrix *RandomMatrix1(i, j)* with random numbers
4. The procedure copies the random numbers from the matrix *RandomMatrix1* to the matrices used in the model, for example the matrix *RandMatComplicat*. This matrix is used to determine the occurrence or not of events.

```
For i = 1 To cycle_no_smoking
  For j = 1 To 11
    RandMatComplicat(i, j) = RandomMatrix1(RandomSeed1 + sample_run + i - 1, j)
  Next j
Next i
```

The key here is that the matrix *RandMatComplicat(i, j)* seems to be filled once and only once at the start of each PSA iteration. Within a given PSA iteration it appears that the values in *RandMatComplicat(i, j)* are constants.

As an example of how this is then used to determine whether an event happens or not for a given patient during a given cycle, the formula for IHD can be taken from the *Higher_steps_diabetes* procedure. The main visual basic code for this appears to be:

*If RandMatComplicat(1 + index_cycle, 1) < MatLTC_IHDprob_event(index_cycle + 1) * MatAE_CVnon_fatal_Spec(1) Then MatLTC_IHD_value(index_cycle + 1) = 1*

The key to this is within the condition the random value that is compared to the calculated event probability is *RandMatComplicat(1 + index_cycle, 1)*. The value indexed is not determined by the patient number, and so is common to all patients across the PSA iteration. In other words, for a given cycle and event, the same random number is used across all patients within a PSA iteration to determine whether an event occurs or not.

At its most extreme, this means that simulating patients who are identical at baseline results in them all following exactly the same path through the model and having the same total costs and QALYs. And this appears to be what happens within the model. 100 identical patients can be simulated by setting the upper bound and lower bound of the continuous patient characteristics equal to the mean

value and any categorical variables equal to 0%[1]. When the model is run, within a PSA iteration the 100 patients have identical paths through the model, and each patient is simulated as having the same total costs and QALYs. The situation exactly parallels the 1,000 prostate cancer patients outlined above; either all are simulated as dying during the first cycle or none are simulated as dying.

This can be further cross checked through additional visual basic code being inserted into the ECEM to output the random numbers being used to assess whether a patient experiences a certain event during a certain cycle. Doing this[2] and running the ECEM with both 1st order and 2nd order sampling confirms that for a given PSA iteration the same set of random numbers is used for each patient.

If the ERG is correct, this is a fundamental error within the implementation of the patient level model. The likely impact of this upon model outputs cannot be formally quantified, though the section on convergence and varying the random number seeds below gives an indication of the possible impact.

4.5.2 Modelling of the evolution of HbA1c

The model can be initially set up to undertake a 3rd line comparison of empagliflozin 25mg with canagliflozin 100mg.

The model can be run setting the risk equations covariance matrix to be zero. The 3rd line baseline patient characteristics and patient disease history can be set to have s.e.s. of zero and their upper and lower bounds to be equal to the mean values. The years since diagnosis can be set to five and the proportion in their second year since diagnosis set to zero. The thresholds for HbA1c can be set to 10% so that patients remain on their 3rd line therapy throughout.

The baseline HbA1c can be further set to 7.5%, the absolute treatment efficacy for HbA1c for empagliflozin set to 1.0 and the relative treatment efficacy for HbA1c for canagliflozin set to 0.5. And again, the upper and lower bound can be set equal to the mean value.

Running 10 patients through the model with 1 PSA iteration results in the same modelled evolution of HbA1c as running 1 patient through the model with 1 PSA iteration.

The modelled evolutions of the HbA1c can be crossed checked by computing the HbA1c evolutions that are implied by the UKPDS 68 equation 11 in an ERG stand-alone excel worksheet.

[1] Note that this also sets the baseline prevalences of the complications of diabetes within the UKPDS 68 to 0%, so these patients should be modelled as being able to experience the entire range of the complications of diabetes during the modelling.
[2] The ERG is happy to supply this to the manufacturer upon request.

These can be further cross checked by comparing the values simulated within the electronic UKPDS Oxford Outcomes Model 1 (OM1). Within this the duration of diabetes has to be adjusted to take into account the stable period implemented within the ECEM model. As the UKPDS outcomes model only deals with full years, i.e. even cycle numbers, this is most easily implemented by assuming a 6 year duration of diabetes at baseline.

Comparing the simulated values in the ERG stand-alone excel worksheet with those of the electronic UKPDS OM1 shows very good agreement, as below.

Table 19. ERG and UKPDS OM1 modelled evolution of HbA1c

	ERG		UKPDS OM1	
Year	Emp.	Can.	Emp.	Can.
1	6.500	6.000	0.000	0.000
2	6.933	6.554	6.934	6.555
3	7.281	6.993	7.283	6.995
4	7.562	7.344	7.565	7.346
5	7.791	7.625	7.794	7.628
6	7.978	7.852	7.981	7.855
7	8.133	8.037	8.136	8.041
8	8.262	8.189	8.265	8.193
9	8.370	8.315	8.374	8.319
10	8.462	8.421	8.466	8.424
11	8.542	8.510	8.545	8.514
12	8.611	8.587	8.614	8.590

The values reported in the *Graphs_data* worksheet can be compared to those suggested by the ERG stand-alone excel worksheet. The latter can be calculated on the basis of taking the initially lagged value of HbA1c from the second cycle, as in the ERG even cycles below. But since the stable period of HbA1c reported it in the *Graphs_data* worksheet is until the third cycle, the more relevant comparison is to take it from the third cycle as in the ERG odd cycles below.

Table 20. ERG and ECEM modelled evolution of HbA1c

		ECEM		ERG even cycles				ERG odd cycles			
Year	Cycle	Emp.	Can.	Emp.	Diff.	Can.	Diff.	Emp.	Diff.	Can.	Diff.
0.5	1	6.500	6.000								
1.0	2	6.500	6.000	6.500	0.000	6.000	0.000				
1.5	3	6.500	6.000					6.500	0.000	6.000	0.000
2.0	4	6.712	6.272	6.933	-0.222	6.554	-0.282				
2.5	5	6.923	6.544					6.943	-0.020	6.564	-0.020
3.0	6	7.145	6.826	7.281	-0.136	6.993	-0.167				
3.5	7	7.367	7.108					7.298	0.069	7.010	0.099
4.0	8	7.547	7.334	7.562	-0.016	7.344	-0.010				
4.5	9	7.727	7.559					7.583	0.144	7.364	0.196
5.0	10	7.861	7.725	7.791	0.070	7.625	0.100				
5.5	11	7.996	7.891					7.813	0.182	7.647	0.243
6.0	12	8.094	8.009	7.978	0.116	7.852	0.157				
6.5	13	8.192	8.127					8.002	0.190	7.876	0.252
7.0	14	8.264	8.212	8.133	0.132	8.037	0.175				
7.5	15	8.337	8.298					8.156	0.180	8.061	0.237
8.0	16	8.391	8.360	8.262	0.130	8.189	0.171				
8.5	17	8.446	8.422					8.285	0.161	8.212	0.210
9.0	18	8.489	8.470	8.370	0.119	8.315	0.155				
9.5	19	8.532	8.517					8.393	0.139	8.338	0.180
10.0	20	8.566	8.555	8.462	0.104	8.421	0.134				
10.5	21	8.601	8.593					8.484	0.117	8.443	0.150
11.0	22	8.630	8.624	8.542	0.089	8.510	0.113				
11.5	23	8.660	8.655					8.563	0.097	8.531	0.123
12.0	24	8.685	8.681	8.611	0.074	8.587	0.094				

For the ERG modelling based upon odd cycles there is a very slight initial discrepancy, but for the bulk of the modelling the values simulated by the ECEM lie above those of the ERG, thereby predicting a slightly steeper rise in HbA1c over time. This apparent discrepancy will apply to all therapies modelled and as a consequence the overall impact upon net quantities may be limited. But modelling the evolution of HbA1c correctly does seem quite a fundamental aspect of any diabetes model.

The ERG has rebuilt the logic of the ECEM visual basic modelling within Excel, and the results of this cross check with the values outputted by the ECEM[3]. But it has also thrown up what appears may be an error within the visual basic code.

The UKPDS 68 equation 11 specifies the following functional form for the evolution of HbA1c in year t[4]:

$H_t - 7.09 = \beta_0 + \beta_1 ln(Dur_t) + \beta_2(Year2_t) + \beta_3(H_{t-1} - 7.09) + \beta_4(H_{Base} - 7.09)$

Where the Dur_t is the duration in years at time t, *Year2* is whether the patient is in the second year since diagnosis and H_{Base} is the HbA1c value at diagnosis. Ignoring the Year2 element since this will generally not apply this simplifies to:

$H_t - 7.09 = \beta_0 + \beta_1 ln(Dur_t) + \beta_3(H_{t-1} - 7.09) + \beta_4(H_{Base} - 7.09)$

The visual basic of the model runs in 6 monthly cycles. As a consequence it calculates an annual step, halves it and applies it over two cycles. For the sake of simplicity this halving and double application can be largely ignored in what follows.

In order to calculate the value for *Mat_Current_HBA1c* for the fifth cycle the visual basic calculates the step based upon the parameter values that apply during the third cycle. This appears to be incorrect. The UKPDS 68 requires that the parameter values of the current cycle be used.

The visual basic calculates the value for *Mat_Current_HBA1c* for the fifth cycle along the following lines. Note that the subscripts here refer to 6 month cycles, hence cycle 5 is year 2.5, cycle 3 is year 1.5 and cycle 1 is year 0.5, all of which are one year apart and so aligned with the UKPDS 68.

(1) Calculate the value at cycle 3 of the UKPDS equation 11

$Equation11_3 = \beta_0 + \beta_1 ln(Dur_3) + \beta_3(H_1 - 7.09) + \beta_4(H_{Base} - 7.09)$

[3] This also implements a replication of the visual basic as would be implied were the model cycle one year rather than six months with this cross checking with the six monthly implementation and compares this with what the ERG views to be the correct implementation of the UKPDS 68 calculated on an annual basis. This is available to the manufacturer upon request.

[4] Note that within table 1 of the UKPDS 68 the definition of H_t is given as the two year moving average of HbA1c. The ERG rebuild and the UKPDS OM1 treat H_t as the value of HbA1c at time t (*personal communication: Alastair Gray*). It seems unlikely that this is the source of the discrepancy between the ECEM and the ERG. It appears that the intended implementation within the ECEM is in line with the ERG rebuild and the UKPDS OM1, but this has not been confirmed with the manufacturer. There is an additional interpretation where the lag of HbA1c could also be interpreted as the lagged 2 year moving average, but again this interpretation does not appear to yield anything in line with the ECEM implementation. If the visual basic matrix of HbA1c values is filled with two year moving averages and this is the interpretation being placed upon it by the manufacturer, the equation structure for the 5th cycle value should be the same as that suggested by the ERG. The inclusion of Dur_3 in the equation for HbA1c in the fifth cycle suggests that there is may be a problem with the lagging of variables.

(2) Calculate the step

$$Step_3 = Equation11_3 - (H_1 - 7.09)$$
$$= \beta_0 + \beta_1 ln(Dur_3) + (\beta_3 - 1)(H_1 - 7.09) + \beta_4(H_{Base} - 7.09)$$

(3) Add the step to the value of HbA1c of the third cycle

$$H_5 = H_3 + Step_3$$
$$= H_3 + \beta_0 + \beta_1 ln(Dur_3) + (\beta_3 - 1)(H_1 - 7.09) + \beta_4(H_{Base} - 7.09)$$

It appears that the correct implementation should be along the following lines

$$H_5 = 7.09 + \beta_0 + \beta_1 ln(Dur_5) + \beta_3(H_3 - 7.09) + \beta_4(H_{Base} - 7.09)$$

The two methods are not equivalent. This appears to be the probable reason for the ECEM modelled evolution of HbA1c being different from that of the ERG and the UKPDS OM1.

4.5.3 Modelling of the evolution of SBP

The same exercise can be performed for the modelling of SBP for a baseline SBP of 143 mmHg with treatment effects of a 3 mmHg reduction and a 5mmHg reduction. This results in the following.

Table 21. ECEM vs. ERG modelling of SBP

		ECEM 3.0 drop	ECEM 5.0 drop	ERG 3.0 drop	ERG 5.0 drop	Diff. 3.0 drop	Diff. 5.0 drop
Year	Cycle						
0.0	0	140.0	140.0	140.0	140.0	0.0	0.0
0.5	1	137.0	135.0				
1.0	2	137.0	135.0	137.0	135.0	0.0	0.0
1.5	3	134.5	132.5				
2.0	4	135.1	133.4	138.2	136.7	-3.0	-3.3
2.5	5	135.7	134.2				
3.0	6	136.6	135.5	139.0	138.0	-2.4	-2.5
3.5	7	137.5	136.7				
4.0	8	138.3	137.7	139.7	139.0	-1.4	-1.3
4.5	9	139.1	138.7				
5.0	10	139.7	139.4	140.2	139.7	-0.5	-0.3
5.5	11	140.3	140.1				
6.0	12	140.6	140.5	140.7	140.3	0.0	0.2
6.5	13	141.0	140.9				
7.0	14	141.2	141.1	141.0	140.7	0.2	0.4
7.5	15	141.4	141.4				
8.0	16	141.5	141.5	141.3	141.1	0.3	0.5
8.5	17	141.7	141.7				
9.0	18	141.7	141.7	141.5	141.3	0.3	0.4
9.5	19	141.8	141.8				
10.0	20	141.9	141.9	141.7	141.6	0.2	0.3
10.5	21	141.9	142.0				
11.0	22	142.0	142.0	141.8	141.8	0.2	0.3
11.5	23	142.0	142.1				
12.0	24	142.1	142.1	142.0	141.9	0.1	0.2
12.5	25	142.1	142.1				
13.0	26	142.2	142.2	142.1	142.0	0.1	0.1
13.5	27	142.2	142.2				
14.0	28	142.3	142.3	142.2	142.2	0.1	0.1

For SBP there are reasonable differences in the early years, though this reduces with time. The reasons for this appear to mainly lie in the ECEM simulating an initial treatment effect and a further treatment effect at the third cycle.

Both the UKPDS OM1 and the ERG are in broad agreement. If the UKPDS OM1 is run with 5 years' duration of diabetes, due to the apparently shorter stable period simulated within the ECEM, the following applies.

Table 22. UKPDS OM1 vs. ERG modelling of SBP

	ERG	ERG	UKPDS OM1	UKPDS 0M1	Diff.	Diff.
Year	1.0% drop	1.5% drop	1.0% drop	1.5% drop	1.5% drop	1.5% drop
0	140.0	140.0	140.0	140.0	0.0	0.0
1	137.0	135.0	137.0	135.0	0.0	0.0
1	138.2	136.7	138.1	136.7	0.1	0.1
2	139.0	138.0	138.9	137.9	0.1	0.1
3	139.7	139.0	139.6	138.9	0.1	0.1
4	140.2	139.7	140.1	139.6	0.1	0.1
5	140.7	140.3	140.5	140.1	0.1	0.1
6	141.0	140.7	140.9	140.6	0.1	0.1
7	141.3	141.1	141.1	140.9	0.1	0.1
8	141.5	141.3	141.4	141.2	0.1	0.1
9	141.7	141.6	141.6	141.5	0.1	0.1
10	141.8	141.8	141.7	141.6	0.1	0.1
11	142.0	141.9	141.9	141.8	0.1	0.1
12	142.1	142.0	142.0	142.0	0.1	0.1
13	142.2	142.2	142.1	142.1	0.1	0.1

There appear to be two errors within the ECEM modelling of SBP. The first related to the step at the third cycle, while the second again relates to quite what should be lagged within the UKPDS 68 calculation. The more serious error is in the calculation of the SBP value for the third cycle, but ease of exposition suggests outlining the other possible error first. This is along similar line to the apparent error within the calculation of the HbA1c values in terms of the lagging of variables.

The UKPDS 68 provides the following equation for the derivation of the value for SBP at time t.

$$(SBP_t - 135.09)/10 = \beta_0 + \beta_1 ln(Dur_t) + \beta_2 (SBP_{t-1} - 135.09)/10 + \beta_3 (SBP_{Base} - 135.09)/10$$

But the ECEM calculates the SBP of, say, the seventh cycle along the following lines.

$$SBP_7 = SBP_5 + 10 * \left[\beta_0 + \beta_1 ln(Dur_7) + \frac{\beta_2 (SBP_3 - 135.09)}{10} + \frac{\beta_3 (SBP_{Base} - 135.09)}{10} - \frac{(SBP_3 - 135.09)}{10}\right]$$

$$= SBP_5 + 10 * \left[\beta_0 + \beta_1 ln(Dur_7) + \frac{(\beta_2 - 1)(SBP_3 - 135.09)}{10} + \frac{\beta_3 (SBP_{Base} - 135.09)}{10}\right]$$

This is not entirely in line with the UKPDS 68 which would suggest

$$SBP_7 = 135.09 + 10 * \left[\beta_0 + \beta_1 ln(Dur_7) + \frac{\beta_2(SBP_5 - 135.09)}{10} + \frac{\beta_3(SBP_{Base} - 135.09)}{10}\right]$$

It appears that the ECEM intention is to calculate SBP_3 along similar line to that outlined above for SBP_7. But it further appears that the sum of the UKPDS 68 equation 12 at this point has not been evaluated. In other words, within the VB

*MatSBP_new(index_cycle + 1) = Mat_Current_SBP(index_cycle + 1) + 10 * (MatSBP_Weighted_Sum(index_cycle + 1) - Mat_L1SBP(index_cycle + 1)) / 2*

it appears that for cycles 1 and 2 *MatSBP_Weighted_Sum* is not evaluated and so by default is zero. For this to be explored, it is necessary to return to 6 monthly cycles and so a halving of the annual step from SBP_2 *to* SBP_3. Given this, the equation for SBP_3 reduces to the following.

$$SBP_3 = SBP_2 + 10 * \left[-\frac{(SBP_{Base} - 135.09)}{10}\right] /2$$

$$= SBP_2 - \frac{(SBP_{Base} - 135.09)}{2}$$

$$= (SBP_{Base} - SBP_{Effect}) - \frac{(SBP_{Base} - 135.09)}{2}$$

$$= \frac{(SBP_{Base} + 135.09)}{2} - SBP_{Effect}$$

This calculation would be incorrect and it appears that it is this that gives rise to the unwarranted step at the third cycle. If it applies it will cause the modelled evolutions of SBP to be incorrect for both arms. It is not obvious whether bias in any particular direction will occur. As for the modelled evolution of HbA1c the impact may tend to net out between the arms. But again, correctly modelling the evolution of SBP seems fairly central to any modelling of diabetes.

4.5.4 Modelling of the evolution of Lipids

The UKPDS 68 gives the following for modelling the evolution of the lipids ratio of total cholesterol to high density lipoprotein (HDL) cholesterol:

$$LPD_t - 5.23 = \beta_0 + \beta_1(LPD_{t-1} - 5.23) + \beta_2(LPD_{Base} - 5.23)$$

As with the modelling of SBP, there may be some problems in terms of quite what is lagged when. The VB of the model appears to calculate LPD_7 along the following lines.

$$LPD_7 = LPD_5 + \left(\beta_0 + \beta_1(LPD_3 - 5.23) + \beta_2(LPD_{Base} - 5.23) - (LPD_3 - 5.23)\right)$$

$$= LPD_5 + \left(\beta_0 + (\beta_1 - 1)(LPD_3 - 5.23) + \beta_2(LPD_{Base} - 5.23)\right)$$

It also seems to be subject to a similar error as the modelling of SBP for the third cycle, in that within the VB

MatLipid_new(index_cycle + 1) = Mat_Current_lipid(index_cycle + 1) + (MatLipid_Weighted_sum(index_cycle + 1) - Mat_L1Lipid(index_cycle + 1)) / 2

the value for *MatLipid_Weighted_sum* is again not evaluated for cycle 1 and cycle 2 and so seems to be by default zero. This again appears to feed into the calculation of the lipid ratio for cycle 3, with this value then flowing on into subsequent calculations.

Given this, the equation for LPD_3 reduces to the following.

$$LPD_3 = LPD_2 + \frac{(-(LPD_2 - 5.23))}{2}$$
$$= \frac{(LPD_2 + 5.23)}{2}$$
$$= \frac{(LPD_{Base} + 5.23)}{2}$$

Reverting to the annual cycle of the UKPDS 68 the ECEM modelled values compared to those implied by the UKPDS are given below.

Table 23. ECEM vs ERG modelling of Lipids' ratio

Year	Cycle	ECEM	ERG	Diff.
0.5	1	4.335	4.335	0.000
1.5	3	4.783	4.512	0.270
2.5	5	4.960	4.605	0.354
3.5	7	4.924	4.654	0.270
4.5	9	4.805	4.680	0.126
5.5	11	4.703	4.694	0.010
6.5	13	4.657	4.701	-0.043
7.5	15	4.660	4.704	-0.045

As with the modelling of HbA1c and SBP, the impact of the identified discrepancies may tend to net out between the arms. But again, correctly modelling the evolution of the lipids ratio seems fairly central to any modelling of diabetes.

4.5.5 Model convergence

The results presented within the submission are based upon 100 patients being sampled for each PSA iteration, and 300 PSA iterations being run. These are unusually low numbers of both patients and PSA iterations. The manufacturer claims that this is due to memory limits in excel limiting the size of matrices. The ERG asked why these limits could not be avoided by constructing the model to run 10,000 patients for one PSA iteration, store the aggregate results for that PSA iteration, discard the

individual results for the 10,000 patients and then run another PSA iteration. The manufacturer responded:

The model was built in a way that we could have access to the information of each of the patients simulated under each of the different samples. By running samples of patients and discarding them afterwards the specific information per patient would have been lost, and it was considered relevant to be able to access and export to Excel if necessary that level of detail for transparency purposes. In any case, it is unclear whether by discarding the individual 10,000 patient simulations per PSA and storing the PSA values only for the 5,000 iterations would have solved the memory issue of Excel (i.e. resulting in an 'out of memory' error in Excel).

The ERG is unclear why the manufacturer wanted to have access to the information of each of the patients simulated under each of the different samples. The submission does not make use of this information. The ERG is also of the opinion that discarding the individual 10,000 patient simulations per PSA iterations and only storing the PSA values for each of the 5,000 iterations probably would solve any memory issue in Excel.

The DSU technical support document number 15 notes that:

Typically the number of patients to sample is left to the discretion of the modeller. However, it would be expected that all modellers justify the number of patients selected. Methods of justification can include a graphical representation of the costs, QALYs and the cost per QALY gained and determining at what number of patients the estimated error in the results appear acceptable.

This is followed in the DSU technical support document by plots of the incremental QALYs, incremental costs and incremental cost effectiveness ratio (ICER) as a function of the numbers of patients simulated.

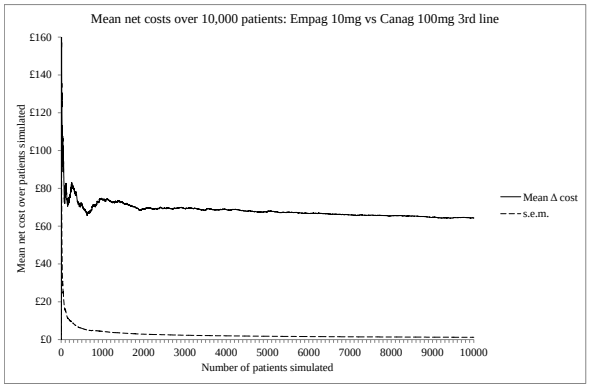
In response to ERG clarification questions, the manufacturer provided an analysis of the convergence of QALY estimates for a deterministic run of the model for a chosen comparison of third line therapies of metformin plus sulfonylurea in combination with either empagliflozin 10mg or sitagliptin 100mg. This presented the mean values, and the upper and lower bounds of the estimates as the number of patients simulated was increased. The estimate of the mean incremental QALY gain shows a reasonable amount of variation, actually flipping from positive to negative as the number of patients that are simulated increased from 100 to 1,000. The incremental QALY estimate appears to only begin to settle once the number of patients simulated rises above about 8,000. While the incremental QALY remains small, for the simulation of 10,000 patients it has more than doubled compared to that of 100 patients. That said, by arm the highest mean is only around 2% above the lowest mean in part illustrating that the model is trying to split very small QALY differences between the arms even for the comparison with sitagliptin.

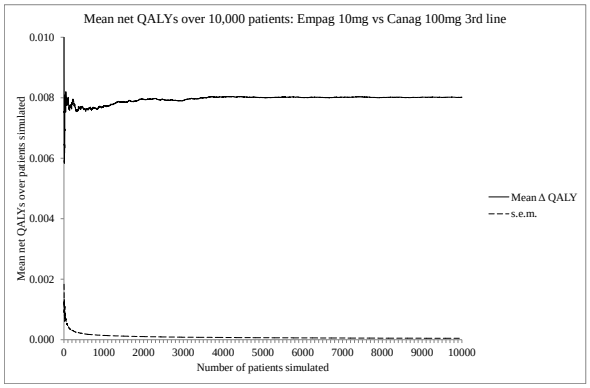
Table 24. Manufacturer simulated QALYs by number of patients simulated

	Empagliflozin 10mg			Sitagliptin 100mg			Incremental		
Patients	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper
100	7.271	0.566	12.916	7.260	0.563	12.880	0.011	-0.007	0.006
1,000	7.147	0.577	12.864	7.150	0.574	12.867	-0.003	-0.007	0.006
2,000	7.262	0.577	13.235	7.242	0.574	13.225	0.020	-0.007	0.006
3,000	7.260	0.851	13.220	7.241	0.847	13.146	0.019	-0.007	0.006
4,000	7.286	0.909	13.147	7.261	0.895	13.141	0.025	-0.007	0.006
5,000	7.292	0.920	13.174	7.271	0.917	13.145	0.022	-0.007	0.006
6,000	7.292	0.899	13.164	7.269	0.895	13.142	0.024	-0.007	0.006
7,000	7.290	0.909	13.166	7.267	0.895	13.145	0.023	-0.007	0.006
8,000	7.284	0.906	13.178	7.255	0.895	13.146	0.029	-0.007	0.006
9,000	7.272	0.882	13.164	7.242	0.850	13.146	0.029	-0.007	0.006
10,000	7.271	0.848	13.159	7.244	0.814	13.146	0.028	-0.007	0.006

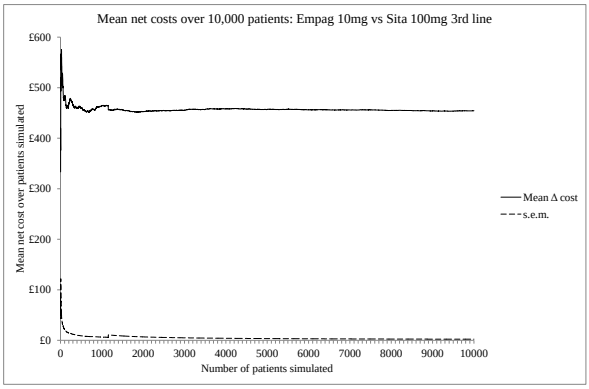
The model can be made to simulate 10,000 patients for a comparison of third line empagliflozin 10 mg with canagliflozin 100 mg. 2nd order sampling can be turned off but 1st order sampling retained with cell AJ48 of the Model_Scope worksheet set to "Yes". The model can then be run for 1 PSA. The seeds used for this run were 32 and 31. A similar exercise can be conducted for a comparison of third line empagliflozin 10 mg with sitagliptin 100 mg.

Empagliflozin 10 mg vs. canagliflozin 100 mg





Empagliflozin 10 mg vs. sitagliptin 100 mg



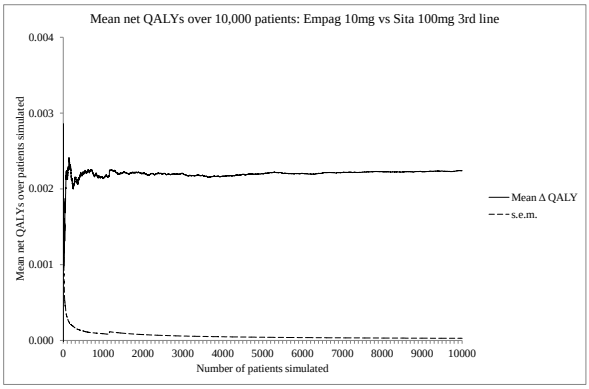


Figure 18. Convergence of values with no 2nd order sampling: deterministic model results

The above figures are based upon there being 1st order sampling of patient variability but no 2nd order sampling; i.e. they are evaluated at central parameter estimates for costs, utilities and treatment effects. The mean values show a fair degree of variability up to and perhaps beyond around 2,000 patients being run through the model. What is clear is that the model has not converged when running only 100 patients through the model. The manufacturer accepts this, noting it its response to clarification question B19 that:

"100 patients per sample did not appear to reach convergence"

Running 100 patients over 300 PSA iterations results in 30,000 patients being run through the model for each arm. The ERG cross check of the calculation of the central estimate for the probabilistic modelling suggests it is based upon the mean cost and mean QALY being calculated for each PSA iteration, and these then subsequently being averaged to give the overall mean costs and mean QALYs from which the central estimate of cost effectiveness is calculated. This is the correct method. Given this, there is no reason to think that the lack of convergence for each PSA iteration would result in a biased central estimate. But whether averaging over 300 PSA iterations is sufficient for the central estimate of the cost effectiveness to have converged is a moot point and one that it is difficult

to formally assess given the current model structure other than by altering the random number seeds of the model.

Any lack of convergence of each PSA iteration if only 100 patients are run through the model does mean that the characterisation of the impact of 2nd order uncertainty upon the uncertainty around the central estimate of cost effectiveness is unreliable; i.e. the scatterplot on the cost effectiveness plane and the CEAC will be unreliable.

But in this check of convergence of the model as a function of the number of patients modelled, it is crucially important that the random numbers that event probabilities are compared against to determine whether an event occurs or not differ between patients. It appears that they do not. As a consequence, the above will greatly overstate the degree of convergence of results for a given number of patients. For instance, if identical patients were being simulated, due to the issue around random number sampling resulting in all these patients having the same course through the model it would have converged with just one patient being simulated. This analysis of convergence should be performed again once populating of the matrix of random numbers that event probabilities are compared against to determine whether an event occurs or not has been corrected.

It may also suggest that the manufacturer should have presented the results of deterministic model with runs of perhaps 10,000 patients with no sampling of 2nd order uncertainty. In response to an ERG clarification question the manufacturer supplied this, but unfortunately did not specify what number of patients were run for the deterministic model and what numbers of patients and iterations were run for the probabilistic model.

As already noted, given the current model structure and random sampling the stability of the central estimate of the cost effectiveness can only really be explored by varying the two random seeds of the model. One random seed[5] drives the sampling of 1st order uncertainty. The other random seed[6] drives the sampling of 2nd order uncertainty and the populating of the matrix of random numbers that event probabilities are compared against to determine whether an event occurs or not[7]. The model can be set to the manufacturer default values with 100 patients and 300 PSA iterations, and a range of random seeds applied. This results in the following.

[5] *RandomSeed#2*
[6] *RandomSeed#1*
[7] *RandMatComplicat*

Table 25. Random seeds' impacts upon model outputs

Random seed		Empagliflozin		Sitagliptin		Net		
1st order	2nd order	QALYs	Cost	QALYs	Cost	QALYs	Cost	ICER
10	31	7.010	£31,235	6.976	£31,192	0.034	£43	£1,249
23	45	7.005	£30,371	6.973	£30,380	0.032	-£9	E. Dom.
31	67	6.965	£30,884	6.939	£30,868	0.026	£17	£633
78	23	6.995	£30,513	6.961	£30,462	0.034	£51	£1,521
22	16	7.021	£30,971	6.990	£30,844	0.031	£128	£4,140
30	22	6.988	£30,559	6.957	£30,524	0.031	£36	£1,142
53	38	7.015	£30,578	6.987	£30,621	0.028	-£41	E. Dom.
27	42	6.979	£30,649	6.949	£30,701	0.030	-£53	E. Dom.
5	33	7.044	£31,224	7.022	£31,305	0.022	-£83	E. Dom.
44	11	6.976	£30,694	6.950	£30,698	0.026	-£5	E. Dom.

The above illustrates the variability in results of the model that result from just changing the random number seeds. Cost estimates in the empagliflozin arm range from £30,373 to £31,235 and in the sitagliptin arm from £30,380 to £31,305: ranges of around £1,000. As would be expected these tend to move together between the two arms and the range of net costs is smaller: from a net saving of £83 to a net cost of £128.

The QALYs appears to show less variation to the random number seeds chosen, and range from 6.965 to 7.044 in the empagliflozin arm and from 6.949 to 7.022 in the sitagliptin arm. Net QALYs range from 0.022 to 0.034.

As a consequence, the cost effectiveness estimates of empagliflozin 10mg relative to sitagliptin 100mg ranges from empagliflozin dominating to £4,140 per QALY.

This illustrates the limits of the current ECEM to reliably differentiate treatments, and the reported central estimates from the PSA being sound. Variation exists between the simulations and this must be borne in mind when considering all the central estimates from the PSA model runs based upon 100 patients and 300 PSA iterations.

It should also be emphasised that the above is only for a small set of 10 random number seed pairs. The range of results that might be simulated given a larger set of random number seed pairs will be larger. A larger set of random number seed pairs might also help illustrate whether the distributions of results are quite spread out or tightly bunched with only a few outliers. But time constraints mean that this has not been examined by the ERG.

The impact of the random seeds can be further explored by arbitrarily selecting the random seed pair (22, 16) and individually individual varying the random seed used for the 1st order sampling, then individually varying the random seed used for the 2nd order sampling and population of the matrix of random numbers that event probabilities are compared against to determine whether an event occurs or not.

Table 26. Random seed for 1st order uncertainty impacts upon model outputs

Random seed		Empagliflozin		Sitagliptin		Net		
1st order	2nd order	QALYs	Cost	QALYs	Cost	QALYs	Cost	ICER
10	16	7.014	£30,895	6.984	£30,752	0.030	£144	£4,857
23	16	7.021	£30,981	6.990	£30,857	0.031	£124	£3,971
31	16	7.018	£31,034	6.986	£30,909	0.032	£125	£3,885
78	16	7.008	£30,970	6.974	£30,849	0.034	£121	£3,527
22	16	7.021	£30,971	6.990	£30,844	0.031	£128	£4,140
30	16	7.018	£31,009	6.987	£30,893	0.031	£116	£3,721
53	16	7.010	£30,954	6.973	£30,855	0.037	£99	£2,699
27	16	7.020	£31,007	6.988	£30,887	0.032	£121	£3,817
5	16	7.017	£30,936	6.985	£30,803	0.031	£132	£4,222
44	16	7.015	£31,007	6.981	£30,901	0.034	£105	£3,100

Table 27. Random seed for remaining sampling impacts upon model outputs

Random seed		Empagliflozin		Sitagliptin		Net		
1st order	2nd order	QALYs	Cost	QALYs	Cost	QALYs	Cost	ICER
22	31	7.021	£31,249	6.988	£31,206	0.034	£43	£1,288
22	45	7.004	£30,367	6.973	£30,379	0.031	-£11	E. Dom
22	67	6.961	£30,863	6.933	£30,870	0.028	-£7	E. Dom
22	23	7.009	£30,768	6.974	£30,749	0.035	£19	£547
22	16	7.021	£30,971	6.990	£30,844	0.031	£128	£4,140
22	22	6.987	£30,453	6.956	£30,393	0.032	£60	£1,884
22	38	7.017	£30,617	6.988	£30,723	0.029	-£107	E. Dom
22	42	6.978	£30,598	6.948	£30,670	0.030	-£72	E. Dom
22	33	7.040	£31,371	7.019	£31,436	0.020	-£65	E. Dom
22	11	6.964	£30,739	6.939	£30,725	0.025	£14	£541

The above appears to confirm that the main variability in results as a function of the random seed that is chosen stems from the random seed that is used for the 2nd order sampling and population of the matrix of random numbers that event probabilities are compared against to determine whether an

event occurs or not. In the opinion of the ERG this is likely to mainly arise due to the impacts upon the sampling for the population of the matrix of random numbers that event probabilities are compared against to determine whether an event occurs or not. But this has not been demonstrated.

With the random seed pair (31, 10) the model can also be run for 100 patients over differing numbers of PSA iterations. This results in the following.

Table 28. PSA iterations and model outputs

	Empagliflozin		Sitagliptin		Net	Net	
PSA iterations	QALYs	Cost	QALYs	Cost	QALYs	Cost	ICER
50	6.921	£30,525	6.887	£30,593	0.034	-£68	E. Dom.
100	6.807	£26,743	6.767	£26,830	0.040	-£87	E. Dom.
150	6.875	£27,731	6.839	£27,670	0.035	£61	£1,721
200	6.945	£29,493	6.915	£29,427	0.029	£66	£2,234
250	6.927	£30,751	6.897	£30,689	0.030	£62	£2,080
300	7.010	£31,235	6.976	£31,192	0.034	£43	£1,249

As the number of PSA iterations is increased it cannot be said that the central cost effectiveness estimate has converged once 300 iterations have been performed. Again, there appears to be more variation in costs than in QALYs.

This above throws into question to what degree the model can reliably distinguish between treatments in terms of costs and effects when running 100 patients over 300 iterations given the treatment effects inputted to the model.

4.5.6 Weight changes, quality of life and costs

For simplicity assume that there is no change of treatment. The matrix *MatWeight_gain_total(i)* appears to store the values according to:

MatWeight_gain_total(1) = MatWeight_gain(identifier_treatment)

MatWeight_gain_total(index_cycle + 1) = MatWeight_gain_total(index_cycle)

And as a consequence is filled with constant values as per *MatWeight_gain(identifier_treatment)*

The key to tracking the handling of weight within the modelling appears to be the matrix of values for *MatAE_weight_gain_cycle(i)*. This holds the values for the weight change for each cycle (i). For instance, for the period of initial weight changes associated with treatment it holds the values according to:

*MatAE_weight_gain_cycle(index_cycle + 1) = MatWeight_gain_total(index_cycle + 1) * cycle_length / weight_years_change*

These values then appear to be loaded into the *MatSum_weight_gain(i)* matrix according to *MatSum_weight_gain(i) = MatAE_weight_gain_cycle(i)*

The ERG cannot find an instance of
MatSum_weight_gain(i) = MatAE_weight_gain_cycle(i) + MatSum_weight_gain(i-1)
which would appear to be the more natural formulation.

This is conditioned by whether weight is above or below the baseline weight according to

MatActualWeight(index_cycle + 1) = MatActualWeight(index_cycle) + MatAE_weight_gain_cycle(index_cycle + 1)
If MatActualWeight(i) - weight_baseline > 0 Then
MatSum_weight_gain_event(i) = 1
Else
MatSum_weight_gain_event(i) = 0
End If
*MatSum_weight_gain_kg(i) = MatSum_weight_gain_event(i) * MatAE_weight_gain_cycle(i)*

Hence if the patient weight is above baseline, the value for *MatAE_weight_gain_cycle(i)* is loaded into *MatSum_weight_gain_kg(i)*. If not, *MatSum_weight_gain_kg(i)* is zero.

The quality of life values are then calculated according to

If MatSpec_Scenario(1) = 1 and MatWeight_gain_total(i) >0 Then
*MatSumQALYs_Weight_gain(i) = MatSum_weight_gain_kg(i) * MatUtilities_Value(24)*
Else
MatSumQALYs_Weight_gain(i) = 0
End If

The condition for quality of life values to be applied is that the choice of scenario selects that it is and that the initial treatment impact upon weight is positive; i.e. that there is an initial weight gain. If so, the utility value is then applied to the *MatSum_weight_gain_kg(i)*, which as already noted is equal to *MatAE_weight_gain_cycle(i)* if the patient weight is above baseline and is zero if it is not. In other words there is a joint condition being applied:

- The initial treatment effect must be a weight gain
- The patient weight must be above baseline

As far as the ERG can discern, the intention within the coding is that when these both apply the impact of the full weight gain including any increases due to natural history has a QALY impact. But note that if the initial treatment effect is a weight loss but the patient weight has risen above baseline due to natural history, the weight gain above baseline will not have a QALY impact. This asymmetric handling of weight gains is not obviously reasonable and may bias the analyses towards treatments that result in an initial weight loss thus favouring the flozins over weight-neutral drugs such as the gliptins, and more so over drugs such as pioglitazone and insulin which cause weight gain

But applying the quality of life value to the *MatSum_weight_gain_kg(i)* will under a number of circumstances be equivalent to applying it to *MatAE_weight_gain_cycle(i).* This seems incorrect.

Following this through to the final computations appears to confirm that the above is the working of the model and not redundant code in that the following appears to apply

*MatSumQALYs_disc_Weight_gain(i) = MatSumQALYs_Weight_gain(i) * MatSumQALYs_Disc_factor(i)*

MatSumQALYs_cum_disc_Weight_gain(i) = MatSumQALYs_disc_Weight_gain(i) + MatSumQALYs_cum_disc_Weight_gain(i - 1)

MatRESULTS_Disc_Cumul_Weight_gain_QALYs = MatSumQALYs_cum_disc_Weight_gain(largest_cycle)

Note also that costs can be associated with weight changes, but that in the base case this cost is assigned to be £0. It appears that these are not handled asymmetrically between treatments associated with an initial weight loss and an initial weight gain once weight rises above the baseline weight.

*MatSumDirC_Weight_gain(i) = MatSum_weight_gain_kg(i) * MatAECosts_Dir(1)*

4.5.7 Halving of the QALY decrements associated with adverse events and complications

It appears that within the *Computations_QALY_LY* the visual basic multiplies the quality of life values applied to UKPDS 68 events by the cycle length of 0.5 to arrive at the appropriate QALY decrement for a given cycle.

*MatSumQALYs_IHD(i) = part5_prod(i, 1) * cycle_length*

*MatSumQALYs_MI(i) = part5_prod(i, 2) * cycle_length*

*MatSumQALYs_CHF(i) = part5_prod(i, 3) * cycle_length*

*MatSumQALYs_stroke(i) = part5_prod(i, 4) * cycle_length*

*MatSumQALYs_Amp(i) = part5_prod(i, 5) * cycle_length*

*MatSumQALYs_Blind(i) = part5_prod(i, 6) * cycle_length*

*MatSumQALYs_Renal(i) = part5_prod(i, 7) * cycle_length*

The values calculated for the adverse events do not require conditioning by the cycle length as the rates are specific to a six month period and the quality of life values inputted to the model are QALY decrements rather than QoL decrements.

But the visual basic code subsequently reapplies the cycle length of 0.5 to both the baseline QoL and the adverse event and complications' QALYs along the following lines.

MatSumQALYs_QALYs(i) =
*cycle_length * MatSumQALYs_Alive(i) **
(*((MatUtilities_Mean(1) + (MatPatDbn(1) + i * cycle_length - RefAge_Utilities) * Age_related_disU_Value)) + MatSumQALYs_Hypo(i) + MatSumQALYs_PeriOed(i) + MatSumQALYs_GI(i) + MatSumQALYs_Fractures(i) + MatSumQALYs_NeedMonitoring(i) + MatSumQALYs_UTI(i) + MatSumQALYs_GTI(i) + MatSumQALYs_AE1(i) + MatSumQALYs_AE2(i) + MatSumQALYs_TIA(i) + MatSumQALYs_Strokes_non_fatal(i) + MatSumQALYs_MI_non_fatal(i) + MatSumQALYs_OtherIsch(i) + MatSumQALYs_Weight_gain(i) + MatSumQALYs_IHD(i) + MatSumQALYs_MI(i) + MatSumQALYs_CHF(i) + MatSumQALYs_stroke(i) + MatSumQALYs_Amp(i) + MatSumQALYs_Blind(i) + MatSumQALYs_Renal(i)* ***)***
+ MatSumQALYs_DiabDeaths(i) + MatSumQALYs_OtherDeaths(i)

Within the above it appears that all the elements within the highlighted brackets are being conditioned by the *cycle_length*: the baseline QoL, the adverse events' QALYs and the complications' QALYs.

This possible error was only identified due to the ERG having difficulty cross checking the detail of the summation of the QALYs. It was also only identified close to deadline for submission of the ERG report, and as a consequence has not been rebuilt in detail by the ERG. But if this is an error, it would largely invalidate the results of the ECEM.

4.6 Minor issues with the ECEM implementation

4.6.1 UKPDS 68 adjustments to age and lipid levels

The ERG questioned the adjustment within the VB of *MatEventDeath_ln_Age_event(1) = Log(MatEventDeath_Age_Event(1) + 52.59) - Log(52.59)* as this does not appear to be in line with the adjustments suggested in table 1 of the UKPDS 68[8]. The manufacturer replied that the adjustment of *Age_Event* – 52.59 as suggested in the UKPDS 68 could result in having to take the log of a negative number. Both the manufacturer and the ERG have received a spreadsheet supplied by

[8] Note that Visual Basic uses Log of the natural logarithm, while Excel uses Ln. Excel also has a Log function, but where the base is not specified it is base 10 rather than base e.

Professor Clarke, one of the UKPDS 68 authors. This spreadsheet does not address *Age_Event* but it does address the lipid ratio of *Total:HDL*. Table 1 of the UKPDS suggests that this should be transformed according to *Total:HDL* – 5.23. Taking the log of this could result in having to take the log of a negative number, but the spreadsheet of Professor Clarke makes the adjustment Ln(*Total:HDL*) – Ln(5.23) = Ln(*Total:HDL* / 5.23).

Within the UKPDS 68 Gompertz of equation 9 applying an *Age_Event* of 60 years and making the adjustment as Ln(*Age_Event* – 52.59) results in a probability of death of 100%. Making the adjustment as Ln(*Age_Event* / 52.59) results in a more reasonable probability of death of around 5%.

The ERG did contact the UKPDS 68 authors, and while some answers were forthcoming none were about this particular aspect. But in the absence of any further information from the UKPDS 68 authors, the ERG is of the opinion that the adjustment made by the manufacturer is the most reasonable to apply.

It should also be noted that in the spreadsheet supplied by Professor Clarke that while Ln(Total:HDL) is adjusted according to Ln(Total:HDL / 5.23) where coefficients are being applied to the unlogged quantity the adjustment is as suggested in Table 1 of the UKPDS 68: Total:HDL - 5.23. In other words the subtraction sign within Table 1 of the UKPDS 68 is not a typo for /. In the light of this, it also seems most reasonable to assume that where coefficients are being applied to the unlogged *Age_Event* the adjustment should be *Age_Event* – 52.59.

4.6.2 Derivation of six-monthly event probabilities

A detailed cross check of the interim values through time and a rebuild of the model code in Excel confirms that it works as intended and is broadly in line with the UKPDS 68.

For the most part the probabilities of the Weibulls of Table 2 of the UKPDS 68 are calculated on the basis of the integrated hazards for the start of the 6 month cycle and the end of the 6 month cycle being inputted into the visual basic

MatLTC_stroke_prob_event(index_cycle + 1) =

1- Exp(MatLTC_stroke_Weighted_sum_start_cycle(index_cycle + 1) - MatLTC_stroke_Weighted_sum_end_cycle(index_cycle + 1))

The above requires a mid-year integrated hazard to be calculated and applied as either the start of cycle value or the end of cycle value. As a consequence, the year start and year end values are complemented by the mid-year value which is typically calculated as:

*sum_weight_half_Stroke = sum_weight_start_Stroke + 1 / 3 * (sum_weight_end_Stroke - sum_weight_start_Stroke) +*

*1 / 3 * (MatLTC_stroke_Weighted_sum_end_cycle(index_cycle)*
- MatLTC_stroke_Weighted_sum_start_cycle(index_cycle))

The calculations underlying the above rely upon logic and estimates that span three six month periods: t-1 to t, t to t+1 and t+1 to t+2.

If the cycle is event the *MatLTC_stroke_Weighted_sum_start_cycle(index_cycle + 1)* takes the start of year value while the *MatLTC_stroke_Weighted_sum_end_cycle(index_cycle + 1))*takes the *sum_weight_half_Stroke* value.

If the cycle is odd the *MatLTC_stroke_Weighted_sum_start_cycle(index_cycle + 1)* takes the *sum_weight_half_Stroke* value while the *MatLTC_stroke_Weighted_sum_end_cycle(index_cycle + 1))* takes the end of year value.

But despite the manufacturer response to the ERG clarification question B32, the ERG still does not follow the logic of spanning three periods when calculating the probability of events. There is no obvious requirement for this, or for moving away from simply calculating the annual probability and applying half of this to the two cycles that fall within it.

The differences in the resulting six monthly probabilities are not large; e.g. 1.7% versus 1.6% in an example of the six-monthly stroke risk calculated by the ERG. But the differences may compound over the cycles of the model.

The manufacturer justifies this move from an annual calculation to a calculation based upon spanning three periods mainly, it appears, on grounds of wanting the probabilities to change as smoothly as possible between cycles. But the ERG cannot think of any requirement for this, and what would be small step changes between the annual probabilities calculated according to the UKPDS 68 would seem to be unobjectionable.

4.6.3 Treatment effects and 2nd order sampling

The treatment effects are estimated as an absolute change for empagliflozin to which additional "relative" treatment effects for the other treatments are added. For instance, suppose that the distribution for the absolute treatment effect of empagliflozin is $e \sim (\mu_e, \sigma_e)$. Suppose further that the distribution for the relative treatment effect of sitagliptin is $s \sim (\mu_s, \sigma_s)$.

If there is no second order sampling the treatment effect of empagliflozin is μ_e. And the treatment effect of sitagliptin is $\mu_e + \mu_s$.

With second order sampling for the ith PSA iteration a value is sampled from each of the distributions e_i and s_i. The value used in the modelling for the treatment effect of empagliflozin is e_i. But the value used for the treatment effect of sitagliptin is $\mu_e + s_i$. It may be the case that this should be $e_i + s_i$. If so, the ECEM systematically understates the degree of 2nd order uncertainty.

4.6.4 Derivation of UKPDS 68 covariances

Within the sampling of 2nd order uncertainty for a number of the UKPDS 68 parameters the ECEM contains not only the variances of the parameters, as derived from the standard errors reported in the UKPDS, but also the covariances between the parameter estimates. The ERG did query this, asking how the covariances had been derived. The manufacturer response was:

"*The co-variances were estimated from the standard errors for the coefficients stated in Clarke et al 2004*."

The ERG remains unclear how the full variance-covariance matrices for the UKPDS 68 parameters within the ECEM have been derived.

4.6.5 Dying twice

It appears that within a cycle a patient may die twice. For instance, with random number seeds 56 and 65 for the comparison of empagliflozin 10 mg with canagliflozin 100 mg during the first iteration it appears that patient number 54 dies of a stroke and also dies of an amputation[9]. The costs of both are incurred.

Given time constraints the ERG has not identified why this occurs within the visual basic. Whether the error in the VB may point to further errors and similar effects also being simulated for other model events is consequently unknown.

In itself, patients dying twice seems likely to be a relatively rare occurrence which may not particularly affect the results of the ECEM. Having confidence that this is the case is hampered by the issue around the random numbers for each patient apparently being the same for all patients within a PSA iteration. But it may point to what may be a more systemic failure of the ECEM development: a possible lack of model validation, stress testing and consequent robustness.

4.6.6 Random ordering of the sequencing of evaluation of events

Due to the UKPDS OM1 being implemented using fixed cycle lengths, as opposed to being a discrete event simulation, it randomly sequences the order in which the assessments of whether an event

[9] Again, a copy of the model that does this with the relevant outputs reported is available to the manufacturer upon request from the ERG.

happens or not during a cycle occurs are made. For instance, for a patient in the fifth cycle it may assess whether that patient experiences a myocardial infarction before it assesses whether that patient experiences a stroke. But in the next cycle these assessments may be randomised to occur such that whether the patient experiences a stroke is assessed prior to whether the patient experiences a myocardial infarction.

The submission notes that the intention of the ECEM it to do likewise. But due to *RandMatSeq_identifier* being initiated only once for each PSA iteration it appears that the sequence in which events are assessed is the same for each patient within a PSA iteration. But note that while this sequence is the same for each patient, it differs between cycles. As a consequence, it is unlikely to have any major impact upon the results of the ECEM.

4.6.7 Random sampling

Most of the distributions are specified as a mean, coupled with an upper and a lower bound. Within the visual basic these upper and lower bounds are treated as the upper and lower confidence limits, with a standard error being estimates as (UCI – LCI) / (2 * 1.96). This estimate of the standard error is then used in conjunction with the mean value to derive the α and β values for gamma and beta distributions. The ERG is not familiar with this method, but cross checking this sampling results in unbiased estimates with a standard error equal to that inferred.

It appears that the sampling procedure for the beta distribution may at times simulate a negative α value. The likelihood of this happening increases as the specified mean move closer to zero or one, and as the specified standard error increases. There is a procedure *sampled_value_beta_adj* for the beta distribution that attempts to correct for this. The ERG has not to date found any instances of this possible error.

Note that the *sampled_value_beta_adj* procedure catches the possibility of a negative α value within an error handler. Provided that this error handler is not invoked, the procedure produces a beta distribution centred around the mean with an appropriate standard error.

But if the error handler is invoked the calculation is then based upon the uniform distribution along the following lines.

```
lb_mean = mean - 1.96 * se
ub_mean = mean + 1.96 * se
If lb_mean < 0 Then
   lb_mean = 0
End If
```

```
If ub_mean > 1 Then
   ub_mean = 1
End If

If se > 0 Then
   sampled_value_beta_adj = (ub_mean - lb_mean) * rnd_nr + lb_mean
Else
   sampled_value_beta_adj = mean
End If
```

The above assume a uniform distribution rather than a beta distribution when the error handler applies. It also curtails the distribution at zero and one. While this ensures that sampled values lie within [0,1] it also changes the mean of the sampled distribution from that specified. The following might correct any bias.

```
lb_mean = mean - 1.96 * se
        ub_mean = mean + 1.96 * se
        if mean > 0.5 and ub_mean > 1 then
                ub_mean = 1
                lb_mean = mean - (ub_mean – mean)
Elseif
        if mean <= 0.5 and lb_mean < 0 then
                lb_mean = 0
                ub_mean = mean + (mean – lb_mean)
        EndIf
```

But since within the current ECEM implementation this is not used, it may not be a concern in any future correct ECEM implementation.

4.7 Elements of the ECEM that appear to work as intended

The model has the facility to age weight utilities. The ERG has not checked the implementation of this age weighting of utilities in any detail. But the ERG can confirm that given the current model inputs there is no age weighting of utilities.

The implementation of the adverse events that are modelled in addition to the UKPDS complications of diabetes; hypos, UTIs and GTIs, does model hypos as being experienced at the specified rates for a given treatment during the cycles that the patient remains on that treatment while UTIs and GTIs are

modelled as only applying during the first six-month cycle of a treatment. The appropriate appear to be applied to these.

The correct costs are applied to the complications of diabetes. This includes halving the first year costs and applying these in the cycle of the event and the cycle subsequent to the event, with the subsequent annual costs also being halved and applied to cycles thereafter.

While largely based upon just an examination of the visual basic code, it appears that the results are summed and discounted correctly. Mean values for costs and QALYs across the number of patients simulated are calculated when the model is run deterministically, with the variation across patients being reported in the model output. When the model is run probabilistically the mean values for costs and QALYs across the number of patients simulated is calculated for each PSA iteration. These are then averaged across the PSA iterations, with the variation across each PSA iteration's mean values being reported in the model output.

5 COST EFFECTIVENESS

Caveat: Everything in this chapter uses the ECEM model as provided by the manufacturer. The ERG concluded that the faults in this model makes it doubtful that its outputs are reliable. The manufacturer will have an opportunity to respond. The data in this chapter will only be reliable if the manufacturer can convince NICE that the model is sound.

The first sections report the manufacturer's methods, assumptions and results, and later ones present some critiques by the ERG. However for convenience some ERG comments are included in the earlier sections but in italics.

5.1 Summary and critique of manufacturer submitted economic evaluation by the ERG

5.1.1 Cost-effectiveness analysis

The cost-effectiveness analysis (CEA) uses a stochastic micro-simulation (the ECEM model) to estimate the cost-effectiveness of empagliflozin compared with SU, pioglitazone, sitagliptin, dapagliflozin, canagliflozin and insulin in adults with T2DM.

5.1.1.1 Comparison of economic submission with NICE reference case.

Table 29. NICE reference case checklist

Attribute	Reference case and TA Methods guidance	Does the *de novo* economic evaluation match the reference case
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice. The scope stipulates: For dual therapy: SUs, pioglitazone, DPP-4, GLP-1 analogues and dapagliflozin For triple therapy: pioglitazone, DPP-4, GLP-1 analogues, insulin For add-on to insulin: insulin	Not entirely. For dual therapy, SUs, pioglitazone and GLP-1 analogues are not considered in the economic evaluation. For triple therapy, GLP-1 analogue is not considered. For add-on to insulin, DPP-4s are considered as a comparator. The GLP-1 analogues are not considered. However, the ERG agrees with the manufacturer that GLP-1 analogues should not be a comparator in dual and triple therapy.
Patient group	As per NICE scope. The scope specifies that for adults with type 2 diabetes that is inadequately controlled on the previous line of therapy.	Yes, the HbA1c therapy switching values are as per the NICE T2DM guideline of 7.5%.
Perspective costs	NHS & Personal Social Services	Yes.
Perspective benefits	All health effects on individuals	Yes.
Form of economic evaluation	Cost-effectiveness analysis	Yes.
Time horizon	Sufficient to capture differences in costs and outcomes	Yes. 40 years.
Synthesis of evidence on outcomes	Systematic review	No
Outcome measure	Quality adjusted life years	Yes.
Health states for QALY	Described using a standardised and validated instrument	The EQ-5D values were derived from the UKPDS study.[35] For the complications of diabetes the manufacturer mainly uses the UKPDS 62 which estimates decrements through EQ-5D using the UK social tariff. A key HRQoL value relates to the direct impact of weight changes. Value for weight changes were taken from Bagust and Beale (2005).[36]
Benefit valuation	Time-trade off or standard gamble	TTO for the UK social tariff. TTO for the HRQoL impacts of weight changes.
Source of preference data for valuation of changes in HRQL	Representative sample of the public	Yes
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes.
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes.

Probabilistic modelling	Probabilistic modelling	Yes.
Sensitivity analysis		A range of univariate sensitivity analyses are undertaken.

5.2 Economic model

Three comparisons are made:

Dual therapy

- Empagliflozin 10 mg and 25 mg plus metformin versus sitagliptin 100 mg plus metformin.
- Empagliflozin 10 mg and 25 mg plus metformin versus dapagliflozin 10 mg plus metformin.
- Empagliflozin 10 mg and 25 mg plus metformin versus canagliflozin 100 mg and 300mg plus metformin.

Triple therapy

- Empagliflozin 10 mg and 25 mg plus metformin plus SU versus metformin plus SU plus sitagliptin 100 mg
- Empagliflozin 10 mg and 25 mg plus metformin plus SU versus metformin plus SU plus dapagliflozin 10mg.
- Empagliflozin 10 mg and 25 mg plus metformin plus SU versus metformin plus SU plus canagliflozin 100 mg and 300 mg.

Add-ons to insulin regimens

- Empagliflozin 10 mg and 25 mg versus metformin plus sitagliptin 100 mg.
- Empagliflozin 10 mg and 25 mg versus metformin plus dapagliflozin 10 mg.
- Empagliflozin 10 mg and 25 mg versus metformin plus canagliflozin 100 mg and 300 mg.

5.3 Population

The population considered in the model mainly reflected the T2DM population from the UKPDS study[37] but the age of patients initiating dual therapy was obtained from the Empa Reg trial 2-year evaluating the long term efficacy of empagliflozin 25 mg once daily compared with glimepiride as add-on therapy to metformin; and the age of patients on triple therapy or insulin was based on the ages used in the T2DM NICE clinical guideline (see Table 30 reproduced from table 49 of the manufacturer submission).

Table 30**. Patient characteristics in the cost effectiveness analysis.**

	Dual therapy	**Triple therapy**
Age when initiating assessed treatment [years]	56‡	58†
Female	45%**	39%†
Afro-Caribbean	1.2%	8%
Asian (mainly south Asian)	1.5%	10%
Smoking status	31%	31%
Peripheral vascular disease	23.5%^	23.5%
HbA1c when starting dual therapy	7.5%	-
HbA1c when starting triple therapy	-	7.5%
Systolic blood pressure when initiating assessed treatment (mm Hg	135†	140†
Total cholesterol:HDL when initiating assessed treatment	5.05**	4.4†
BMI when initiating assessed treatment [kg/m2]	27.5†	30.42†
	From Table 50	
IHD	3.2%	
stroke	1.4%	
Blindness in one eye	1.2%	

5.4 Interventions and comparators

In all the comparisons that follow, the evidence is taken from the NMA.

5.4.1 Dual therapy

Empagliflozin plus metformin is compared with:

- Metformin plus sitagliptin
- Metformin plus dapagliflozin
- Metformin plus canagliflozin

The combination of a SU and empagliflozin was not included as there is no evidence from trials.

5.4.2 Triple therapy

Empagliflozin plus SU plus metformin is compared with:

- Sitagliptin plus SU plus metformin
- Canagliflozin plus SU plus metformin

Empagliflozin plus pioglitazone plus metformin is compared with:

- Sitagliptin plus pioglitazone plus metformin
- Canagliflozin plus pioglitazone plus metformin

5.4.3 Insulin regimens

Empagliflozin as an add- on to insulin is compared with:

- Sitagliptin plus insulin
- Dapagliflozin plus insulin
- Canagliflozin plus insulin

The MS presents results from the economic evaluation for the above comparisons. We focus on the comparisons with sitagliptin, dapagliflozin and canagliflozin in dual therapy, triple therapy and add-on with insulin.

5.5 Perspective, time horizon and discounting

The perspective is as per the NICE guidelines: the patient perspective for benefits and the NHS/PSS for costs. The time horizon is 40 years, with costs and benefits being discounted at an annual rate of 3.5%.

5.6 Treatment effectiveness and extrapolation

5.6.1 Treatment effectiveness: empagliflozin as an add-on to metformin

The effectiveness estimates are drawn from the NMA. Events in the model were implemented as relative risk applied to the baseline rate for metformin and placebo from the NMA. The treatment effects were modelled using data from patients who received treatment for 24 weeks; and treatment effects for 52 weeks data were used in a sensitivity analysis. Efficacy was estimated firstly in terms of the effect of the alternative treatments on HbA1c levels.

For the impact upon weight, the treatment could result in a weight loss, weight increase or maintenance of weight. Two scenarios were defined to determine the progress of weight over time. (Pages 228-9). For the first scenario which was also assumed to be the base case scenario, the manufacturer assumed that every time a patient initiated a new treatment, the weight change associated with the new treatment would happen gradually (in the model the number of years to achieve weight change was entered as one year) as long as the patient continued receiving the same treatment. The weight change was assumed to be maintained over 6 months after the full weight change has been achieved. .

The second scenario assumed that the full weight change would happen during the first cycle after initiating the new treatment and would maintain that weight change as long as the patient continued receiving the same treatment. The main difference between the scenarios is duration of weight change.

The model also assumed that whenever a patient switched to a different treatment, the weight at the beginning of the new treatment would be equal to their baseline weight plus 0.1 kg multiplied by the number of years since treatment initiation plus the impact of weight of the new treatment.

Figures 42, 45 and 51 of the submission are reproduced below and show weight change with the main comparators. Colours do not reproduce but in the period from baseline to 2 years, there is 2-3 kg weight loss with all three flozins (the five lowest lines) but little weight loss with sitagliptin (the top line in that period)

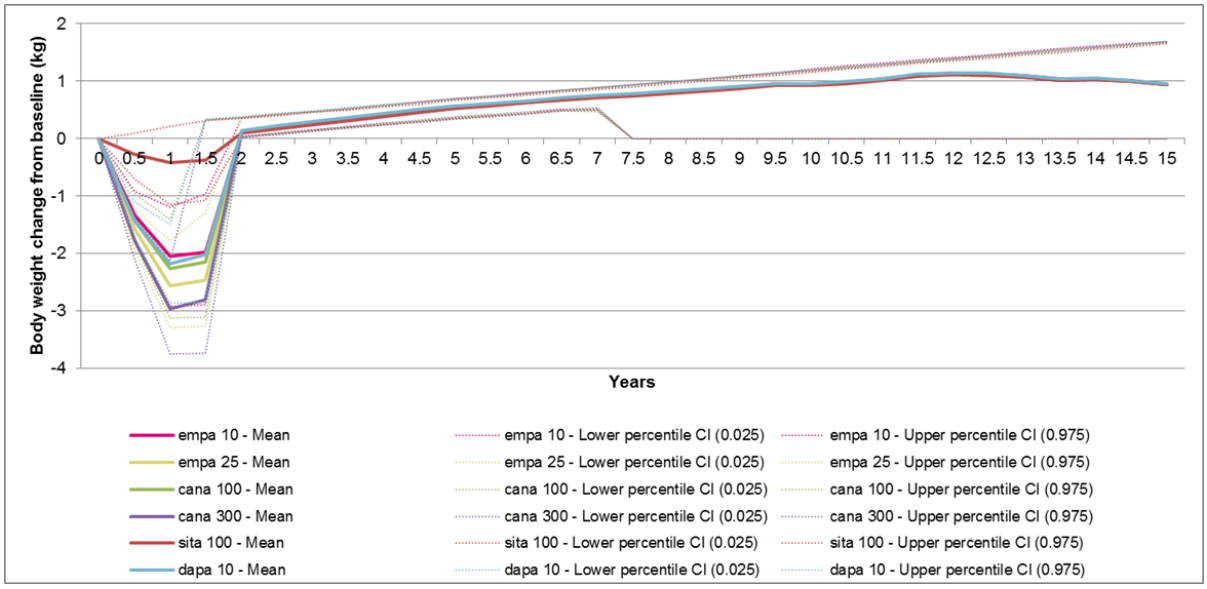


Figure 19. Dual therapy – Metformin – Weight

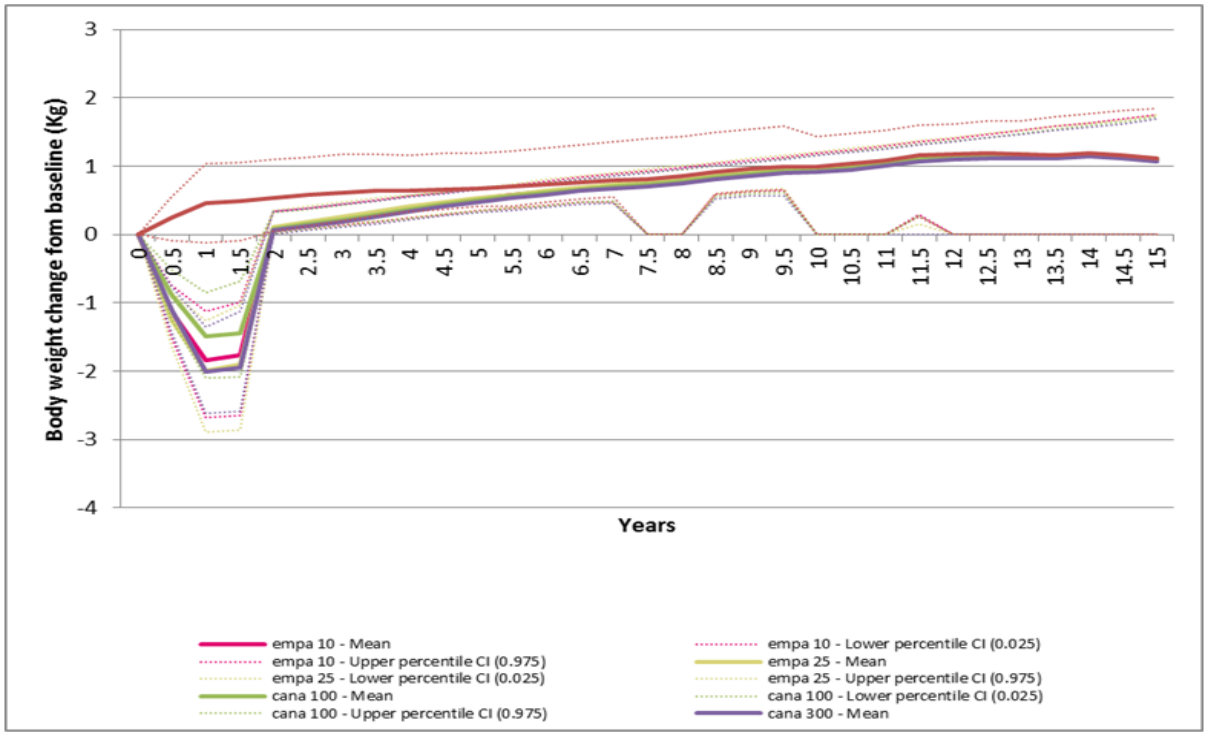


Figure 20. Triple therapy – Metformin and SU – Weight over time

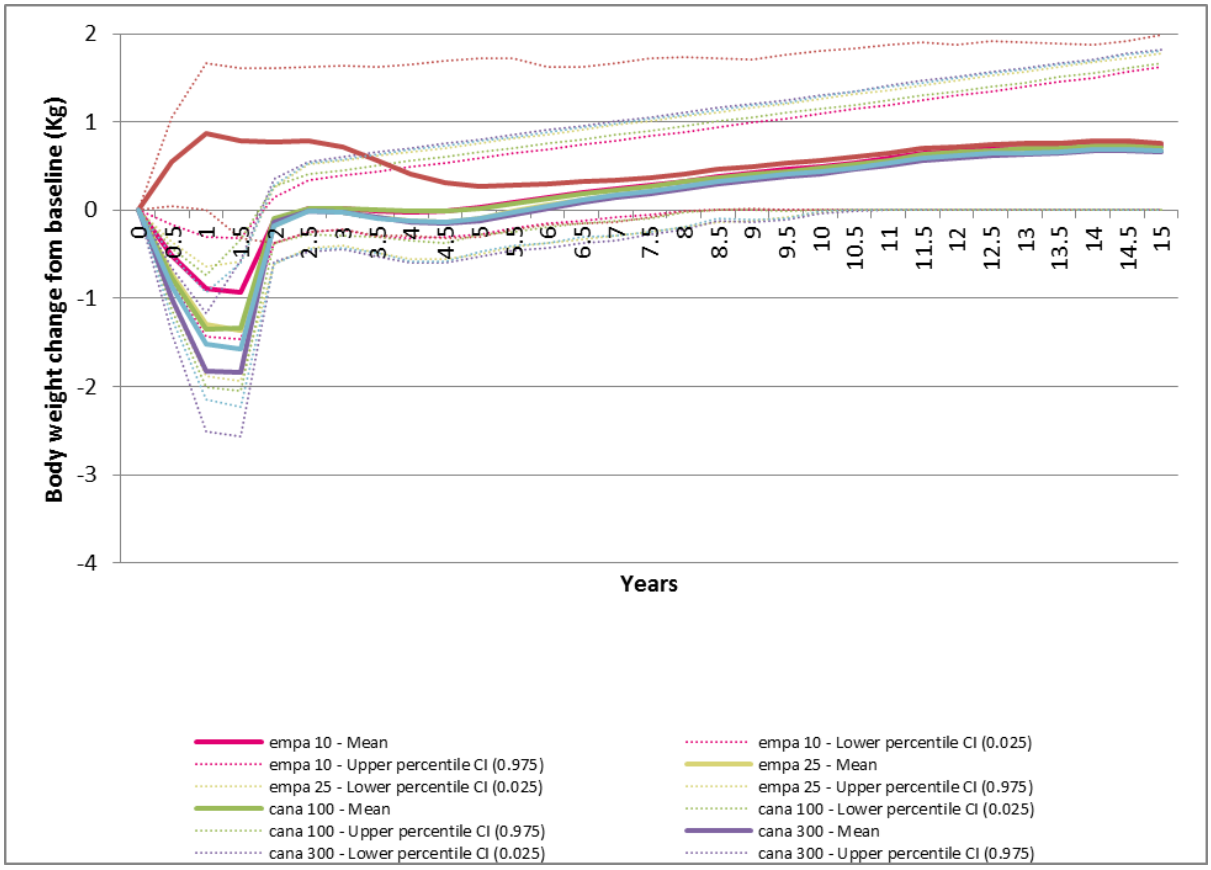


Figure 21. Insulin combinations – Weight over time

Table 31. Empagliflozin as an add-on to Metformin: treatment effect (model inputs) on HbA1c, SBP, weight and adverse events (drawn from table 56 of manufacturer's submission).

Risk factors (change from baseline)				
	Met+Sita (relative response to Empag 10 mg)	**Met+Empag 10 mg (absolute response)**	**Met+Empag 25mg (absolute response)**	**Met + Sita (relative response to Empag 25 mg)**
HbA1C	0.19	0.65	0.75	0.09
SBP	-2.12	6.504	7.304	-2.9
weight loss	2.1	-2.63	-3.13	2.62
	Met+Dapa (relative response to Empag 10 mg)	**Met+Empag 10 mg (absolute response)**	**Met+Empag 25mg (absolute response)**	**Met + Dapa (relative response to Empag 25 mg)**
HbA1C	-0.04	0.65	0.75	-0.14
SBP	-0.52	6.504	7.304	-1.33
weight loss	-0.27	-2.63	-3.13	0.23
	Met+Cana 100mg (relative response to Empag 10 mg)	**Met+Empag 10 mg (absolute response)**	**Met+Empag 25mg (absolute response)**	**Met + Cana 300mg (relative response to Empag 25 mg)**
HbA1C	0.06	0.65	0.75	0.09
SBP	-0.15	6.504	7.304	0.24
weight loss	-0.27	-2.63	-3.13	-0.45
Adverse events				
	Met+Sita (relative response to Empag 10 mg)	**Met+Empag 10 mg (absolute response)**	**Met+Empag 25mg (absolute response)**	**Met + Sita (relative response to Empag 25 mg)**
Discontinuation	0.008	0.0276	0.0329	0.008
Hypoglycaemia Non severe	0.0089	0.0738	0.0508	0.0089
UTI	0.027	0.0357	0.0338	0.027
GTI	-	0.0369	0.0467	-
	Met+Dapa (relative response to Empag 10 mg)	**Met+Empag 10 mg (absolute response)**	**Met+Empag 25mg (absolute response)**	**Met + Dapa (relative response to Empag 25 mg)**
Discontinuation	0.059	0.0276	0.0329	0.059
Hypoglycaemia Non severe	0.0244	0.0738	0.0508	0.0244
UTI	0.0685	0.0357	0.0338	0.0685
GTI	0.1230	0.0369	0.0467	0.1230
	Met+Cana 100mg (relative response to Empag 10 mg)	**Met+Empag 10 mg (absolute response)**	**Met+Empag 25mg (absolute response)**	**Met + Cana 300mg (relative response to Empag 25 mg)**
Discontinuation	0.041	0.0276	0.0329	0.051
Hypoglycaemia Non severe	0.0402	0.0738	0.0508	0.0244
UTI	0.0634	0.0357	0.0338	0.0629
GTI	0.0634	0.0369	0.0467	0.0629

ERG comment. The ERG found these figures puzzling, because they do not fit with what was reported in the NMA in Table 31 of the submission. For example, the HbA1c figures in the third row of the table above are 0.65 and 0.75%, but those from the NMA are 0.60 and 0.70%. The manufacturer has explained that the figures of 0.60 and 0.70% are the differences from the change in the metformin and placebo comparator, which was 0.05%. The manufacturer explained in the response to clarification that;

"The value of 0.05% reflects the full treatment effect of metformin + placebo, while 0.60 reflects the relative effect of empagliflozin 10g vs. metformin+ placebo. Therefore the absolute effect for met + empagliflozin will be 0.05 + 0.60 = 0.65"

The source of the 0.05% is unclear – the ERG found this figure only in the model, not the text, but it looks plausible.

We note that in the ER MET trial, the reduction in HbA1c was 0.70% with 10mg empagliflozin and 0.13% with placebo, a difference of 0.57%.(Table 6, ERG report). This is contrasts with the figure of 0.65% used in the modelling. For the 25mg dose, the effect size from the trial was 0.64%, compared to the 0.75% used in the modelling.

The same method seems to have been used throughout the modelling. However, this affects the figures in columns 3 and 4 above, but not the absolute differences reported in columns 2 and 5, which match the figures in table 31.

5.6.2 Treatment effectiveness: Empagliflozin in triple therapy

The clinical effectiveness (model inputs) are shown in Table 32 drawn from Table 57 of the manufacturer's submission.

The values reported in the table below are the clinical efficacy estimates based on the results from the NMA. The efficacy estimates in the table below relates to the clinical efficacy estimates of the reduction from baseline in HbA1c, SBP and weight loss for empagliflozin treatment combinations, and the relative responses of Met+ SU+ sitagliptin, canagliflozin and dapagliflozin when compared to empagliflozin combinations.

Table 32. Empagliflozin in triple therapy: treatment effect (model inputs) on HbA1c, SBP, weight and adverse events.

Risk factors : Baseline (Met +SU)				
	Met+SU+Sita (relative response to Empag 10 mg)	**Met+SU+Empag10 mg (absolute response)**	**Met+SU+Empag 25mg (absolute response)**	**Met+SU+Sita (relative response to Empag 25 mg)**
HbA1C	0.16	0.79	0.74	0.21
SBP	-6.93	3.26	2.64	-6.31
weight loss	2.74	-2.26	-2.49	2.97
	Met+SU+Dapa 10mg (relative response to Empag 10 mg)	**Met+SU+Empag10 mg (absolute response)**	**Met+SU+Empag 25mg (absolute response)**	**Met+SU+ Dapa 10mg (relative response to Empag 25 mg)**
HbA1C	0.36	0.79	0.74	0.41
SBP	-0.94	3.26	2.64	-0.35
weight loss	0.07	-2.26	-2.49	0.31
	Met+SU+Cana100mg (relative response to Empag 10 mg)	**Met+SU+Empag10 mg (absolute response)**	**Met+SU+Empag 25mg (absolute response)**	**Met+SU+ Cana 300mg (relative response to Empag 25 mg)**
HbA1C	0.11	0.79	0.74	0.41
SBP	-0.37	3.26	2.64	-0.35
weight loss	0.51	-2.26	-2.49	0.31
Adverse events : (Met+ SU)				
	Met+SU+Sita (relative response to Empag 10 mg)	**Met+SU+Empag10 mg (absolute response)**	**Met+SU+Empag 25mg (absolute response)**	**Met+SU+Sita (relative response to Empag 25 mg)**
Discontinuation	0.008	0.0268	0.0323	0.008
Hypoglycaemia Non severe	0.1345	0.1394	0.1071	0.1345
UTI	0.0298	0.0806	0.0645	0.298
GTI	-	0.0268	0.0230	-
	Met+SU+Dapa 10mg (relative response to Empag 10 mg)	**Met+SU+Empag10 mg (absolute response)**	**Met+SU+Empag 25mg (absolute response)**	**Met+SU+ Dapa 10mg (relative response to Empag 25 mg)**
Discontinuation	0.0323	0.0268	0.0323	0.323
Hypoglycaemia Non severe	0.2291	0.1394	0.1071	0.2291
UTI	0.0629	0.0806	0.0645	0.0629
GTI	0.0629	0.0268	0.0230	0.0629
	Met+SU+Cana100mg (relative response to Empag 10 mg)	**Met+SU+Empag10 mg (absolute response)**	**Met+SU+Empag 25mg (absolute response)**	**Met+SU+ Cana 300mg (relative response to Empag 25 mg)**
Discontinuation	0.0268	0.0268	0.0323	0.0323
Hypoglycaemia Non severe	0.1884	0.1394	0.1071	0.2291
UTI	0.0634	0.0806	0.0645	0.0629
GTI	0.0634	0.0268	0.0230	0.0629

In table 59 of the MS, the probability of experiencing an event of non-severe hypoglycemia for dapagliflozin is missing. The values reported above were sourced from the model. The adverse events for canagliflozin 300 mg was assumed to be equivalent to dapagliflozin 10 mg.

The weight losses were modelled similar to dual therapy, assumed to be maintained for 6 months. As noted earlier, this may be unduly pessimistic, based on the extension study which showed weight loss maintained till at least 76 weeks in dual therapy.

5.6.3 Treatment effectiveness: Empagliflozin as an add-on to insulin

The effectiveness estimates were drawn from the NMA. The clinical effectiveness estimates are shown in Table 33 based on table 59 of the Boehringer submission

Table 33. Empagliflozin as an add-on to insulin: treatment effect (model inputs) on HbA1c, SBP, weight and adverse events (sourced from table 59 of manufacture's submission).

	Insulin+ Sita (relative response to Empag 10 mg)	Insulin+Empag 10 mg	Insulin+Empag 25mg	Insulin+Sita (relative response to Empag 25 mg)
Risk factors (change from baseline)				
HbA1C	0.10	0.69	0.76	0.03
SBP	-4.751	4.751	4.141	-4.141
weight loss	2.07	-1.037	-1.457	2.49
Adverse events				
Discontinuation	0.0479	0.0538	0.0476	0.0479
Hypoglycaemia Non severe	0.2338	0.2555	0.2579	0.2338
Hypoglycaemia severe	0.0096	0.0096	0.096	0.096
UTI	0.0630	0.1559	0.1534	0.063
GTI	0.003	0.0430	0.0952	0.0030
	Insulin+ Dapa 10mg (relative response to Empag 10 mg)	Insulin+Empag 10 mg	Insulin+Empag 25mg	Insulin+ Dapa 10mg (relative response to Empag 25 mg)
HbA1C	0.07[a]	0.69	0.76	0.00[c]
SBP	0.27	4.751	4.141	0.88
weight loss	-0.67[b]	-1.037	-1.457	-0.26
Adverse events				
Discontinuation	0.0510	0.0538	0.0476	0.0051
Hypoglycaemia Non severe	0.2458	0.2555	0.2579	0.2458
Hypoglycaemia severe	0.0096	0.0096	0.0096	0.0096
UTI	0.056	0.1559	0.1534	0.056
GTI	0.0629	0.0430	0.0952	0.0629
Risk factors (change from baseline)				
	Insulin+ Cana 100mg (relative response to Empag 10 mg)	Insulin+Empag 10 mg	Insulin+Empag 25mg	Insulin+ Cana 300mg (relative response to Empag 25 mg)
HbA1C	0.14	0.69	0.76	0.16
SBP	-0.26	4.751	4.141	2.14
weight loss	-0.49	-1.037	-1.457	-0.55
Adverse events				
Discontinuation	0.0708	0.0538	0.0476	0.0051
Hypoglycaemia Non severe	0.3470	0.2555	0.2579	0.3347
Hypoglycaemia severe	0.0096	0.0096	0.0096	0.0096
UTI	0.063	0.1559	0.1534	0.0629
GTI	0.063	0.0430	0.0952	0.0629

[a] The value in the model states 0.07 but ERG think the value should be 0.001.[b] The value in the model states -0.67 but ERG think the value should be -0.07.[c] The value in the model states 0.00 but ERG think the value should be 0.069.

ERG. As before, the change seen on insulin alone is added to the effect size of insulin + empagliflozin, but this does not seem to affect the figures in columns 2 and 5.

To recap, the differences seen in the trial of addition of empagliflozin to basal insulin regimens were 0.56% for 10mg (reduction on placebo 0.01, on empa 0.57%) and 0.70% for the 25mg dose (reduction on empa 0.71%), somewhat smaller than the figures used in the modelling – differences of 0.13% and 0.6% respectively.

5.7 Extrapolation: therapy switch

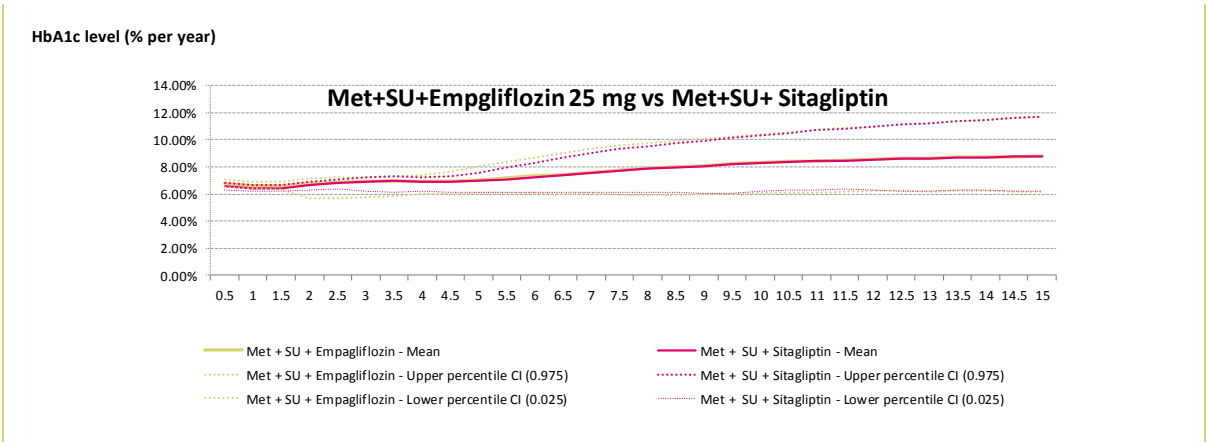
The model contains the facility for patients to switch therapy once their HbA1c rises above a user specified threshold level, assumed to be 7.5%.

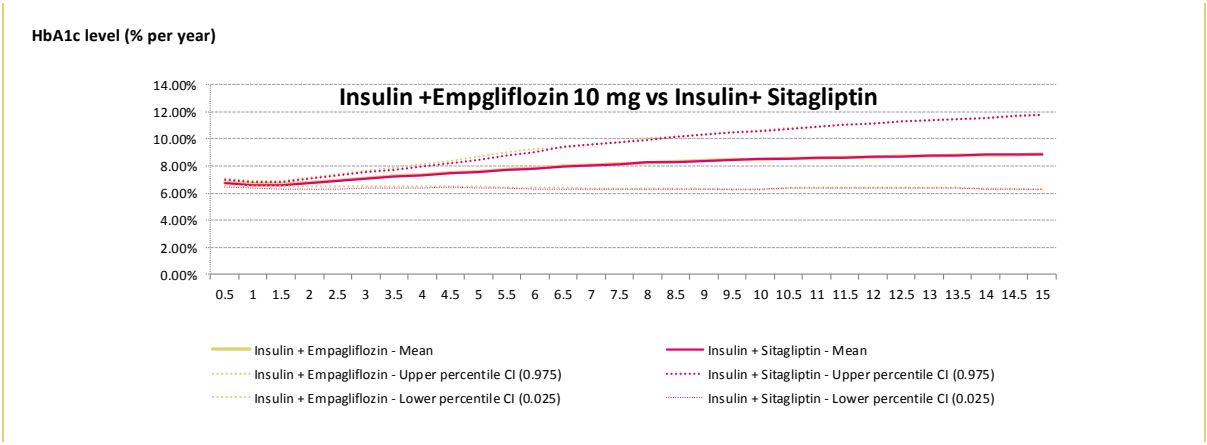
5.7.1 Extrapolation: HbA1c

HbA1c is modelled using equation 11 of the UKPDS 68.[37]

Figure 17, reproduced by running the ECEM model of the submission gives a graphical presentation of the extent to which the modelling permits the HbA1c to rise above the NICE guideline of 7.5% The evolution of HbA1c level does not differ amongst the dual, triple and insulin comparisons of empagliflozin with sitagliptin. The HbA1c levels were similar for other dual, triple and add-on insulin comparison (See Figure 23, Figure 24 and Figure 25 reproduced from figure 40, 43 and 49 of the submission).

The extrapolation of HbA1c levels across treatments (from ECEM model).





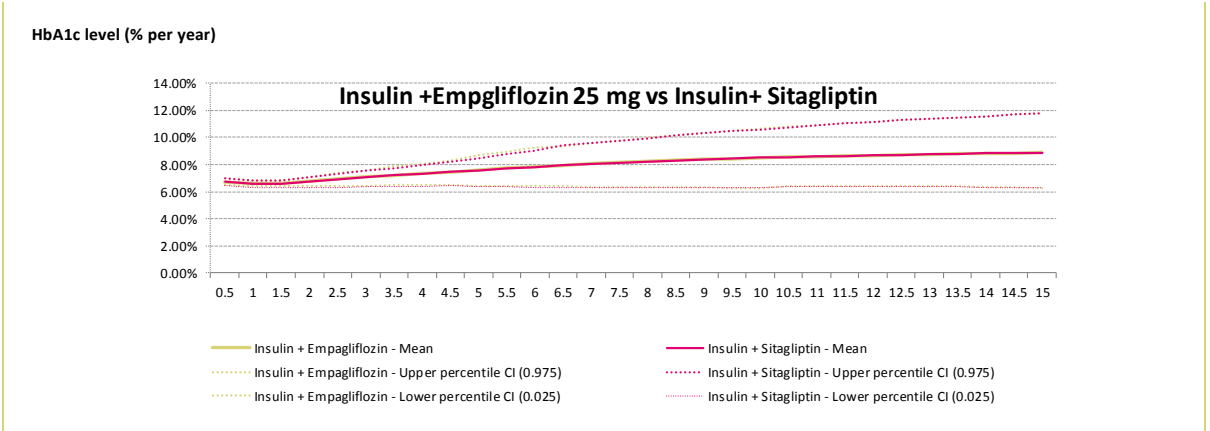


Figure 22. The extrapolation of HbA1c levels across treatments (from ECCM model).

The extrapolation of HbA1c levels across treatments (reproduced from MS – figure 40,43 & 49)

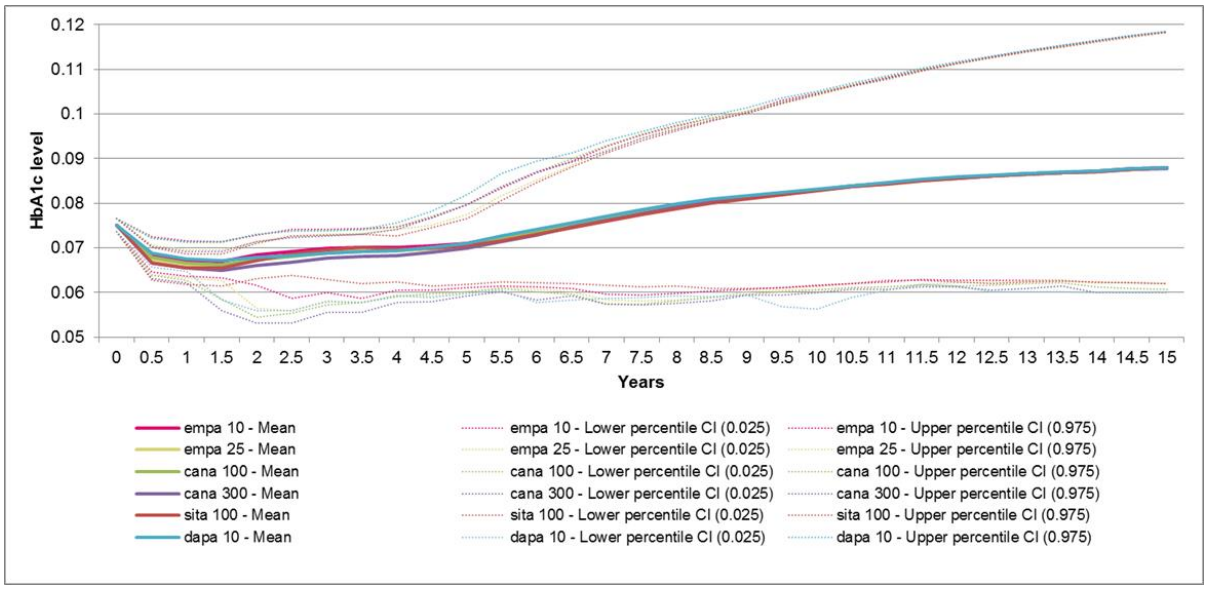


Figure 23. Dual therapy – Metformin - HbA1c level over time from MS

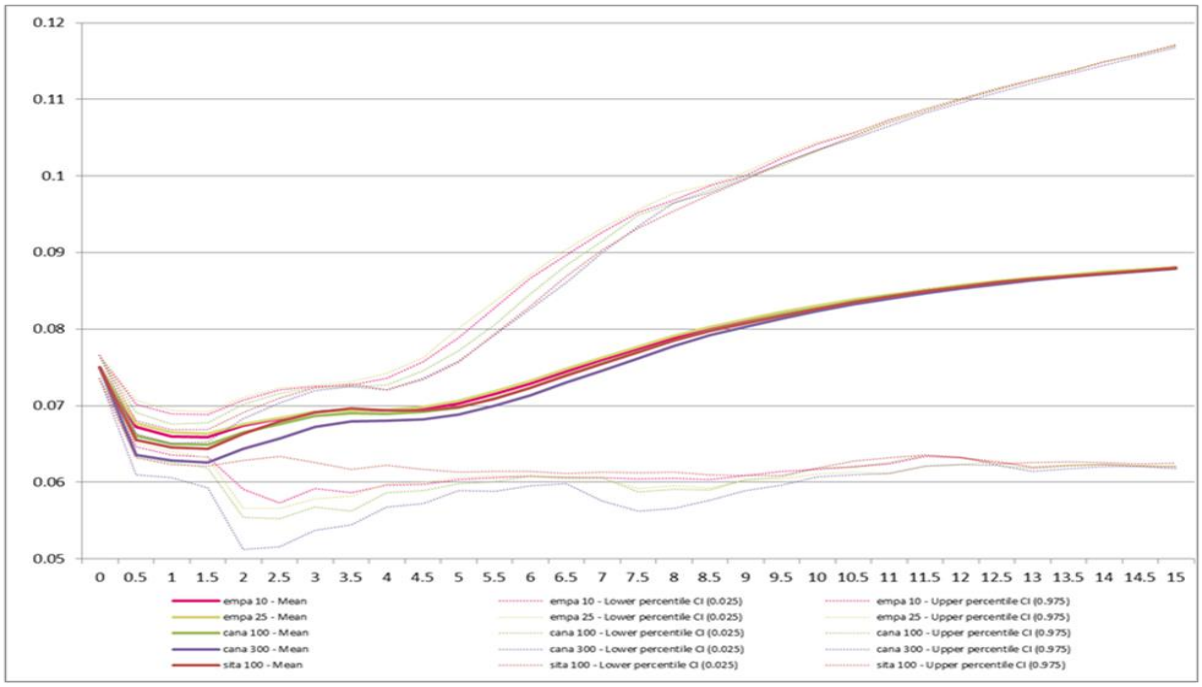


Figure 24. Triple therapy – Metformin and SU - HbA1c level over time from MS

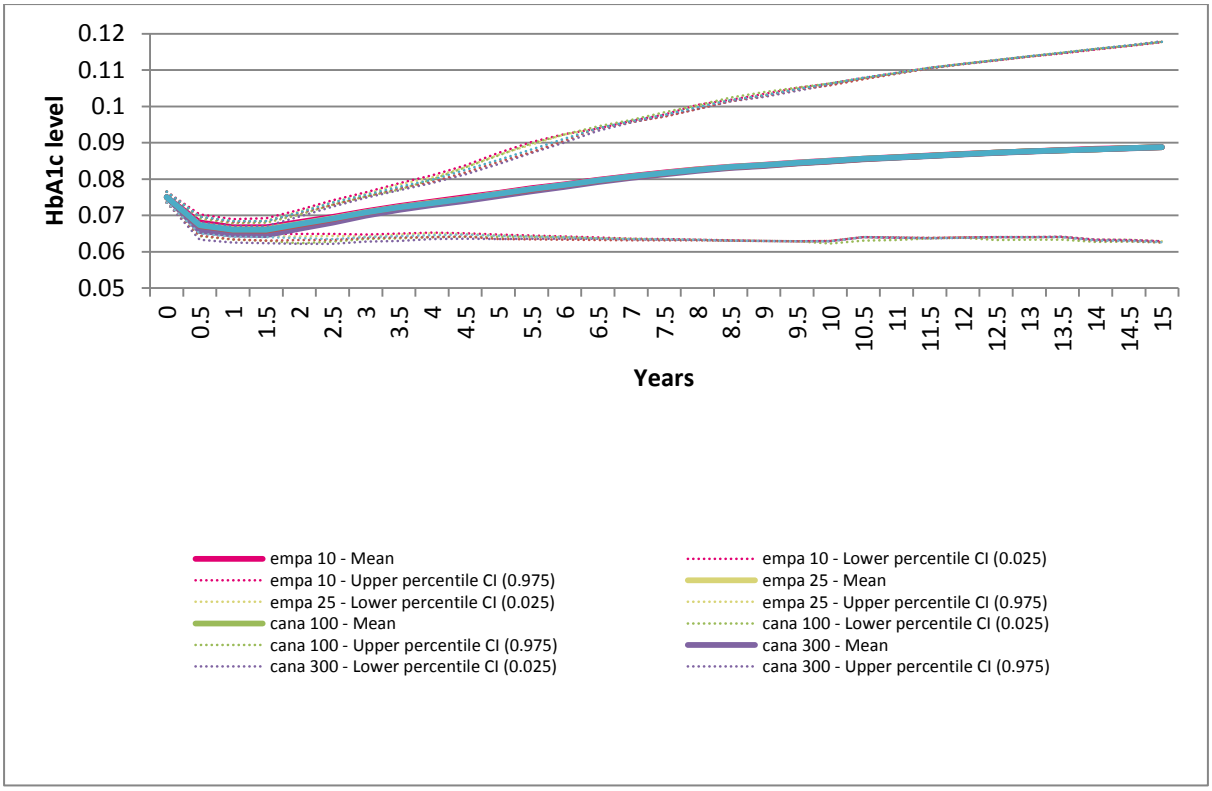


Figure 25. Insulin combination - HbA1c level over time from MS

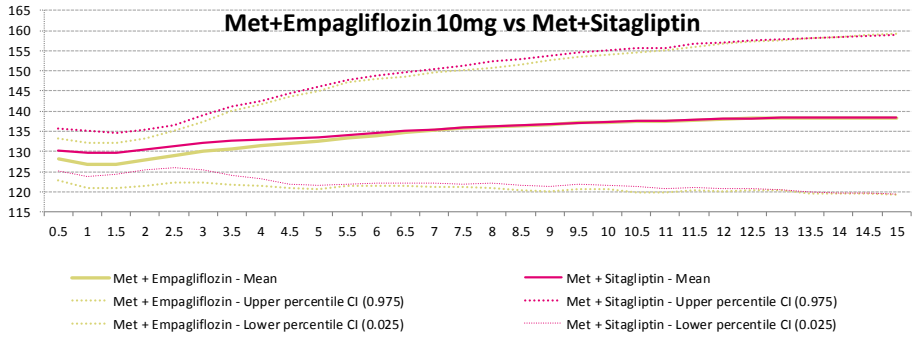
5.8 Extrapolation: SBP

SBP is modelled using equation 12 of the UKPDS 68.[37] There is a difference in the evolutions of SBP between the arms for the dual therapy comparison of empagliflozin with sitagliptin and the triple therapy comparison of empagliflozin with sulphonylurea (see Figure 26). The difference in SBP is significant for 3 years but after 3 years SBP converge.

The extrapolation of SBP levels across treatments (from ECCM model).

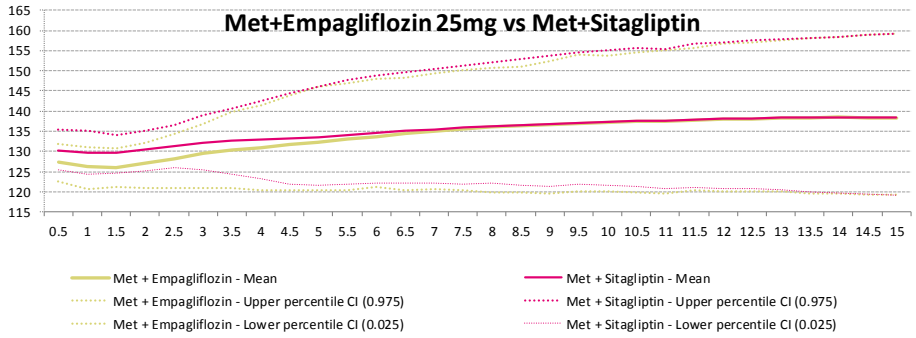
SBP level (mmHg per year)

Top ↑



SBP level (mmHg per year)

Top ↑



SBP level (mmHg per year)

Top ↑

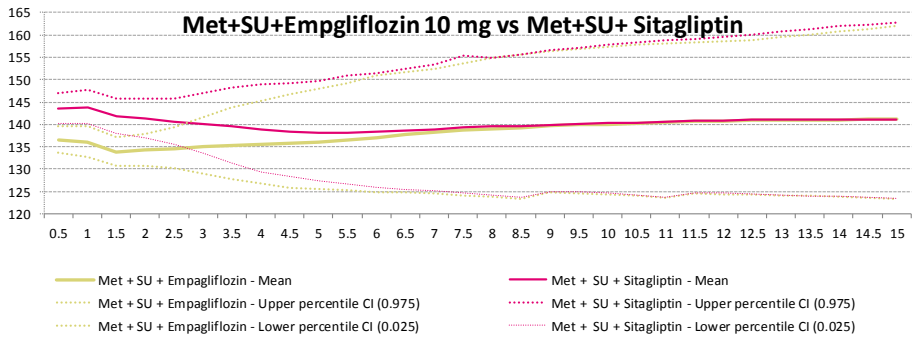


Figure 26. The extrapolation of SBP levels across treatments (from ECCM model).

The differences in SBP were similar for other dual, triple and add-on insulin comparisons (See Figure 27, Figure 28 and Figure 29 reproduced from figure 41, 44 and 50 of the submission).

The extrapolation of SBP across treatments (reproduced from MS – figure 41, 44 & 50)

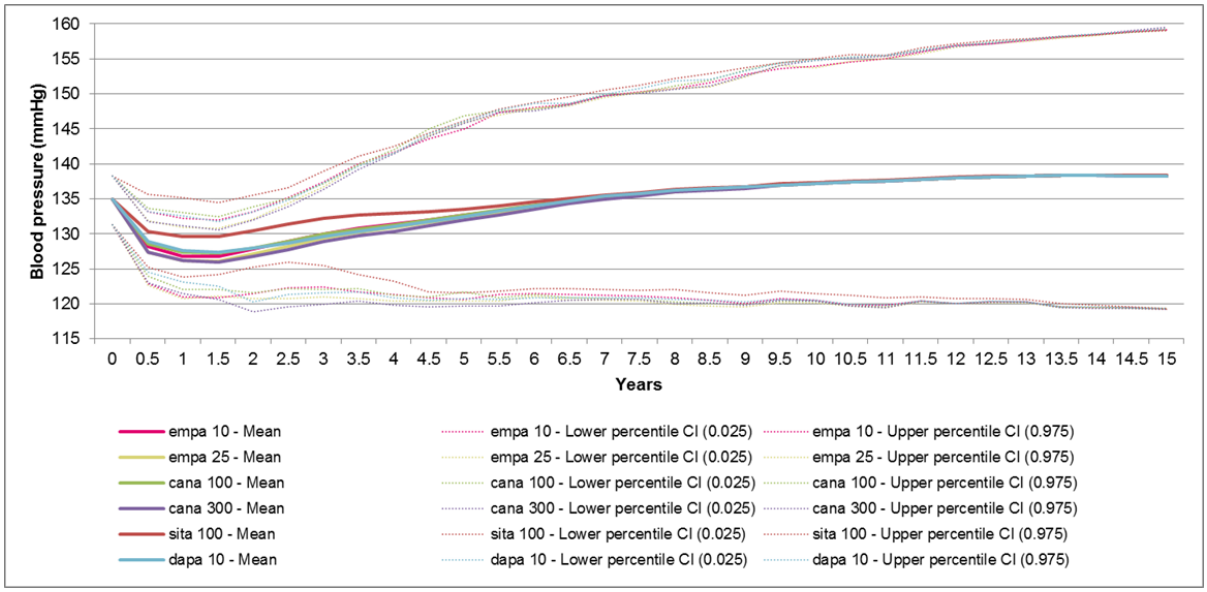


Figure 27. Dual therapy – Metformin – Systolic Blood Pressure from MS.

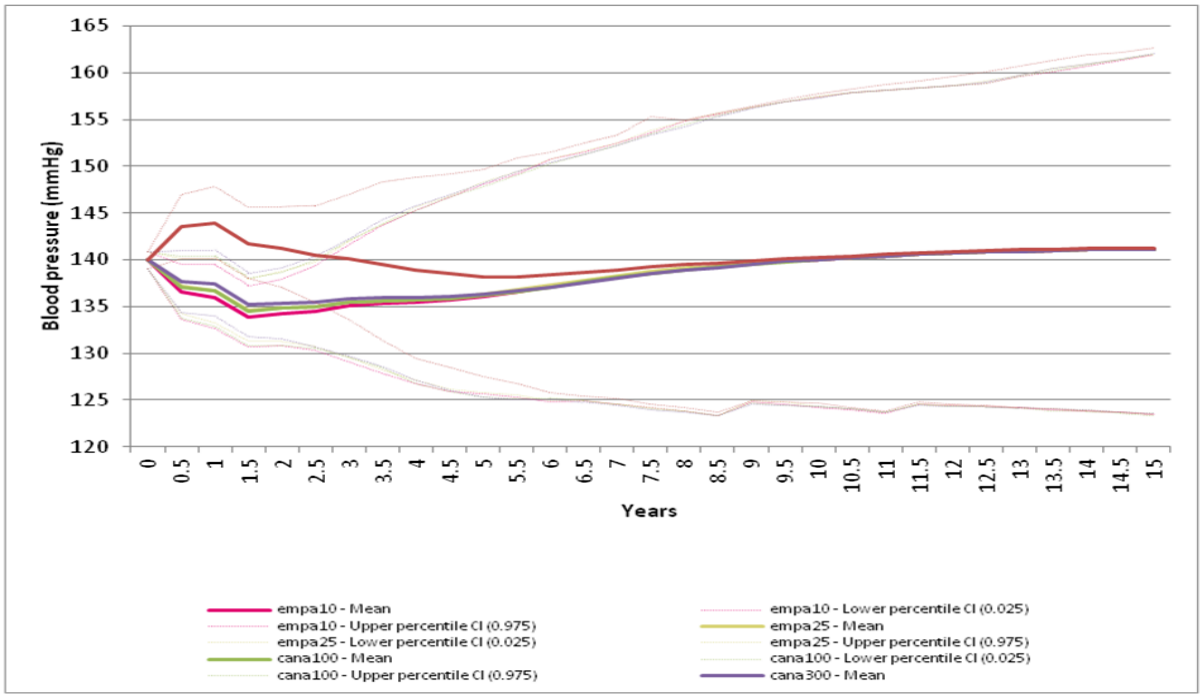


Figure 28. Metformin and SU – SBP over time from MS

Curiously, SBP rises initially on empagliflozin 10mg

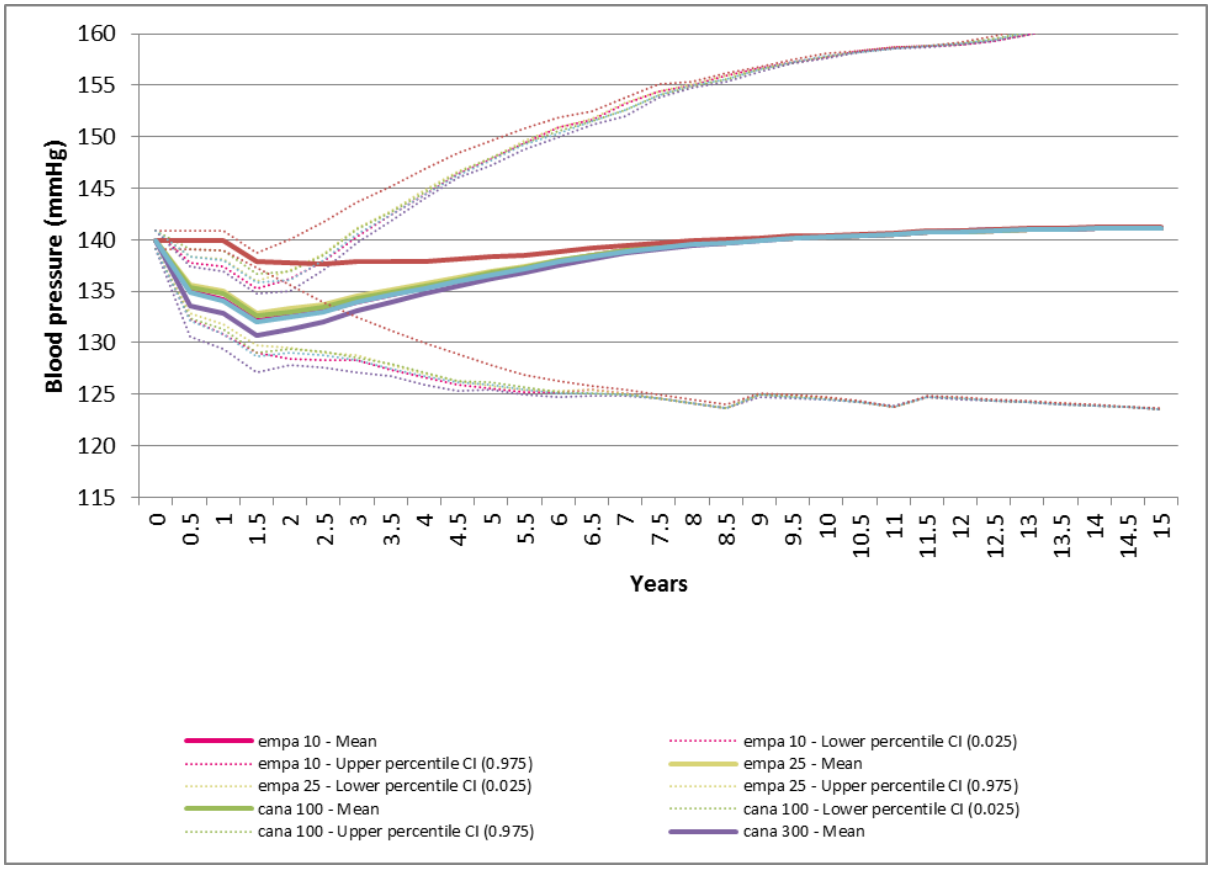


Figure 29. Insulin combination – SBP over time from MS

5.9 Extrapolation: Weight

The model applies the initial treatment effect, and there is a plateau followed by a loss of effect for those treatments that reduce weight and this is maintained for 6 months. For those treatments that increase weight the model assumes a similar effect.

The graphs presenting the evolution of weight over time demonstrate that for the comparison of empagliflozin with sitagliptin all the graphs demonstrate an initial decrease in weight over two years and then gradually increasing back to the baseline level. Thereafter there is a convergence in weight across all the treatments by 7 years (See figure 3).

The extrapolation of weight change levels across treatments (from ECCM model).

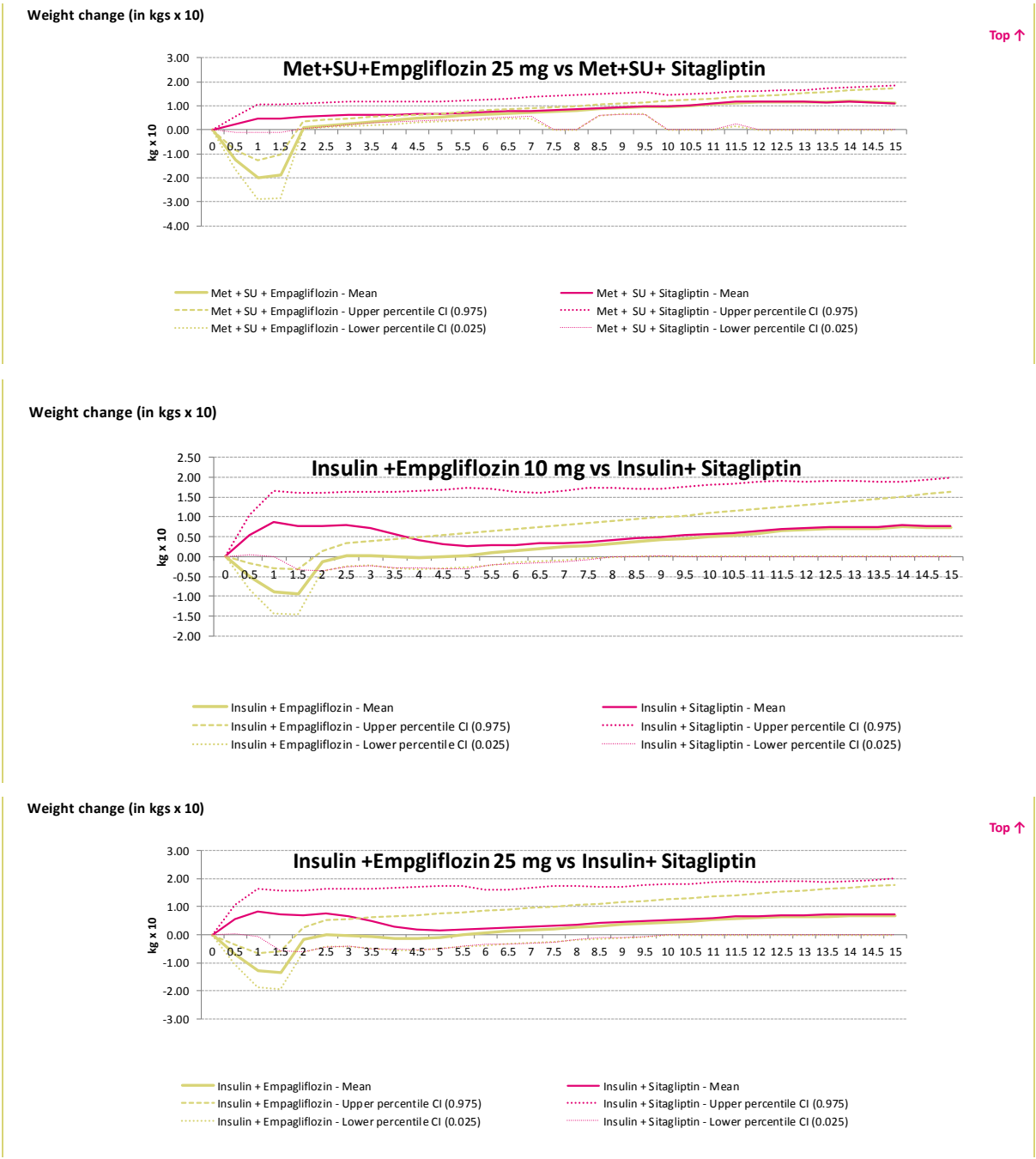


Figure 30. The extrapolation of weight change levels across treatments (from ECCM model).

The evolution of weight were similar for other dual, triple and add-on insulin comparison (See Figure 31, Figure 32 and Figure 33 reproduced from figure 42 , 45 and 51 of the submission).

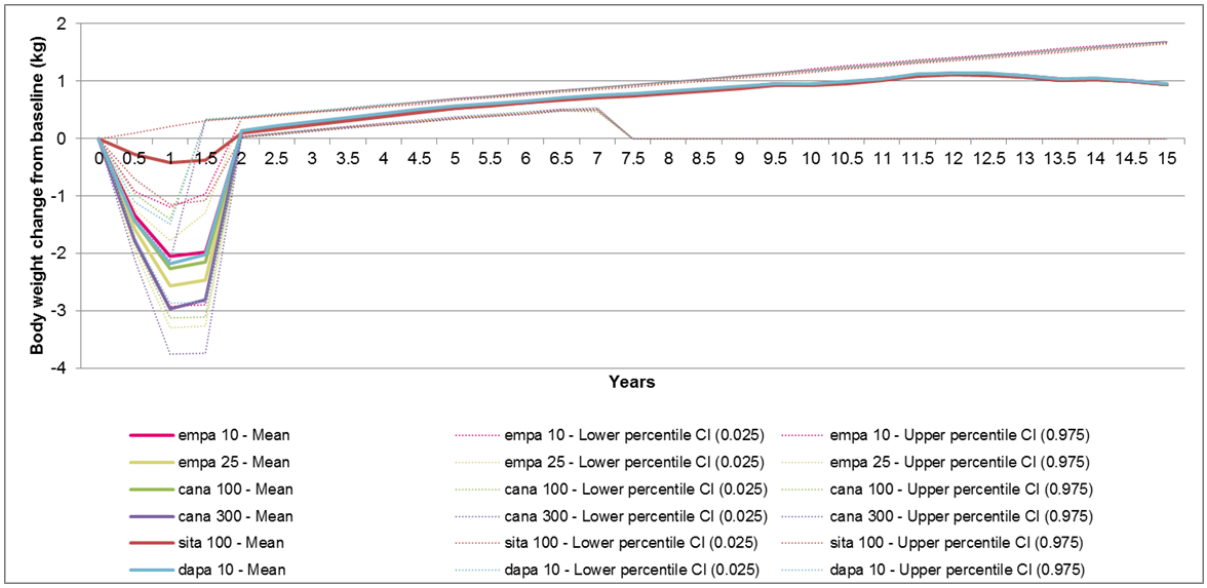


Figure 31. Dual therapy – Metformin – Weight

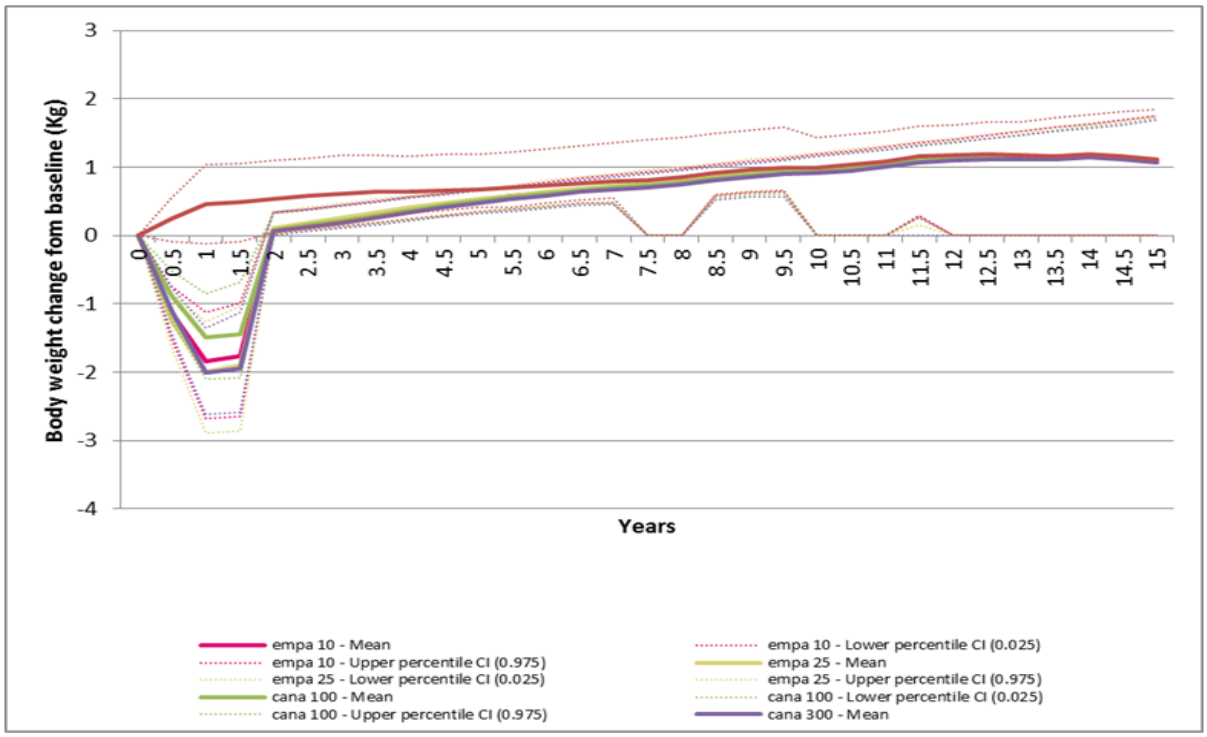


Figure 32. Triple therapy – Metformin and SU – Weight over time

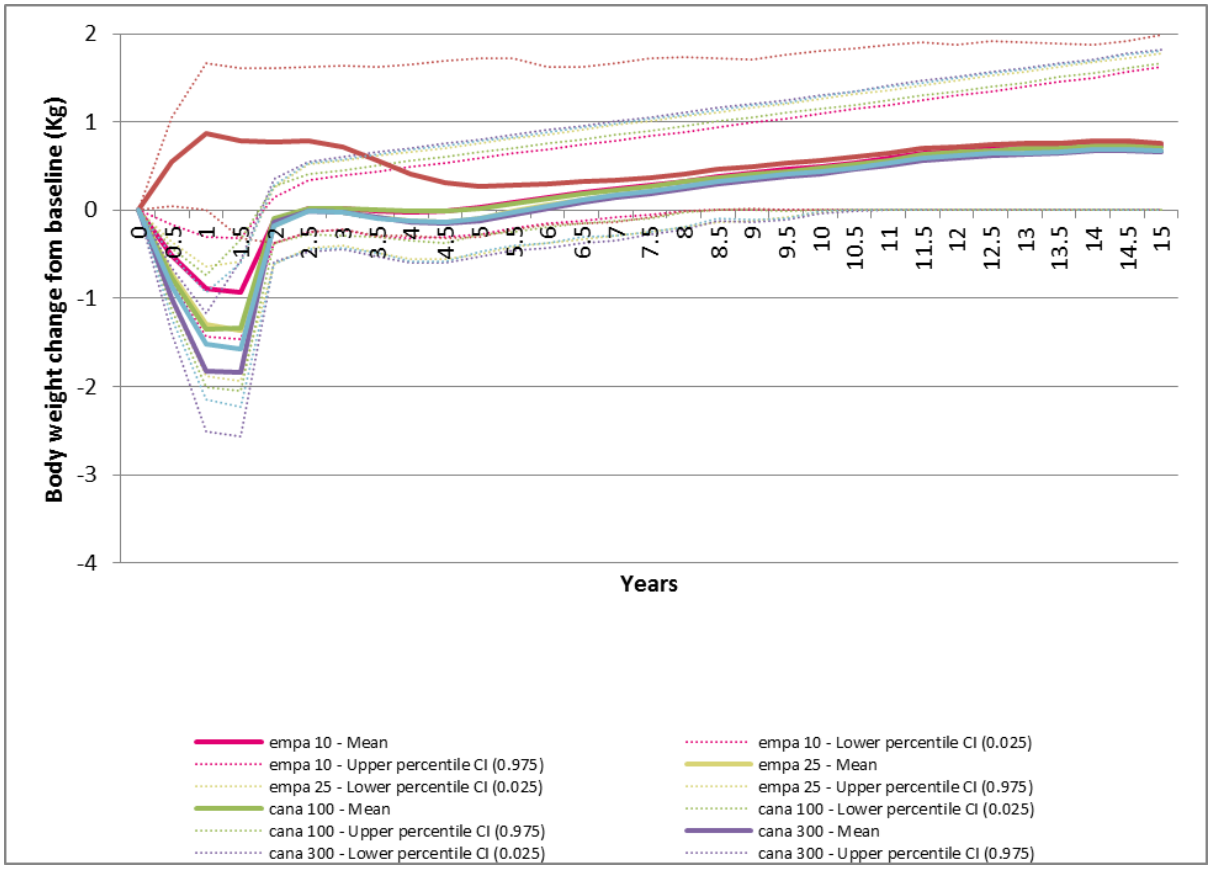


Figure 33. Insulin combinations – Weight over time

5.10 Health related quality of life

5.10.1 Baseline HRQoL

The EQ-5D data from the clinical trials demonstrated no difference between the treatment arms or the placebo/comparator treatment. However, the EQ-5D scores from the trial were not used for the economic evaluation. The EQ-5D values were derived from the UKPDS study.[35] The manufacturer states that the EQ-5D value from the UKPDS study was chosen because it is a well validated study and patients were relevant to a UK setting. The ERG believes this is reasonable.

5.10.2 HRQoL impact on adverse events

The adverse events considered in the model were hypoglycaemia, UTIs and genital tract infections. In the model hypoglycaemia is assumed to occur in all cycles as long as patient continues to receive the treatment that is causing the event; and UTIs and genital infections are assumed to occur only in the first cycle after the patient initiates the treatment associated with the AE. The values for these adverse events were obtained from published sources.

Table 34. HRQoL impacts of the complications of diabetes

State	Utility value	Reference
Diabetes without complications	0.77	UKPDS 62[35]
IHD	-0.09	Dapa ERG (UKPDS 62)
MI	-0.055	Dapa ERG (UKPDS 62)
CHF	-0.108	Dapa ERG (UKPDS 62)
Stroke	-0.164	Dapa ERG (UKPDS 62)
Amputation	-0.28	Dapa ERG (UKPDS 62)
Blindness in one eye	-0.074	Dapa ERG (UKPDS 62)
Renal failure	-0.0963	Sullivan et al 2011[38]
Hypoglycaemic events (non-severe)	-0.0035	NICE clinical guideline
Hypoglycaemic events (severe)	-0.01	NICE clinical guideline
UTIs	-0.00282	Dapa ERG (UKPDS 62)
GTIs	-0.00282	Dapa ERG (UKPDS 62)
Weight gain (per kg)	-0.0159	Baugust and Beale (2005)[36]

The model applied disutility as disutilities per event for adverse events related to hypoglycaemia (severe and non-severe), UTIs and GTIs; and disutility was also applied per year for patients with long term complications.

The model also assumed that weight loss did not have an impact on utilities but attached a disutility whenever the weight gain was higher than the baseline weight of the patient as a result of a treatment related weight gain.

5.11 Resources and costs

5.11.1 Direct treatment costs

Treatment costs for the drugs were obtained from the BNF.

Table 35. Medication costs reproduced from the table 74 of the MS

Drug	Strength (mg) or average daily IU	Pills per pack	Price per pack (£)	Daily dose (mg)	Pills per day	Net cost per day (£)	Net cost per year (£)
Metformin	500 mg	84	£0.81	1270 mg	1	£0.02	£8.95
SU	80 mg	60	£1.61	160 mg	2	£0.05	£19.60
Empagliflozin	10mg	28	£36.59	10mg	1	£1.31	£477.30
Empagliflozin	25mg	28	£36.59	25mg	1	£1.31	£477.30
Dapagliflozin	10mg	28	£36.59	10mg	1	£1.31	£477.30
Canagliflozin	100mg	30	£39.20	100mg	1	£1.31	£477.30
Canagliflozin	300mg	30	£49.99	300mg	1	£1.66	£608.63
Sitagliptin	100mg	28	£33.26	100mg	1	£1.19	£433.85
Insulin NPH	31.7 IU	N/A	N/A	N/A	N/A	£0.44	£396.20
Insulin glargine	31.7 IU	N/A	N/A	N/A	N/A	£0.86	£557.55

SU-Gliclazide; Insulin NPH- Insulatard

The yearly cost of insulin per patient included insulin pen costs £31, needles £5 and test strips £180 per year.

5.11.2 Costs of diabetic complications

The unit costs for each adverse events were derived from the UKPDS 65 study.[39] The model applies the UKPDS 65 estimated annual mean in-patient costs during the year of incidence and subsequent years. The costs were inflated to the 2013 prices (approximately increased by 60%) by using the hospital and community health services (HCHS) pay and price inflation.

A renal failure cost of £34,488 is drawn from the NICE costing report for chronic kidney disease. The cost of renal failure represented the annual cost of haemodialysis.

The submission states (page 294, section 7.5.7) that weight does not affect costs, but one might expect some increase in daily insulin dose with weight gain.

Table 36. Adverse event costs in first year reproduced from MS.

Event	Costs in British Pounds (£)	Lower bound	Upper bound	Reference
IHD	£7,628.14	£5,987.13	£9,717.86	[39]
MI	£8,179.41	£7,189.04	£9,304.40	Clarke, 2003 #47}
CHF	£6,773.98	£5,498.35	£8,346.08	[39]
Stroke	£10,932.6	£8,421.40	£14,195.39	[39]
Amputation (one leg)	£13,556.0	£8,365.31	£21,964.55	[39]
Blindness (one eye)	£7,003.14	£5,158.61	£9,507.93	[39]
Renal failure	£34,488.31			NICE CG for diabetes
Fatal MI	£2,511.20	£2,049.66	£3,075.29	[39]
Fatal CHF	£2,511.20	£2,049.66	£3,075.29	Assumption based on [39]
Fatal stroke	£6,774	£3,932.66	£11,666.56	[39]
Fatal amputation	£6,280	£3,875.65	£10,176.18	Assumption based on [39]
Fatal renal failure	£34,488			Assumption based on NICE CG for diabetes
Other mortality related to diabetes	£2,511	£2,049.66	£3,075.29	Assumption based on [39]

The cost of renal failure represented the annual cost of haemodialysis, which has been reported to be the main mode of renal replacement therapy in the UK. The cost of fatal renal failure was assumed to be the same as the annual cost of renal failure, as it was assumed in the economic model developed in the NICE clinical guidelines for diabetes.

The cost of fatal CHF was assumed to be the same as that of fatal MI. The cost of fatal amputation was assumed to be 46.33% of that of non-fatal amputation and non-fatal renal failure, respectively. This proportion was estimated as the average between the proportion of the fatal cost over the non-fatal cost for MIs (30.70%) and strokes (61.96%). The cost of other mortality related to diabetes was assumed to be similar to the cheapest mortality cost identified from Clarke 2003 (i.e. that of MI). CHF = congestive heart failure; IHD = ischaemic heart disease; MI = myocardial infarction

5.11.3 Cost of treatment related adverse events

A cost per severe event of hypoglycaemia of £335 is drawn from NICE CG87, updated to 2013 prices using HCHS pay and price inflation, resulted in £360 per event. UTIs and GTIs are assumed to require one GP consultation at a cost of £36 as drawn from dapagliflozin ERG submission.

5.12 Cost effectiveness results

The following results report the total modelled events, QALYs and cost over the 40 year time horizon of the modelling.

5.12.1 Cost effectiveness: Empagliflozin as an add – on to metformin vs. metformin plus sitagliptin

For the comparison with metformin plus sitagliptin, the events rates in the empagliflozin plus MET arm and the net impact compared to the metformin plus sitagliptin arm are as below.

Table 37. Clinical events for Empagliflozin as dual therapy (from ECCM model)

Events	Met+ Empag 10mg	Met+ Empag 25mg	Met+ Sita	Difference Met + Empa 10 mg	Difference Met+ Empa 25 mg
Long term complications: Mean total number of events per patient					
IHD	0.090	0.089	0.091	-0.001	-0.001
MIs	0.222	0.220	0.224	-0.001	-0.004
CHFs	0.076	0.075	0.077	-0.001	-0.002
Strokes	0.174	0.177	0.174	-	0.003
Amputations	0.082	0.081	0.081	0.001	-
Blindness in one eye	0.041	0.041	0.040	0.001	0.001
Renal failure	0.019	0.019	0.019	0	-
Years free of complications	7.484	7.507	7.497	-0.013	0.009
Adverse events: Mean total number of events per patient over what time period					
Hypoglycaemic events (non-severe)	28.2	27.10	27.424	0.775	-0.296
Hypoglycaemic events (severe)	0.178	0.176	0.160	0.017	0.016
UTIs	0.448	0.348	0.361	0.087	-0.013
GTIs	0.084	0.141	0.005	0.078	0.136

The net impacts of most events are relatively minor. Note that the reporting of GTIs events implies that compared to sitagliptin, GTI events are 16 times for Met+ Empa 10 mg and 28 times higher for Met + Empa 25mg per patient over the period of modelling.

The net impact of most events of other dual therapy comparisons (dapagliflozin and canagliflozin) were relatively minor. The only difference was across the reporting of GTI/UTIs per patient over the period of modelling (see Table 38 reproduced from MS table 78).

Table 38. Clinical events for Empagliflozin as dual therapy for other dual therapy comparison (reproduced from MS table 78)

	Dual therapy - Metformin				
	Empagliflozin 10mg	Empagliflozin 25mg	Canagliflozin 100mg	Canagliflozin 300mg	Dapagliflozin 10mg
QALY	7.424	7.433	7.415	7.418	7.404
Life Years	9.952	9.958	9.937	9.941	9.919
LTCs: Total number of events	0.704	0.702	0.710	0.705	0.713
IHDs	0.090	0.089	0.090	0.090	0.090
MIs	0.222	0.220	0.223	0.221	0.225
CHFs	0.076	0.075	0.076	0.075	0.076
Strokes	0.174	0.177	0.178	0.179	0.179
Amputations	0.082	0.081	0.082	0.081	0.083
Blindness in one eye	0.041	0.041	0.040	0.040	0.041
Renal failures	0.019	0.019	0.019	0.019	0.019
Years free of complication	7.484	7.507	7.480	7.510	7.457
AEs: Total number of events	28.91	27.77	29.72	31.25	26.47
Hypoglycaemic events (non-severe)	28.20	27.10	29.25	30.76	25.74
Hypoglycaemic events (severe)	0.178	0.176	0.175	0.178	0.184
UTIs	0.448	0.348	0.179	0.189	0.365
GTIs	0.084	0.141	0.112	0.117	0.179
Total deaths	1.000	1.000	1.000	1.000	1.000
Diabetes-related deaths	0.270	0.270	0.273	0.272	0.275
Other deaths	0.730	0.730	0.727	0.728	0.725

5.12.2 Empagliflozin as an add – on to metformin vs. Metformin plus Sitagliptin: QALYS

Table 39. QALYs for Empagliflozin as dual therapy

QALYS lost due to AEs	Met+ Empag 10mg	Met+ Empag 25 mg	Met+ Sita	Difference from sita Met+Empag 10 mg	Difference from sita Met+ Empag 25mg
Hypoglycaemic events (non-severe)	-0.0758	-0.0704	-0.0693	0.0065	0.0011
Hypoglycaemic events (severe)	-0.0015	-0.0015	-0.0014	0.0001	0.0001
UTIs	-0.0104	-0.0080	-0.0142	-0.0038	-0.0062
GTIs	-0.0020	-0.0033	-0.0001	0.0019	0.0032
Weight gain	-0.0145	-0.0144	-0.0142	0.0003	0.0002
Total QALYs	7.42	7.43	7.42	-	0.01

Please note the net QALY decrement will not in general equal the overall net QALYS, as the net QALYS incorporate survival effects. Note that there is no impact of weight gain on QALY in the Met+ Sitagliptin arm over the period of the modelling.

5.12.3 Empagliflozin as an add – on to metformin vs. Metformin plus Sitagliptin: Costs

Table 40. Costs for Empagliflozin as dual therapy

Events	Met+ Empag 10mg (£)	Met+ Empag 25mg (£)	Met+ Sita (£)	Difference Met+ Empag 10 mg(£)	Difference Met+ Empag 25 mg (£)
Diabetes without complications	4,725	4,728	4,720	5	8
Long term complications: Mean total cost per patient					
IHD	4,177	4,145	4,173	4	-28
MIs	4,691	4,669	4,722	-31	-53
CHFs	1,877	1,846	1,874	3	-28
Strokes	3,934	3,948	3,739	195	209
Amputations	2,289	2,254	2,226	63	28
Blindness in one eye	1,462	1,458	1,438	24	20
Renal failure	2,248	2,204	2,205	43	-1
Subtotal	**20,677**	**20,524**	**20,376**	**-0.013**	**148**
Adverse events: Mean total cost per patient					
Hypoglycaemic events (severe)	11	11	10	1	1
UTIs	15	12	12	3	-
GTIs	3	5	0	3	5
Subtotal	**29**	**28**	**23**	**7**	**6**
Treatment costs: Mean total cost per patient					
First treatment	1,592	1,605	1,712	-120	-107
Subsequent treatments	5,203	5,044	4,789	415	255
Subtotal	**6,796**	**6,649**	**6,501**	**295**	**148**
Deaths (mortality related to diabetes)	427	424	426	1	-2
Total NHS and PSS perspective	**32,654**	**32,353**	**32,046**	**603**	**307**

Note that the average time spent on the dual therapy before the switch to insulin is not reported.

In terms of costs and QALYs there does not appear to be any significant difference between other dual therapy comparisons (dapagliflozin and canagliflozin) (see Table 41 reproduced from MS table 82).

Table 41. Disaggregated QALYs and costs – metformin background [Total cost and QALYs for other dual therapy comparisons (reproduced from MS table 82)]

		Empagliflozin 10 mg	Empagliflozin 25 mg	Canagliflozin 100 mg	Canagliflozin 300 mg	Dapagliflozin 10 mg
Deaths	**Diabetes-related deaths**	0.27	0.27	0.27	0.27	0.28
	Other deaths	0.73	0.73	0.73	0.73	0.73
Life Years	**Total LYs**	9.95	9.96	9.94	9.94	9.92
QALYs	**Lost due to LTCs**					
	IHD	-0.07	-0.06	-0.06	-0.06	-0.06
	MI	-0.05	-0.05	-0.05	-0.05	-0.05
	CHF	-0.03	-0.03	-0.03	-0.03	-0.03
	Strokes	-0.13	-0.13	-0.13	-0.13	-0.13
	Amputation	-0.10	-0.10	-0.10	-0.09	-0.10
	Blindness in one eye	-0.02	-0.02	-0.02	-0.02	-0.02
	Renal failure	-0.01	-0.01	-0.01	-0.01	-0.01
	Subtotal	-0.40	-0.40	-0.40	-0.39	-0.40
	Lost due to AEs					
	Hypoglycaemic events (non-severe)	-0.08	-0.07	-0.08	-0.08	-0.07
	Hypoglycaemic events (severe)	0.00	0.00	0.00	0.00	0.00
	UTIs	0.00	0.00	0.00	0.00	0.00
	GenIs	0.00	0.00	0.00	0.00	0.00
	Weight gain	-0.01	-0.01	-0.01	-0.01	-0.02
	Subtotal	-0.09	-0.09	-0.09	-0.10	-0.08
	Lost QALYs due to deaths					
	Diabetes-related deaths	0.00	0.00	0.00	0.00	0.00
	Other deaths	0.00	0.00	0.00	0.00	0.00
	Subtotal	0.00	0.00	0.00	0.00	0.00
	Total QALYs	7.42	7.43	7.41	7.42	7.40
Costs						
	Diabetes w/o complications	£4,725.00	£4,728.00	£4,718.00	£4,720.00	£4,709.00
	LTCs					
	IHD	£4,177.00	£4,145.00	£4,160.00	£4,136.00	£4,159.00
	MI	£4,691.00	£4,669.00	£4,670.00	£4,619.00	£4,685.00
	CHF	£1,877.00	£1,846.00	£1,874.00	£1,829.00	£1,867.00
	Strokes	£3,934.00	£3,948.00	£3,938.00	£3,913.00	£3,942.00
	Amputation	£2,289.00	£2,254.00	£2,226.00	£2,165.00	£2,262.00
	Blindness in one eye	£1,462.00	£1,458.00	£1,471.00	£1,462.00	£1,471.00
	Renal failure	£2,248.00	£2,204.00	£2,277.00	£2,239.00	£2,267.00
	Subtotal	£20,677.00	£20,524.00	£20,616.00	£20,363.00	£20,652.00
	AEs					
	Hypoglycaemic events (non-severe)	£0.00	£0.00	£0.00	£0.00	£0.00
	Hypoglycaemic events (severe)	£11.00	£11.00	£11.00	£11.00	£12.00
	UTIs	£15.00	£12.00	£6.00	£6.00	£13.00

GenIs	£3.00	£5.00	£4.00	£4.00	£6.00
Subtotal	£29.00	£27.00	£21.00	£21.00	£30.00
Treatment costs					
First treatment	£1,592.00	£1,605.00	£1,510.00	£1,909.00	£1,329.00
Subsequent treatments	£5,203.00	£5,044.00	£5,010.00	£5,296.00	£5,021.00
Subtotal	£6,796.00	£6,649.00	£6,520.00	£7,205.00	£6,350.00
Deaths	£427.00	£424.00	£426.00	£428.00	£437.00
Total NHS and PSS costs	£32,654.00	£32,353.00	£32,300.00	£32,737.00	£32,179.00

LTC = long term conditions; QALY = quality adjusted life years; CHF = congestive heart failure; IHD = ischaemic heart disease; MI = myocardial infarction; UTIs = urinary tract infection; GTIs = genital infection

5.12.4 Incremental cost-effectiveness results in dual therapy:

Table 42. Cost – effectiveness results for Empagliflozin as dual therapy

	Met+ Empa 10mg	**Met+ Sita**	**Incremental**
Cost (£)	32,654	32,046	608
QALYs	7.42	7.42	-
ICER			Dominated by sitagliptin
	Met+Empa 25mg	**Met+ Sit**	**Incremental**
Cost (£)	32,353	32,046	307
QALYs	7.433	7.425	0.008
ICER			38,375

The above cost effectiveness results are probabilistic results and not deterministic. When the model was run deterministically, the results obtained were different to probabilistic analysis. The probability of empagliflozin 10 mg being cost effective for a willingness to pay of £20,000 per QALY is estimated to be 20% and 23% for empagliflozin 25 mg. Note the totals have been drawn directly from the model and may not exactly equal the totals from the written submission.

The table below (reproduced from Ms-table 88) compares different treatment comparisons and shows the sitagliptin 100 mg is the most cost effective option.

Table 43. Base-case results dual therapy – combined results (reproduced from MS table 88)

	QALY	Δ QALYs	Total costs NHS perspective	Δ Costs	ICER
Met + Sitagliptin 100 mg	7.425	-	£32,046	-	-
Met + Dapagliflozin 10 mg	7.404	-0.021	£32,179	£133	Dominated by sitagliptin 100 mg
Met + Canagliflozin 100 mg	7.415	-0.01	£32,300	£255	Dominated by sitagliptin 100 mg
Met + Empagliflozin 25 mg	7.433	0.008	£32,353	£307	£38,425
Met + Empagliflozin 10 mg	7.424	-0.009	£32,654	£301	Dominated by empagliflozin 25 mg
Met + Canagliflozin 300 mg	7.418	-0.015	£32,737	£384	Dominated by empagliflozin 25 mg

ICER = incremental costs-effectiveness ratio; Met = metformin; NHS = National Health Service QALY = quality-adjusted life year

5.12.5 Empagliflozin in triple therapy: Events

Metformin and SU combinations: metformin plus SU plus empagliflozin vs. metformin plus SU plus sitagliptin

Table 44. Clinical events for Empagliflozin as triple therapy

Events	Met+ SU+Empag 10mg	Met+ SU+Empag 25mg	Met+ SU+Sita	Difference Met + SU+Empa 10 mg	Difference Met+ SU+Empa 25 mg
Long term complications: Mean total number of events per patient					
IHD	0.093	0.094	0.096	-0.003	-0.002
MIs	0.211	0.212	0.216	-0.005	-0.004
CHFs	0.094	0.094	0.097	-0.003	-0.003
Strokes	0.179	0.183	0.180	-0.001	-0.003
Amputations	0.080	0.080	0.080	-	-
Blindness in one eye	0.046	0.047	0.046	-	-0.001
Renal failure	0.019	0.019	0.019	-	-
Years free of complications	6.978	6.957	6.926	0.052	0.028
Adverse events: Mean total number of events per patient					
Hypoglycaemic events (non-severe)	22.532	25.217	26.941	-4.409	-1.724
Hypoglycaemic events (severe)	0.163	0.167	0.147	0.016	0.02
UTIs	0.272	0.248	0.132	0.14	0.116
GTIs	0.063	0.117	0.013	0.05	0.104

As in dual therapy analyses, the net impact on events are relatively minor.

The net impact of most events of canagliflozin as triple therapy comparisons was relatively minor. The only difference was across the reporting of stroke, non-severe hypoglycaemic events and GTI/UTIs per patient over the period of modelling (see Table 45 reproduced from MS table 78). The events for dapagliflozin is not reported as the events were assumed to be similar to canagliflozin 300mg.

Table 45. Breakdown of clinical events for analysis of triple therapy - metformin plus SU [Clinical events for Empagliflozin as triple therapy for canagliflozin comparison (reproduced from MS table 78)]

	Triple therapy - Metformin plus Sulphonylureas			
	Empagliflozin 10mg	Empagliflozin 25mg	Canagliflozin 100mg	Canagliflozin 300mg
QALY	6.991	6.978	6.980	6.976
Life Years	9.369	9.360	9.362	9.364
LTCs: Total number of events	0.722	0.728	0.721	0.725
IHDs	0.093	0.094	0.094	0.094
MIs	0.211	0.212	0.211	0.211
CHFs	0.094	0.094	0.094	0.094
Strokes	0.179	0.183	0.178	0.183
Amputations	0.080	0.080	0.078	0.078
Blindness in one eye	0.046	0.047	0.046	0.045
Renal failures	0.019	0.019	0.019	0.019
Years free of complication	6.978	6.957	6.982	6.983
AEs: Total number of events	23.030	25.748	28.719	31.738
Hypoglycaemic events (non-severe)	22.532	25.217	28.2823	31.294
Hypoglycaemic events (severe)	0.163	0.167	0.155	0.152
UTIs	0.272	0.248	0.169	0.172
GTIs	0.063	0.117	0.112	0.120
Total deaths	1.000	1.000	1.000	1.000
Diabetes-related deaths	0.268	0.269	0.269	0.267
Other deaths	0.732	0.731	0.731	0.733

QALY = quality adjusted life years; CHF = congestive heart failure; IHD = ischaemic heart disease; MI = myocardial infarction; UTIs = urinary tract infection; GTIs = genital infection

5.12.6 Empagliflozin in triple therapy: QALYs

Table 46. QALYs for Empagliflozin as triple therapy

QALYS lost due to AEs	Met+ SU+ Empag 10mg	Met+ SU+Empag 25 mg	Met+ SU+Sita	Difference Met+SU+Empag 10 mg	Difference Met+ SU+Empag 25mg
Hypoglycaemic events (non-severe)	-0.0591	-0.0660	-0.0704	-0.0113	-0.0044
Hypoglycaemic events (severe)	-0.0014	-0.0014	-0.0013	0.0001	-0.0001
UTIs	-0.0061	-0.0056	-0.0029	0.0032	-0.0027
GTIs	-0.0015	-0.0027	-0.0003	0.0012	-0.0024
Weight gain	-0.0133	-0.0136	-0.0233	-0.01	-0.0097
Total QALYs	6.99	6.97	6.96	0.03	0.01

The difference in QALY is very minimal between comparisons.

5.12.7 Empagliflozin in triple therapy: Costs

Table 47. Costs for Empagliflozin as triple therapy

Events	Met+ SU+Empag 10mg (£)	Met+ SU+Empag 25mg (£)	Met+ SU+Sita (£)	Difference Met+ SU+Empag 10 mg(£)	Difference Met+ SU+Empag 25 mg (£)
Diabetes without complications	4,441	4,437	4,428	4	9
Long term complications: Mean total cost per patient					
IHD	4,161	4,174	4,215	-54	-41
MIs	4,344	4,352	4,418	-74	-66
CHFs	2,371	2,374	2,455	-84	-81
Strokes	3,665	3,784	3,692	-27	92
Amputations	2,124	2,139	2,124	-	15
Blindness in one eye	1,455	1,473	1,433	22	40
Renal failure	1,958	1,971	1,978	-20	-7
Subtotal	**20,078**	**20,268**	**20,314**	**-236**	**-46**
Adverse events: Mean total cost per patient					
Hypoglycaemic events (severe)	10	10	9	1	1
UTIs	9	8	4	5	4
GTIs	2	4	0	2	4
Subtotal	**21**	**22**	**14**	**7**	**8**
Treatment costs: Mean total cost per patient					
First treatment	1,696	1,611	1,801	-105	-190
Subsequent treatments	4,735	4,778	4,369	366	409
Subtotal	**6,430**	**6,390**	**6,170**	**260**	**220**
Deaths (mortality related to diabetes)	438	440	438	-	2
Total NHS and PSS perspective	**32,654**	**31,557**	**31,365**	**1289**	**192**

In terms of costs and QALYs there does not appear to be any significant difference between triple therapy comparison (canagliflozin) (see Table 48 reproduced from MS table 83).

Table 48. Disaggregated QALYs and costs – metformin and SU

		Empagliflozin 10 mg	Empagliflozin 25 mg	Canagliflozin 100 mg	Canagliflozin 300 mg
Deaths	**Diabetes-related deaths**	0.268	0.269	0.269	0.267
	Other deaths	0.732	0.731	0.732	0.733
Life Years	**Total LYs**	9.370	9.360	9.360	9.360
QALYs	**Lost due to LTCs**				
	IHD	-0.065	-0.065	-0.065	-0.065
	MI	-0.045	-0.045	-0.044	-0.044
	CHF	-0.042	-0.042	-0.042	-0.041
	Strokes	-0.118	-0.122	-0.117	-0.122
	Amputation	-0.091	-0.091	-0.089	-0.087
	Blindness in one eye	-0.023	-0.024	-0.023	-0.023
	Renal failure	-0.006	-0.006	-0.006	-0.006
	Subtotal	-0.389	-0.394	-0.386	-0.389
	Lost due to AEs				
	Hypoglycaemic events (non-severe)	-0.059	-0.066	-0.074	-0.082
	Hypoglycaemic events (severe)	-0.001	-0.001	-0.001	-0.001
	UTIs	-0.001	-0.001	0.000	0.000
	GenIs	0.000	0.000	0.000	0.000
	Weight gain	-0.013	-0.014	-0.013	-0.013
	Subtotal	-0.075	-0.082	-0.089	-0.097
	Lost QALYs due to deaths				
	Diabetes-related deaths	0.000	0.000	0.000	0.000
	Other deaths	0.000	0.000	0.000	0.000
	Subtotal	0.000	0.000	0.000	0.000
	Total QALYs	6.990	6.980	6.980	6.980
Costs					
	Diabetes w/o complications	£4,441	£4,437	£4,438	£4,439
	LTCs				
	IHD	£4,161	£4,174	£4,184	£4,180
	MI	£4,344	£4,352	£4,324	£4,268
	CHF	£2,371	£2,374	£2,349	£2,336
	Strokes	£3,665	£3,784	£3,623	£3,794
	Amputation	£2,124	£2,139	£2,091	£2,043
	Blindness in one eye	£1,455	£1,473	£1,443	£1,439
	Renal failure	£1,958	£1,971	£1,962	£1,996
	Subtotal	£20,078	£20,268	£19,976	£20,057
	AEs				
	Hypoglycaemic events (non-severe)	£0	£0	£0	£0

Hypoglycaemic events (severe)	£10	£10	£10	£9
UTIs	£9	£8	£6	£6
GenIs	£2	£4	£4	£4
Subtotal	£21	£22	£19	£19
Treatment costs				
First treatment	£1,696	£1,611	£1,757	£2,321
Subsequent treatments	£4,735	£4,778	£4,592	£4,824
Subtotal	£6,430	£6,390	£6,349	£7,145
Deaths	£438	£440	£434	£428
Total NHS and PSS costs	£31,408	£31,557	£31,217	£32,087

LTC = long term conditions; QALY = quality adjusted life years; CHF = congestive heart failure; IHD = ischaemic heart disease; MI = myocardial infarction; UTIs = urinary tract infection; GTIs = genital infection

5.12.8 Incremental Cost effectiveness results in triple therapy

Table 49. Cost – effectiveness results for Empagliflozin as triple therapy

	Met+ SU+Empa 10mg	**Met+ SU+Sita**	**Incremental**
Cost (£)	31,408	31,365	43
QALYs	6.99	6.96	0.03
ICER			1,433
	Met+SU+Empa 25mg	**Met+ SU+Sit**	**Incremental**
Cost (£)	31,557	31,365	192
QALYs	6.97	6.96	0.01
ICER			19,200

Note the totals have been drawn directly from the model and may not exactly equal the totals from the written submission. However, it should be noted that due to small difference in QALYs means caution must be applied in interpreting the ICER. The ICER is highly variable for every third decimal place difference in mean QALYS.

The probability of Met + SU + empagliflozin 10 mg being cost effective for a willingness to pay of £20,000 per QALY is estimated to be 34% and 32% for empagliflozin 25 mg.

The table below (reproduced from MS-table 92) compares different treatment comparisons and shows that Met + SU + canagliflozin 100 mg is the least costly Met + SU + empagliflozin 10 mg is the most effective option.

Table 50. Base-case results triple therapy (met + SU combinations) reproduced from MS table – 92

	QALY	Δ QALYs	Total costs NHS perspective	Δ Costs	ICER
Met+SU+ Canagliflozin 100 mg	6.98	-	£31,217	-	-
Met+SU+ Sitagliptin 100 mg	6.959	-0.021	£31,365	£148	Dominated by canagliflozin 100 mg
Met+SU+ Empagliflozin 10 mg	6.991	0.011	£31,409	£192	£17,445
Met+SU+ Empagliflozin 25 mg	6.978	-0.013	£31,557	£148	Dominated by empagliflozin 10 mg
Met+SU+ Canagliflozin 300 mg	6.976	-0.015	£32,087	£679	Dominated by empagliflozin 10 mg

ICER = incremental costs-effectiveness ratio; Met = metformin; NHS = National Health Service QALY = quality-adjusted life year; SU = sulphonylurea

5.12.9 Cost effectiveness: Empagliflozin as an add-on to insulin

For the comparison of add-on to insulin the events rates, cost and QALYS in the empagliflozin plus insulin arm and the net impact compared to the sitagliptin plus insulin arm are as below.

Table 51. Clinical events for Empagliflozin as an add-on to insulin

Events	Insulin+Empag 10mg	Insulin+Empag 25mg	Insulin+Sita	Difference Insulin+Empa 10 mg	Difference Insulin+Empa 25 mg
Long term complications: Mean total number of events per patient					
IHD	0.098	0.098	0.099	-0.001	-0.001
MIs	0.219	0.218	0.222	-0.003	-0.004
CHFs	0.095	0.095	0.098	-0.003	-0.003
Strokes	0.189	0.189	0.192	-0.003	-0.003
Amputations	0.081	0.081	0.082	-0.001	-0.001
Blindness in one eye	0.049	0.049	0.048	0.001	0.001
Renal failure	0.019	0.019	0.019	-	-
Years free of complications	6.876	6.881	6.836	0.04	0.045
Adverse events: Mean total number of events per patient					
Hypoglycaemic events (non-severe)	28.189	29.491	31.519	-3.33	-2.028
Hypoglycaemic events (severe)	0.231	0.231	0.229	0.002	0.002
UTIs	0.430	0.406	0.333	0.097	0.073
GTIs	0.064	0.150	0.020	0.044	0.13

As before, the net impacts on most events are relatively minor.

The net impact of most events of dapagliflozin and canagliflozin plus insulin combination were relatively minor (see Table 52 reproduced from MS table 81).

Table 52. Breakdown of clinical events for analysis of insulin combinations (reproduced from MS table 81).

	Insulin combinations				
	Empagliflozin 10mg	Empagliflozin 25mg	Canagliflozin 100mg	Canagliflozin 300mg	Dapagliflozin 10mg
QALY	6.947	6.948	6.944	6.957	6.953
Life Years	9.325	9.329	9.325	9.344	9.335
LTCs: Total number of events	0.749	0.748	0.749	0.743	0.748
IHDs	0.098	0.098	0.098	0.098	0.098
MIs	0.219	0.218	0.218	0.217	0.218
CHFs	0.095	0.095	0.094	0.094	0.094
Strokes	0.189	0.189	0.191	0.189	0.190
Amputations	0.081	0.081	0.080	0.080	0.081
Blindness in one eye	0.049	0.049	0.049	0.048	0.049
Renal failures	0.019	0.019	0.019	0.019	0.019
Years free of complication	6.876	6.881	6.890	6.919	6.892
AEs: Total number of events	28.914	30.278	31.583	33.419	29.924
Hypoglycaemic events (non-severe)	28.189	29.491	30.950	32.784	29.296
Hypoglycaemic events (severe)	0.231	0.231	0.232	0.231	0.231
UTIs	0.430	0.406	0.315	0.300	0.294
GTIs	0.064	0.150	0.086	0.103	0.104
Total deaths	1.000	1.000	1.000	1.000	1.000
Diabetes-related deaths	0.278	0.278	0.278	0.275	0.277
Other deaths	0.722	0.722	0.722	0.725	0.723

QALY = quality adjusted life years; CHF = congestive heart failure; IHD = ischaemic heart disease; MI = myocardial infarction; UTIs = urinary tract infection; GTIs = genital infection

5.12.10 Empagliflozin as an add-on to insulin: QALYs

Table 53. QALYs for Empagliflozin as an add-on to insulin

QALYS lost due to AEs	Insulin+ Empag 10mg	Insulin+Empag 25 mg	Insulin+Sita	Difference Insulin+Empag 10 mg	Difference Insulin+Empag 25mg
Hypoglycaemic events (non-severe)	-0.0777	-0.0819	-0.0886	0.0109	0.0067
Hypoglycaemic events (severe)	-0.0021	-0.0021	-0.0021	-	-
UTIs	-0.0099	-0.0093	-0.0075	-0.0024	-0.0018
GTIs	-0.0015	-0.0037	-0.0005	-0.001	-0.0032
Weight gain	0	0	-0.0186	0.0186	0.0186
Total QALYs	6.94	6.94	6.91	0.03	0.03

The direct HRQoL effects of weight gain changes contribute to the minimal anticipated gains from empagliflozin.

5.12.11Empagliflozin as an add-on to insulin: costs

Table 54. Costs for Empagliflozin as an add-on to insulin

Events	Insulin+Empag 10mg (£)	Insulin+Empag 25mg (£)	Insulin+Sita (£)	Difference Insulin+Empag 10 mg(£)	Difference Insulin+Empag 25 mg (£)
Diabetes without complications	4,420	4,422	4,408	12	14
Long term complications: Mean total cost per patient					
IHD	4,280	4,276	4,285	-5	-9
MIs	4,402	4,396	4,405	-3	-9
CHFs	2,435	2,426	2,515	-80	-89
Strokes	3,909	3,913	3,963	-54	-50
Amputations	2,179	2,189	2,208	-29	-19
Blindness in one eye	1,504	1,500	1,492	12	8
Renal failure	1,981	1,981	1,983	-2	-2
Subtotal	**20,688**	**20,681**	**20,851**	**-163**	**-170**
Adverse events: Mean total cost per patient					
Hypoglycaemic events (severe)	16	16	16	0	0
UTIs	14	14	11	3	3
GTIs	2	5	1	1	4
Subtotal	**32**	**35**	**27**	**5**	**8**
Treatment costs: Mean total cost per patient					
First treatment	2,467	2,636	2,538	-71	98
Subsequent treatments	2,507	2,432	2,405	102	27
Subtotal	**4,974**	**5,067**	**4,942**	**32**	**125**
Deaths (mortality related to diabetes)	450	448	453	-3	-5
Total NHS and PSS perspective	**30,564**	**30,653**	**30,682**	**-118**	**-29**

In terms of costs and QALYs there does not appear to be any significant difference between canagliflozin and dapagliflozin as –an add on to insulin (see Table 55 reproduced from MS table 85).

Table 55. Disaggregated QALYs and costs – Insulin

		Empagliflozin 10 mg	Empagliflozin 25 mg	Canagliflozin 100 mg	Canagliflozin 300 mg	Dapagliflozin 10 mg	Sitagliptin 100 mg
Deaths	**Diabetes-related deaths**	0.278	0.278	0.278	0.275	0.277	0.282
	Other deaths	0.722	0.722	0.722	0.725	0.723	0.718
Life Years	**Total LYs**	9.330	9.330	9.330	9.340	9.330	9.300
QALYs	**Lost due to LTCs**						
	IHD	-0.067	-0.067	-0.066	-0.066	-0.066	-0.067
	MI	-0.045	-0.045	-0.045	-0.045	-0.045	-0.045
	CHF	-0.043	-0.043	-0.042	-0.042	-0.042	-0.044
	Strokes	-0.126	-0.126	-0.126	-0.125	-0.126	-0.127
	Amputation	-0.093	-0.093	-0.092	-0.090	-0.093	-0.094
	Blindness in one eye	-0.024	-0.024	-0.024	-0.024	-0.024	-0.024
	Renal failure	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006
	Subtotal	-0.404	-0.404	-0.401	-0.398	-0.403	-0.408
	Lost due to AEs						
	Hypoglycaemic events (non-severe)	-0.078	-0.082	-0.087	-0.093	-0.081	-0.089
	Hypoglycaemic events (severe)	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	UTIs	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	GenIs	0.000	0.000	0.000	0.000	0.000	0.000
	Weight gain	0.000	0.000	0.000	0.000	0.000	-0.019
	Subtotal	-0.081	-0.085	-0.090	-0.096	-0.084	-0.110
	Lost QALYs due to deaths						
	Diabetes-related deaths	0.000	0.000	0.000	0.000	0.000	0.000
	Other deaths	0.000	0.000	0.000	0.000	0.000	0.000
	Subtotal	0.000	0.000	0.000	0.000	0.000	0.000
	Total QALYs	6.950	6.950	6.940	6.960	6.950	6.910
Costs							
	Diabetes w/o complications	£4,420.00	£4,422.00	£4,420.00	£4,429.00	£4,425.00	£4,408.00
	LTCs						
	IHD	£4,280.00	£4,276.00	£4,261.00	£4,264.00	£4,268.00	£4,285.00
	MI	£4,402.00	£4,396.00	£4,365.00	£4,355.00	£4,387.00	£4,405.00
	CHF	£2,435.00	£2,426.00	£2,374.00	£2,346.00	£2,389.00	£2,515.00
	Strokes	£3,909.00	£3,913.00	£3,910.00	£3,882.00	£3,925.00	£3,963.00
	Amputation	£2,179.00	£2,189.00	£2,163.00	£2,122.00	£2,174.00	£2,208.00
	Blindness in one eye	£1,504.00	£1,500.00	£1,487.00	£1,482.00	£1,491.00	£1,492.00
	Renal failure	£1,981.00	£1,981.00	£1,993.00	£1,966.00	£1,975.00	£1,983.00
	Subtotal	£20,688.00	£20,681.00	£20,552.00	£20,417.00	£20,610.00	£20,851.00
	AEs						
	Hypoglycaemic events (non-severe)	£0.00	£0.00	£0.00	£0.00	£0.00	£0.00
	Hypoglycaemic events (severe)	£16.00	£16.00	£16.00	£16.00	£16.00	£16.00
	UTIs	£14.00	£14.00	£10.00	£10.00	£10.00	£11.00

	GenIs	£2.00	£5.00	£3.00	£4.00	£4.00	£1.00
	Subtotal	£32.00	£35.00	£29.00	£29.00	£29.00	£27.00
	Treatment costs						
	First treatment	£2,467.00	£2,636.00	£2,458.00	£3,121.00	£2,573.00	£2,538.00
	Subsequent treatments	£2,507.00	£2,432.00	£2,511.00	£2,402.00	£2,463.00	£2,405.00
	Subtotal	£4,974.00	£5,067.00	£4,969.00	£5,524.00	£5,035.00	£4,942.00
	Deaths	£450.00	£448.00	£448.00	£442.00	£446.00	£453.00
	Total NHS and PSS costs	£30,564.00	£30,653.00	£30,418.00	£30,842.00	£30,545.00	£30,682.00

LTC = long term conditions; QALY = quality adjusted life years; CHF = congestive heart failure; IHD = ischaemic heart disease; MI = myocardial infarction; UTIs = urinary tract infection; GTIs = genital infection

5.12.12Incremental Cost effectiveness results add-on to insulin

Table 56. Cost – effectiveness results for Empagliflozin as an add-on to insulin

	Insulin+Empa 10mg	**Insulin+Sita**	**Incremental**
Cost (£)	30,564	30,682	-118
QALYs	6.94	6.91	0.03
ICER			Dominated by Empagliflozin
	Insulin+Empa 25mg	**Insulin+Sit**	**Incremental**
Cost (£)	30,653	30,682	-29
QALYs	6.94	6.91	0.03
ICER			Dominated by Empagliflozin

The probability of Insulin+ empagliflozin 10 mg being cost effective for a willingness to pay of £20,000 per QALY is estimated to be 75% and 65% for empagliflozin 25 mg.

The table below (reproduced from Ms-table 100) shows that dapagliflozin is the most cost effective treatment.

Table 57. Base case results for insulin combinations (reproduced from MS – table 100)

	QALY	Δ QALYs	Total costs NHS perspective	Δ Costs	ICER
Insulin + Canagliflozin 100 mg	6.944	-	£30,418	-	-
Insulin + Dapagliflozin 10 mg	6.953	0.009	£30,545	£128	£14,178
Insulin + Empagliflozin 10 mg	6.947	-0.006	£30,564	£18	Dominated by dapagliflozin 10 mg
Insulin + Empagliflozin 25 mg	6.948	-0.005	£30,653	£108	Dominated by dapagliflozin 10 mg
Insulin + Sitagliptin 100 mg	6.912	-0.041	£30,682	£137	Dominated by dapagliflozin 10 mg
Insulin + Canagliflozin 300 mg	6.957	0.004	£30,842	£296	£74,075

ICER = incremental costs-effectiveness ratio; NHS = National Health Service QALY = quality-adjusted life year

5.13 Sensitivity analyses

A range of sensitivity analyses were undertaken:

- Rather than applying the 24 weeks data applying the 52 weeks data for metformin, metformin and SU and insulin.
- Disutilities values taken from the UK catalogue of utilities with and without applying age-related disutility.
- Assuming discontinuation rate similar to empagliflozin 10 mg.
- Varying the adverse events by applying the lowest number of adverse events observed across treatments.
- BMI changes did not impact the incidence of CHF or the disutility of patients due to weight changes.
- Varying the assumption around weight change assuming that the weight change would initially occur over the first cycle and the weight change would be maintained over one year.
- Assuming the duration of treatment effect would last for two years rather than one year.
- Varying the life time horizon from 40 years to 10 years.
- Assuming different discount rates (0% and 6%) on costs and effects.

All performed sensitivity analyses confirmed no significant difference between treatments and the difference in cost and QALYS were insignificant across comparison.

5.14 ERG cross check and critique

5.14.1 Patient group

The patient group included in the economic evaluation is adults with T2DM. The principal source used to describe the patients' initial characteristics comes from the UKPDS study.[37] It should be remembered that the median age at recruitment to UKPDS was 54, that the main results were based on 10 years of follow-up and that therefore the recruits were younger than average. People of 66 and

over were not recruited. The mean HbA1c of participants in the included trials were above the NICE current guidelines threshold of HbA1c >7.5%, indicating a need for further therapies.

5.14.2 Comparators

The economic modelling comparators are relevant and routinely used in the NHS (See Table 58). Glimepiride was the SU used in the trial, although gliclazide is the SU with the highest level of UK prescribing. In the triple therapy comparison NPH insulin was used but the MS does not report the brand type/name.

Table 58. Economic modelling comparisons conducted

Intervention	**Dual therapy (comparators)**		
Empag 10 mg	Met + Sitagliptin	Met+Canagliflozin (100mg)	Met+Dapagliflozin 10mg
Empag 25 mg	Met+ Sitagliptin	Met+ Canagliflozin (300mg)	
	Triple therapy (comparators)		
Empag 10 mg	Met+SU +Sitagliptin	Met +SU+ Canagliflozin (100mg)	Met+SU+Dapagliflozin 10mg
Empag 25 mg	Met+ SU+ Sitagliptin	Met +SU+ Canagliflozin (100mg)	
Empag 10 mg	Met+Pioglitazone+Sitagliptin	Met+Pioglitazone+Canagliflozin	Met+Pioglitazone+Dapagliflozin 10mg
Empag 25 mg	Met+Pioglitazone+Sitagliptin	Met+Pioglitazone+Canagliflozin	
	Insulin add on (comparators)		
Empag 10 mg	Insulin+ Sitagliptin	Insulin+Canagliflozin (100mg)	Insulin+Dapagliflozin 10mg
Empag 25 mg	Insulin+ Sitagliptin	Insulin+ Canagliflozin (300mg)	

5.14.3 Overview of treatment effects: Data inputs:

The primary outcomes data used in the model is the change in HbA1c, SBP and weight change. The change in lipid ratio was not modelled due to (page 227) "lack of information identified through the systematic review of clinical evidence". The ERG considers that it would have been more appropriate to use direct trial data which were available from at least the ER Met trial, as reported in table 3. However the data more or less match the assumption.

The model requires treatment effectiveness versus placebo to be used but the model reports treatment effectiveness relative to the comparator.

Some of the adverse events are modelled using observed rates (non-severe hypoglycaemic events) and some are modelled as differences from baseline values as drawn from NMA (UTI s). The discontinuation rates for empagliflozin were taken from clinical trials or based upon clinical assumption and the ERG considers that they are reasonable.

The ERG cross-checked the clinical effectiveness values used in the model data input sheet against the NMA results and they match.

Overall the ERG considers that model clinical effectiveness data are appropriate although the principal treatment effectiveness data (HbA1c, SBP and weight gain) taken from the NMA for dual, triple and add-on insulin therapy reports a baseline figure. In the model the placebo baseline figure has been added to the reductions on active drugs. As previously noted, the manufacturer argues that the placebo baseline figures should to be added to the reductions on active drugs.

The ERG has run the model was re-run taking after removing the baseline effect. The results are that the absolute value in terms of health benefits were smaller and the corresponding costs higher than those for the analyses presented in the submission (See Table 59). However the differences in QALYs are negligible and the differences in cost small (under 2% of the totals), so this adjustment makes no real difference.

Table 59. Cost-effectiveness results including and removing placebo effect for dual therapy comparisons

Considering placebo effect	**Met+ Empa 10mg**	**Met+ Sita**	**Incremental**
Mean Cost (£)	32,654	32,046	608
Mean QALYs	7.424	7.425	-0.001
			Dominated by sitagliptin
Removing placebo effect			
	Met+Empa 10mg	**Met+ Sit**	**Incremental**
Mean Cost (£)	32,833	32,258	575
Mean QALYs	7.406	7.409	-0.003
ICER			Dominated by sitagliptin

5.14.4 Overview of treatment effects: Treatment related weight change

The ERG observes that the clinical trials report a decrease in mean body weight (kg) for empagliflozin 10 mg and 25 mg but this has not been adequately captured in the model. The treatment related weight change has been modelled using two scenarios. The first scenario assumed that the weight change associated with the new treatment would happen gradually during the cycles as long as the patient continued receiving the same treatment. The second scenario assumed the change of weight was fixed and the full weight change would happen during the first cycle after initiating a new treatment. In both the scenarios, once the patient had achieved the full weight change, it was assumed that weight would be maintained for a 6 months which has minimal effect on the 40-year model output and the weight changes seen in trials and in NMA outputs have no effect in the model for comparison of treatments.

5.14.5 Overview of Heath related quality of life

HRQoL is applied in the model using utility from the UKPDS study[35] and these appear to correspond to estimates applied in the UKPDS study.

Disutility for weight gain was taken from Bagust and Beale (2005).[36] This study reports the disutility per increased unit of BMI (-0.0061), while the model takes into account weight-related disutility. The MS conducted an additional adjustment to report disutility per kilogram of weight increased. The ERG reassessed the disutility and disagrees with the disutility -0.0159 per kilogram of weight increase.

When considering the QALY adjustment related to weight changes, the submission assumes that weight loss did not have an impact on utilities while weight gain would result in a decrease of utility. ERG suggests that this might be a rather conservative assumption which disadvantages the flozins.

The baseline utilities used in the model were cross checked and while most cases appear to be correct, it is noted that the MS reports EQ-5D disutility of 0.00282 for UTIs/GTIs from a published source[38] but this value does not match with source. The ERG has re-run the model with a disutility value (-0.0246) as reported in the published source. Results are given in Table 60.

Table 60. Comparison of mean QALY obtained with change of disutility for UTI/GTI.

	Met+ Empa 10mg	**Met+ Sita**	**Difference**
Mean QALYs MS	7.424	7.425	-0.001
ERG analysis			
Mean QALYs	7.419	7.421	-0.002
	Met+Empa 25 mg	**Met+ Sit**	**Difference**
Mean QALYs MS	7.433	7.425	0.008
ERG analysis			
Mean QALYs	7.428	7.421	0.007
	Met+ SU+Empa 25mg	**Met+ SU+Sita**	**Difference**
Mean QALYs MS	**6.978**	**6.959**	0.019
ERG analysis			
Mean QALYs	**6.974**	**6.957**	0.017

The differences are trivial. ICERs are highly unstable with such tiny differences in mean QALYS.

5.14.6 Resource use and costs

The costs analysis was conducted from an NHS and PSS perspective and the overall assumptions appear to be reasonable. No drug administration resources or costs were applied in the model.

The source of each of the costs provided in the MS has been checked and in general the resource requirement and costs are accurate. One possible exception is that the annual cost of insulin treatment was sourced from their own study. The insulin pen and needle cost is similar to the cost reported in

the NICE clinical guidelines 87 but the test strip costs reported by the MS are on the higher side. However, this will not affect the overall results.

5.14.7 Base case results

ERG re-ran the model removing second order uncertainty. The results obtained between the probabilistic and deterministic analysis were rather different in terms of mean QALYs and cost.

Re-running the base case analyses produces different probabilistic and deterministic results. Overall the differences in total cost are only a few hundreds of pounds, and differences between QALYs amounts are tiny - fewer than a few days between treatments. However the second or third decimal place differences in incremental QALY can cause large but largely meaningless variations in ICERs. A QALY difference of 0.01, as reported in the comparison of sitagliptin and canagliflozin in table 86, is 3.65 days of perfect health. The QALY difference between sitagliptin and empagliflozin in the same table is about 9 hours of perfect health.

Summary.

If the ERG is right about the flaws in the model, then nothing in this chapter can be taken as reliable.

It the ERG is wrong, and the model is reliable, the findings are much as could be expected from a situation where both the effectiveness and the costs of the drugs are similar. Empagliflozin veers from dominating sitagliptin, to being dominated by it – the analysis is best regarded as showing no difference

6 Conclusions

Empagliflozin is effective in reducing hyperglycaemia, and appears similar in effect to the two flozins already approved by NICE, dapagliflozin and canagliflozin

Adverse effects are as expected, with a slight increase in urinary tract infections in women, for example, 12% versus 8% on placebo,[28] and in genital infections, mainly in women, 7-10% versus none on placebo.[28]

Until we have much longer experience with empagliflozin (and other flozins), it is not possible to be sure that there will be no unexpected serious adverse effects. These could be class effects (all flozins liable to cause AE) or single drug effects (for example, practolol was the only beta-blocker to cause sclerosing peritonitis).

Hypoglycaemia

In the trials, some hypoglycaemia was reported, but mainly in patients on insulin or sulphonylurea, in which case the hypoglycaemia should be attributed to those drugs.

In the dual therapy trial or empagliflozin + metformin versus metformin + placebo, hypoglycaemia was reported, in 0.5% on placebo, 1.8% on empagliflozin 10 mg and 1.4% on empagliflozin 25 mg. However the definition used PG $\underline{<}$ 3.9mmol/l or requiring assistance. None required assistance so the hypoglycaemia reporting is based on the glucose levels. These are not given.

Normal blood glucose levels are often stated to be in the range 3.5 to 5.5moml/l but as Amiel has noted, people without diabetes may often have levels below that.[40] However the ADA has recommended using a threshold of 3.9 mmol/l.[41]

The definition of hypoglycaemia remains under debate. There is evidence of impairment of brain function once plasma glucose (PG) falls below 3 mmol/l so we would want to take action before it fell below that. But to class people with blood glucose in the lowest band of the normal range (3.5 to 3.9 mmol/l) as hypoglycaemic seems inappropriate. A threshold of <3.5 mmol/l would fit better with the normal range.

The ERG thinks it would be reasonable to argue that the flozins do not cause hypoglycaemia.

The ECEM model

The ERG's concerns have been detailed earlier. The ERG has cross checked a number of elements of the visual basic (VB) implementation of the ECEM. This has identified what may be a number of serious issues: random sampling at the patient level, modelling of the evolution of the risk factors, model convergence, model sensitivity to the random seeds chosen, questionable handling of the

application of quality of life values to weight changes and a possible halving of the quality adjusted life year (QALY) decrements associated with adverse events and the complications of diabetes. If the manufacturer confirms that many of these are indeed errors, it will largely invalidate the submitted results. The ECEM has also been constructed so that it can only simulate 100 individual patients if 300 probabilistic sensitivity analysis (PSA) iterations are being conducted. These are unusually low numbers and may limit the ability of the ECEM to reliably discriminate between the overall impacts of different therapies.

Due to the extent and complexity of the coding of this new model, the ERG has not had time to parse all of VB code. It appears that there may have been a lack of validation and stress testing of the model, which may further call into question the robustness and reliability of the remaining code.
The problems with the model bring the ICERs and also the estimates of the uncertainty surrounding them into question. This does not necessarily mean that the conclusion of equivalence, based on clinical trial data and the NMA, is incorrect, but the model does not appear to be capable of demonstrating cost-effectiveness equivalence in a robust way. For example, the error in converting utility per BMI point into utility per Kg change affects the estimation of the effect of weight change in the modelling, and weight loss is one of the main advantages of the flozins.

Renal impairment
The ER Renal trial reported that HbA1c was improved in patients with moderate renal impairment, but the EMA statement says that the action of empagliflozin depends on renal function and that
"Empagliflozin should not be initiated in patients with an eGFR <60 ml/min/1.73 m2 or CrCl <60 ml/min"
though it can be continued in patients in whom it had been started before renal function declined. The EMA recommendation does not seem to fit with the clinical trial results.

Research needs
As always, the ERG would like to see head to head trials, in particular of flozins versus gliptins, rather than rely on NMAs. Long-term data from larger numbers of patients will be required to ensure that there are no adverse events not yet detected.

The ERG has suggested in several previous reports, for STAs, MTAs and the type 2 diabetes guideline development, that there should be a trial in the UK of the effectiveness of intensive lifestyle intervention in type 2 diabetes not controlled on drug treatment. Aas and colleagues carried out a trial of lifestyle intervention in people with poor control on combinations of oral drugs and showed that it was as effective as starting insulin.[22, 42]

As reported, in the clinical effectiveness section, many patients do not reach the HbA1c target on combination therapy with metformin or metformin and a sulphonylurea. One option for them is to add a gliptin, and the results of a trial of adding a fixed dose combination of a gliptin (linagliptin) and a flozin (empagliflozin) were presented at the recent ADA annual meeting.[43] In a 52-week RCT of add-ons to metformin, patients were randomised to empagliflozin 10mg and 25mg, linagliptin 5mg, or the combination of empagliflozin 10 or 25 mg + linagliptin 5mg.

Greater proportions of patients who had HbA1c > 7% at baseline (mean was 8.0%) got down to under 7.0% on the combinations (figures are rounded);

- Empagliflozin 25 + linagliptin 5mg 62%
- Empagliflozin 10mg + linagliptin 5mg 58%
- Empagliflozin 25mg 33%
- Empagliflozin 10mg 28%
- Linagliptin 5mg 36%

Reducing the number of tablets to be taken each day usually helps with adherence so more research into such combinations would be worthwhile.

7 REFERENCES

1. YHPHO. Diabetes Prevalence Model (APHO). 2013 11th April 2013 [cited 10th June 2014]; Available from: http://www.yhpho.org.uk/resource/view.aspx?RID=81090.

2. Hex N, Bartlett C, Wright D, Taylor M, Varley D. Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabetic medicine : a journal of the British Diabetic Association*. 2012;**29**(7):855-62.

3. National Institute for Health and Care Excellence. CG87 Type 2 diabetes: The managemet of type 2 diabetes. NICE guideline. 2009 [cited 10th June 2014]; Available from: http://www.nice.org.uk/guidance/CG87.

4. Saint Denis: L'Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM). Use of Medications Containing Pioglitazone (Actos®, Competact®) Suspended 2011 [cited 9th June 2014]; Available from: http://ansm.sante.fr/var/ansm_site/storage/original/application/4e293bcd0814c025b94d46d7502a0958.pdf.

5. Wright EM, Turk E. The sodium/glucose cotransport family SLC5. *Pflugers Archiv : European journal of physiology*. 2004;**447**(5):510-8.

6. Santer R, Kinner M, Lassen CL, Schneppenheim R, Eggert P, Bald M, *et al.* Molecular analysis of the SGLT2 gene in patients with renal glucosuria. *Journal of the American Society of Nephrology : JASN*. 2003;**14**(11):2873-82.

7. Hanefeld M, Forst T. Dapagliflozin, an SGLT2 inhibitor, for diabetes. *Lancet*. 2010;**375**(9733):2196-8.

8. Polidori D, Sha S, Mudaliar S, Ciaraldi TP, Ghosh A, Vaccaro N, *et al.* Canagliflozin Lowers Postprandial Glucose and Insulin by Delaying Intestinal Glucose Absorption in Addition to Increasing Urinary Glucose Excretion Results of a randomized, placebo-controlled study. *Diabetes care*. 2013;**36**(8):2154-61.

9. National Institute for Health and Care Excellence. NICE Guidance TA288. Dapagliflozin in combination therapy for treating type 2 diabetes. . 2013 [cited 8th June 2014]; Available from: http://www.nice.org.uk/guidance/TA288/chapter/1-guidance.

10. National Institute for Health and Care Excellence. NICE Guidance TA315. Canagliflozin in combination therapy for treating type 2 diabetes. 2014 [cited 27th June 2014]; Available from: http://www.nice.org.uk/guidance/TA315/chapter/1-guidance.

11. British National Formulatory. BNF 67. March - September. 2014 [cited 10th June 2014]; Available from: http://www.bnf.org/bnf/index.htm.

12. Copley V, Pickett K, Shepher 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 [cited 5th June 2014]; Available from: http://www.nice.org.uk/guidance/ta315/resources/diabetes-type-2-canagliflozin-evaluation-report4.

13. Van den Donk M, Sandbaek A, Borch-Johnsen K, Lauritzen T, Simmons R, Wareham N, *et al.* Screening for Type 2 diabetes. Lessons from the ADDITION-Europe study. *Diabetic Medicine*. 2011;**28**(11):1416-24.

14. Waugh N, Cummins E, Royle P, Clar C, Marien M, Richter B, *et al.* Newer agents for blood glucose control in type 2 diabetes: systematic review and economic evaluation. *Health technology assessment (Winchester, England)*. 2010;**14**(36):1-248.

15. Peyrot M, Rubin RR, Lauritzen T, Skovlund SE, Snoek FJ, Matthews DR, *et al.* Resistance to insulin therapy among patients and providers: results of the cross-national Diabetes Attitudes, Wishes, and Needs (DAWN) study. *Diabetes care*. 2005;**28**(11):2673-9.

16. Alvarez Guisasola F, Mavros P, Nocea G, Alemao E, Alexander CM, Yin D. Glycaemic control among patients with type 2 diabetes mellitus in seven European countries: findings from the Real-Life Effectiveness and Care Patterns of Diabetes Management (RECAP-DM) study. *Diabetes, obesity & metabolism*. 2008;**10 Suppl 1**:8-15.

17. Rubino A, McQuay LJ, Gough SC, Kvasz M, Tennis P. Delayed initiation of subcutaneous insulin therapy after failure of oral glucose-lowering agents in patients with Type 2 diabetes: a population-based analysis in the UK. *Diabetic medicine : a journal of the British Diabetic Association*. 2007;**24**(12):1412-8.

18. Calvert MJ, McManus RJ, Freemantle N. Management of type 2 diabetes with multiple oral hypoglycaemic agents or insulin in primary care: retrospective cohort study. *The British journal of general practice : the journal of the Royal College of General Practitioners*. 2007;**57**(539):455-60.

19. European Medicines Agency (EMA). European Medicines Agency recommends new contraindications and warnings for pioglitazone to reduce small increased risk of bladder cancer. 2011 [cited 11th June 2014]; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2011/07/WC500109176.pdf.

20. Association of British Clinical Diabetologists (ABCD). Conclusions of European Medicines Authoristy safety review of pioglitazone welcomed by Association of British Clinical Diabetologists. 2011 [cited 15th June 2014]; Available from: https://www.diabetologists.org.uk/Shared_Documents/position_papers/ABCD_pioglitazone_press_release_23.7.11.pdf.

21. Health and Social Care Information Centre. Prescribing for Diabetes, England 2005-6 to 2012-13. 2013 [cited 15th June 2014]; Available from: http://www.hscic.gov.uk/catalogue/PUB11422/pres-diab-eng-2005-2013-rep.pdf.

22. Aas AM, Bergstad I, Thorsby PM, Johannesen O, Solberg M, Birkeland KI. An intensified lifestyle intervention programme may be superior to insulin treatment in poorly controlled Type 2 diabetic patients on oral hypoglycaemic agents: results of a feasibility study. *Diabetic medicine : a journal of the British Diabetic Association*. 2005;**22**(3):316-22.

23. European Medicines Agency (EMA). Byetta (Exenatide):EPAR Summary for the Public. 2012 [cited 12th June 2014]; Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000698/human_med_000682.jsp&mid=WC0b01ac058001d124.

24. European Medicines Agency (EMA). Committee for Medicinal Products for Human Use (CHMP). Outcome of extension of indication application for Victoza (liraglutide). 2011 [cited 11th June 2014]; Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2011/09/news_detail_001338.jsp&mid=WC0b01ac058004d5c1.

25. Black C, Donnelly P, McIntyre L, Royle PL, Shepherd JP, Thomas S. Meglitinide analogues for type 2 diabetes mellitus. *The Cochrane database of systematic reviews*. 2007(2):Cd004654.

26. Canadian Agency for Drugs and Technologies in Helath (CADTH). CADTH Therapeutic Review Clinical Review: Third-Line Therapy for Patients with Type 2 Diabetes Inadequately Controlled with Metformin and a Sulfonylurea. 2010 [cited 8th June 1014]; Available from: http://www.cadth.ca/media/pdf/Diabetes_TR_Clinical_Report_Final_e.pdf.

27. Haring HU, Merker L, Seewaldt-Becker E, Weimer M, Meinicke T, Woerle HJ, *et al.* Empagliflozin as add-on to metformin plus sulfonylurea in patients with type 2 diabetes: a 24-week, randomized, double-blind, placebo-controlled trial. *Diabetes care*. 2013;**36**(11):3396-404.

28. Haring HU, Merker L, Seewaldt-Becker E, Weimer M, Meinicke T, Broedl UC, *et al.* Empagliflozin as add-on to metformin in patients with type 2 diabetes: a 24-week, randomized, double-blind, placebo-controlled trial. *Diabetes care*. 2014;**37**(6):1650-9.

29. Rosenstock J, Jelaska A, Frappin G, Salsali A, Kim G, Woerle HJ, *et al.* Improved glucose control with weight loss, lower insulin doses, and no increased hypoglycemia with empagliflozin added to titrated multiple daily injections of insulin in obese inadequately controlled type 2 diabetes. *Diabetes care*. 2014;**37**(7):1815-23.

30. Higgins J, Green S, Collaboration TC. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011 [cited 3rd June 2014]; Available from: http://www.cochrane.org/handbook.

31. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S. Evidence synthesis for decision making 7: a reviewer's checklist. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2013;**33**(5):679-91.

32. Donegan S, Williamson P, Gamble C, Tudur-Smith C. Indirect comparisons: a review of reporting and methodological quality. *PloS one*. 2010;**5**(11):e11054.

33. McIntosh B, Cameron C, Singh SR, Yu C, Ahuja T, Welton NJ, *et al.* Second-line therapy in patients with type 2 diabetes inadequately controlled with metformin monotherapy: a systematic review and mixed-treatment comparison meta-analysis. *Open medicine : a peer-reviewed, independent, open-access journal*. 2011;**5**(1):e35-48.

34. Davis S, Stevenson M, Tappenden P, Wailoo AJ. NICE DSU technical support document 15: cost-effectiveness modelling using patient-level simulation. 2014 [cited 15th June 2014]; Available from: http://www.nicedsu.org.uk/TSD15_Patient-level_simulation.pdf.

35. Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). *Medical Decision Making*. 2002;**22**(4):340-9.

36. Bagust A, Beale S. Modelling EuroQol health-related utility values for diabetic complications from CODE-2 data. *Health Economics*. 2005;**14**(3):217-30.

37. Clarke P, Gray A, Briggs A, Farmer A, Fenn P, Stevens R, *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**(10):1747-59.

38. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. *Medical Decision Making*. 2011;**31**(6):800-4.

39. Clarke P, Gray A, Legood R, Briggs A, Holman R. The impact of diabetes-related complications on healthcare costs: results from the United Kingdom Prospective Diabetes Study (UKPDS Study No. 65). *Diabetic Medicine*. 2003;**20**(6):442-50.

40. Amiel SA. It's all in the name: what exactly is hypoglycaemia? *Diabetic Hypoglycaemia*. 2013;**5**(3):15-8.

41. Workgroup on Hypoglycemia/American Diabetes Association. Defining and reporting hypoglycemia in diabetes: a report from the American Diabetes Association Workgroup on Hypoglycemia. *Diabetes care*. 2005;**28**(5):1245-9.

42. Aas AM, Seljeflot I, Torjesen PA, Diep LM, Thorsby PM, Birkeland KI. Blood glucose lowering by means of lifestyle intervention has different effects on adipokines as compared with insulin treatment in subjects with type 2 diabetes. *Diabetologia*. 2006;**49**(5):872-80.

43. DeFronzo R, Lewin A, Patel S, Liu D, Kaste R, Woerle HJ, *et al.*, editors. Fixed-Dose Combinations of Empagliflozin/Linagliptin for 24 Weeks as Add-On to Metformin in Patients with Type 2 Diabetes (T2DM). 130-LB 74th American Diabetes Association scientific sessions; 2014 June 13 to 17; San Francisco.

8 APPENDICES

Appendix 1. Quality assessment of included studies using Cochrane risk of bias

Table 61. Study quality (1245.23- Met and Met+SU)

Items	Description	High/Low/Unclear
Adequate sequence generation	Interactive voice and web response system (IXRS)	Low
Allocation concealment	IXRS	Low
Masking	Patient, investigator and outcome assessor	Low
Incomplete outcome data addressed	Primary analyses undertaken on the full analysis set (FAS); missing values were imputed using LOCF method; adequate description of loss to follow-up	Low
Free of selective reporting	All pre-defined and pre-specified outcomes were reported	Low
Free of other bias (e.g. similarity at baseline, power assessment)	Baseline characteristics similar across all treatment groups; Power calculation done	Low
Funder	Boehringer Ingelheim Ltd	

Table 62. Study quality (1245.33- Basal insulin)

Items	Description	High/Low/Unclear
Adequate sequence generation	Interactive voice and web response system (IXRS)	Low
Allocation concealment	IXRS	Low
Masking	Patient, investigator and outcome assessor	Low
Incomplete outcome data addressed	Primary analyses undertaken on the full analysis set (FAS); missing values were imputed using LOCF method; adequate description of loss to follow-up	Low
Free of selective reporting	All pre-defined and pre-specified outcomes were reported	Low
Free of other bias (e.g. similarity at baseline, power assessment)	Baseline characteristics similar across all treatment groups; Power calculation done	Low
Funder	Boehringer Ingelheim Ltd	

Table 63. Study quality (1245.49 - MDI)

Items	Description	High/Low/Unclear
Adequate sequence generation	Interactive voice and web response system (IXRS)	Low
Allocation concealment	IXRS	Low
Masking	Patient, investigator and outcome assessor	Low
Incomplete outcome data addressed	Primary analyses undertaken on the full analysis set (FAS); missing values were imputed using LOCF method; adequate description of loss to follow-up	Low
Free of selective reporting	All pre-defined and pre-specified outcomes were reported	Low
Free of other bias (e.g. similarity at baseline, power assessment)	Baseline characteristics similar across all treatment groups except gender; Power calculation done	Low
Funder	Boehringer Ingelheim Ltd	