A novel model-based meta-analysis to indirectly estimate the comparative efficacy of two medications: an example using DPP-4 inhibitors, sitagliptin and linagliptin, in treatment of type 2 diabetes mellitus

Jorge Luiz Gross,1 James Rogers,2 Daniel Polhamus,2 William Gillespie,2 Christian Friedrich,3 Yan Gong,4 Brigitta Ursula Monz,4 Sanjay Patel,5 Alexander Staab,3 Silke Retlich3

ABSTRACT

Objectives: To develop a longitudinal statistical model to indirectly estimate the comparative efficacies of two drugs, using model-based meta-analysis (MBMA). Comparison of two oral dipeptidyl peptidase (DPP)-4 inhibitors, sitagliptin and linagliptin, for type 2 diabetes mellitus (T2DM) treatment was used as an example.

Design: Systematic review with MBMA.


Eligibility criteria for selecting studies: Double-blind, randomised controlled clinical trials, ≥12 weeks’ duration, that analysed sitagliptin or linagliptin efficacies as changes in glycated haemoglobin (HbA1c) levels, in adults with T2DM and HbA1c >7%, irrespective of background medication.

Model development and application: A Bayesian model was fitted (Markov Chain Monte Carlo method). The final model described HbA1c levels as function of time, dose, baseline HbA1c, washout status/duration and ethnicity. Other covariates showed no major impact on model parameters and were not included. For the indirect comparison, a population of 1000 patients was simulated from the model with a racial composition reflecting the average racial distribution of the linagliptin trials, and baseline HbA1c of 8%.

Results: The model was developed using longitudinal data from 11 234 patients (10 linagliptin, 15 sitagliptin trials), and assessed by internal evaluation techniques, demonstrating that the model adequately described the observations. Simulations showed both linagliptin 5 mg and sitagliptin 100 mg reduced HbA1c by 0.81% (placebo-adjusted) at week 24. Credible intervals for participants without washout were −0.88 to −0.75 (linagliptin) and −0.89 to −0.73 (sitagliptin), and for those with washout, −0.91 to −0.76 (linagliptin) and −0.91 to −0.75 (sitagliptin).

ARTICLE SUMMARY

Article focus

- In the absence of evidence from head-to-head trials, indirect and mixed treatment comparisons can be used for drug comparisons.
- The aim of this study was to develop an approach, using Bayesian methodology (Markov Chain Monte Carlo method) to indirectly estimate the comparative efficacy of two compounds, incorporating longitudinal dose–response data.

Key messages

- A longitudinal statistical model was developed for the indirect comparison of two pharmaceutical compounds (oral DPP-4 inhibitors linagliptin and sitagliptin), with respect to changes in glycated haemoglobin (HbA1c) levels in patients with type 2 diabetes mellitus (T2DM).
- The model was evaluated by comparing model predictions with observed values.
- The model demonstrated that both linagliptin and sitagliptin reduced HbA1c levels by 0.8% (placebo-adjusted) when administered to patients with T2DM for 24 weeks, irrespective of background medication.

Conclusions: This study demonstrates the use of longitudinal MBMA in the field of diabetes treatment. Based on an example evaluating HbA1c reduction with linagliptin versus sitagliptin, the model used seems a valid approach for indirect drug comparisons.

INTRODUCTION

Ideally, head-to-head, randomised controlled trials should be conducted to estimate the
Novel MBMA for indirect comparison of diabetes treatments

ARTICLE SUMMARY

Strengths and limitations of this study
- This study represents a novel use of longitudinal model-based meta-analysis in the field of diabetes treatment, being the only instance to date that adequately accounts for longitudinal correlations in each treatment arm, which is a prerequisite to the correct characterisation of uncertainty in estimation of drug effects.
- When relevant head-to-head comparisons are not available, the model described in this study could have an important role in treatment decision-making.
- Although the analysis included a large sample of 11,234 patients with T2DM, its applicability to the general population of patients with T2DM might be limited by the relatively selected patient populations in the included trials. Additionally, while our analysis adjusts for key differences in study designs, there remains the possibility of bias attributable to covariate effects that could not be estimated with the available data.

comparative efficacy of different treatments. However, it is not always feasible to conduct direct comparisons among all available treatment options. Network meta-analysis (mixed treatment comparisons) has been used to estimate relative efficacy when there are no direct comparative data, to provide the best available evidence to facilitate decision-making by physicians and other stakeholders, such as payers. However, these approaches have certain limitations, including the risk of bias arising from inherent differences in the designs of the included studies, and the difficulties of finding appropriate summary statistics to compare the findings of individual trials. In particular, endpoint-based approaches cannot be sensibly applied when the studies involved in the review vary substantially with respect to treatment duration.

An approach, recently described as model-based meta-analysis (MBMA), has been developed and used to estimate the comparative efficacy of two medications. MBMA can be used to provide a mechanism for integrating information from heterogeneously designed trials and thus to evaluate outcomes with different drugs that have not been compared directly. MBMA is distinguished from the methodology of conventional meta-analysis by the manner in which it incorporates longitudinal and/or dose-response data. By modelling the response as a parametric function of time, MBMA allows the integration of information from trials of different durations and with different sampling time-points. This enables the use of less restrictive inclusion/exclusion criteria for study selection, and more efficient use of data from the studies that are selected, thereby resulting in a particularly comprehensive summary of all relevant data.

In response to the growing worldwide epidemic of diabetes mellitus, new antihyperglycaemic agents are continuously being developed. The dipeptidyl peptidase (DPP)-4 inhibitors are a relatively new class of oral antihyperglycaemic drugs developed for the treatment of type 2 diabetes mellitus (T2DM) that are increasingly being used in clinical practice because of their clinically meaningful efficacy, promising tolerability, safety and convenience—in particular, a virtually absent risk of hypoglycaemia or weight gain. Although several DPP-4 inhibitors are already available in many countries, to date, only one published trial has been conducted to directly compare individual drugs within this class. Therefore, further research is needed to understand the comparative effects of the drugs within this class.

The model developed in this study incorporates Bayesian methodology and aims to provide a valid approach to estimate the comparative efficacy of different compounds. Bayesian approaches are acknowledged by the Cochrane collaboration to have a role in meta-analysis, particularly in the setting of indirect comparison.

This approach to drug comparison employs a mathematical model to describe the timecourse of glycated haemoglobin (HbA1c) levels, and is being increasingly used to characterise longitudinal data. The general meta-analytic methodology of Ahn and French has previously been used to successfully describe longitudinal metadata from clinical trials in Alzheimer’s disease, rheumatoid arthritis, lipid disorders, glaucoma and chronic obstructive pulmonary disease. Similar approaches have been used to perform dose–response meta-analyses in a range of therapeutic areas, including migraine, postoperative anticoagulant therapy and rheumatoid arthritis. This analytical approach has also been used in the field of diabetes in a recent study by Gibbs et al., which evaluated the relationship between DPP-4 inhibition and HbA1c reduction using data obtained from clinical trials of four drugs in this class.

Objective
To use an MBMA approach to develop a longitudinal statistical model for the comparison of the efficacy of two oral DPP-4 inhibitors, shown by changes in HbA1c levels, in patients with T2DM who had started treatment with one of two DPP-4 inhibitors, regardless of background medication. The two drugs evaluated were linaagliptin, which has recently been approved for clinical use in several jurisdictions, and sitagliptin, the most commonly used DPP-4 inhibitor.

METHODS
Data sources
Sitagliptin studies were identified from a systematic search in MEDLINE, EMBASE, studies listed on http://www.ClinicalTrials.gov that included a reference to publication, the latest-date Cochrane review of DPP-4 inhibitors for T2DM and details of sitagliptin trials on the Food and Drug Administration website, to December 2011. Details of the search strategy used are provided in the appendix (see online supplementary table S1).

Results of the relevant studies for linagliptin were obtained from the manufacturer’s database, several of...
which have been subsequently published as full papers\textsuperscript{18–21} or abstracts.\textsuperscript{22–24}

**Study selection**

Included studies were double-blind, randomised controlled trials of \(\geq 12\) weeks' duration that analysed the efficacy of sitagliptin or linagliptin in the reduction of HbA1c levels in adults with T2DM and HbA1c \(>7\%\), irrespective of background medication. Excluded studies were: open-label studies (and data from open-label extensions to double-blind studies) and extension studies that used patient response in the initial study to determine eligibility in the extension phase of the study (eg, if the extension phase included only those who did not require rescue medication during the initial study). Other excluded study types were special population studies (eg, studies in patients with declining renal function) and phase IV studies or study arms in which patients were randomly assigned to initial combination therapies.

Two independent reviewers extracted aggregated data from all selected studies, according to treatment arm (sitagliptin, linagliptin or placebo). We extracted data on: the first author's name, year of publication of the trial, comparator, dose(s) of sitagliptin or linagliptin evaluated, trial duration, number of participants and their gender, ethnicity, duration of T2DM, mean age, baseline HbA1c (%), HbA1c at evaluated time-points, baseline body mass index (BMI, kg/m\(^2\)), fraction of patients on previous anti-hyperglycaemic therapy, and the presence and duration of washout and concomitant medication. A common data template was defined. The main outcome of interest was HbA1c, the primary endpoint of all included studies. Intention-to-treat populations were included whenever possible and group means, as reported, were used or were calculated, using the last observation carried forward approach. The analyses were conducted using the maximum licensed dose of sitagliptin (100 mg) and the licensed dose of linagliptin (5 mg). However, when data at other dose levels were available, they were included in the analysis, and appropriate adjustments were made via the dose–response terms in the model.

**Data selection process**

For the linagliptin studies, the dataset was built from the original Boehringer Ingelheim database using SAS scripting. The quality of the dataset was assured by an independent script review. For the sitagliptin studies, the dataset was built manually by collecting information given in the different source publications. If the results were available as numbers in the publications, these numbers were included in the dataset. Where the results were only available as graphics, the corresponding data were collected using GetData Graph Digitizer, V.2.24 software (http://www.getdata-graph-digitizer.com). The quality of the manually built sitagliptin dataset was assured by an independent second reviewer. The initial dataset consisted of HbA1c data, presented as either the change from baseline and/or the actual HbA1c measurements, depending on the information provided in the publication, R scripting (R V.2.10.1, The R Foundation for Statistical Computing, Vienna, Austria) was then used to obtain an analysis-ready dataset with consistent encoding of information (eg, baseline values were added to changes from baseline in order to obtain actual HbA1c measurements for all records).\textsuperscript{25}

**Statistical analysis**

**Model development**

The statistical models that were considered represent a particular class of non-linear mixed-effects models in which model precision terms are scaled according to sample sizes. Sample size adjustments are carried out in a manner that approximately estimates and adjusts for longitudinal correlations, following an approach described elsewhere.\textsuperscript{3}

Initial exploratory data analyses were used to derive a suitable parametric (algebraic) description of the average HbA1c trends as a function of time, dose, washout status/duration and ethnic origin. Qualitative prior information was also used to guide the initial selection of parametric forms. The following assumptions were made: (1) given the known properties of measured HbA1c, it was assumed that in the absence of additional interventions, HbA1c levels for patients washing out prior antidiabetes medication (during the study washout/run-in phase) would rise for some time until achieving a plateau, and (2) the incremental (placebo-adjusted) effect of DPP-4 inhibitors on HbA1c was expected to approach a plateau during the time frame of interest (24 weeks). Bayesian prior distributions for parameters describing the magnitude and onset of drug effects were specified separately and independently for linagliptin and sitagliptin. Magnitudes of drug effect were parameterised as fractional reductions from baseline and were assigned uniform prior distributions between zero and one, implying that both drugs have some beneficial effects (a defensible assumption for marketed drugs) and that neither can reduce HbA1c levels below zero (patently true), and assigning equal likelihood to all possibilities between these two extremes.

The model was fitted using Bayesian Markov Chain Monte Carlo methodology. The computations were carried out using OpenBUGS V.3.2.1 (2010) software (Free Software Foundation, Boston, Massachusetts, USA). Final inferences were based on 1000 approximately independent draws from the posterior (after discarding burn-in samples and thinning to de-correlate samples)\textsuperscript{26}. The model was adjusted for baseline HbA1c and washout status/duration. Other covariates considered were: standard covariates including demographics, such as ethnicity, age, BMI and gender, antihyperglycaemic background medication, duration of T2DM and the fraction of patients who underwent washout of previous antihyperglycaemic therapy. The OpenBUGS code is available from the authors on request.

Model selection and evaluation
Following a ‘full model estimation approach’, initial preference was given to a full model, meaning one that includes all terms of potential interest. In order to achieve stable parameter estimation, selective simplifications were applied, guided by exploratory data analysis, to the full model until we obtained satisfactory convergence diagnostics. Covariates were excluded from the model for the purpose of achieving stable parameter estimation; however, each excluded covariate was evaluated graphically to ensure that it was not associated with model residuals (differences between the observed values and those predicted by the model). A graphic representation of the final model, for patients with or without a prerandomisation washout period, is shown in figure 1A,B.

The final model was evaluated using posterior predictive check methodology in order to assess whether the observed data were consistent with the range of expectation implied by the model. This model inherently adjusted for baseline HbA1c and washout status/duration. The other covariates (see previous page), with the exception of ethnicity, showed no major impact on the model parameters and were therefore not included in the final model. Further details of the mathematical and statistical specifications of the final model are presented in the online supplementary technical appendix.

Figure 1  (A) Graphic representation of the components of the final model, for study arms that included patients washing out their prior antihyperglycaemic medication in the run-in period. (B) Graphic representation of the components of the final model, for study arms that included patients who were treatment-naïve or had completely washed out their prior antihyperglycaemic medication before enrolment.

Model summary and inference
Since the mean predicted values are not directly available as model parameters, these were estimated by taking averages of values that were simulated from the fitted model. In the same way that variances can be appropriately scaled according to sample size during model fitting, variances were scaled during simulation to simulate trial arms of different sizes. This included scaling simulation variances to correspond to n=1, which we conceptualised as the simulation of an individual patient.

In order to assess the efficacy of the two DPP-4 inhibitors in comparable patients under similar conditions, a population of 1000 patients was simulated from the model under reference conditions and the average HbA1c level was computed at each time-point for this simulated population. Data for each patient were simulated as if arising from an individual trial, so that the resulting inference represents an average over the expected range of intertrial variation. The simulation of this population average was then repeated for each of the 1000 different parameter configurations represented in the posterior sample (the entire posterior simulation therefore involved a total of 10^6 simulated patients), resulting in inferences that reflect posterior parameter uncertainty as well as intertrial and interpatient variation. The reference racial composition for this simulated population was 61.5% White, 1.5% Black and 37% Asian, reflecting the average enrolment distribution in linagliptin trials. The median simulated baseline HbA1c (%) in this population was 8 Results are expressed as mean differences, with 95% credible intervals (the Bayesian equivalent of CIs).

RESULTS
A total of 31 sitagliptin studies were assessed for eligibility for inclusion in the analysis, and 16 were excluded on the basis of the study design that did not meet our inclusion criteria (see online supplementary table S2). A further 10 linagliptin studies were included.

The included studies were between 12 and 26 weeks’ duration, with one exception (the study by Seck et al lasted 104 weeks; table 1).

Data from a total of 11 234 participants were included in the analysis, arising from 25 randomised trials (10 linagliptin and 15 sitagliptin). The mean age at baseline of all study participants was 56.5 years, with reported means for treatment arms of the included studies ranging from 50.9 to 62 years; the proportion of women across all study participants was 56.5%, with reported proportions for study groups ranging from 22.8% to 64%; and the mean BMI was 29.7 kg/m^2, with reported means for treatment arms ranging from 24.1 to 32.7 kg/m^2. Mean baseline HbA1c was 8%, with reported means for treatment arms ranging from 7.49% to 8.87%. The most commonly used background medication was metformin monotherapy. Metformin was also used in combination with glimepiride or pioglitazone, and one study included patients receiving initial monotherapy with pioglitazone.
<table>
<thead>
<tr>
<th>Study</th>
<th>Drug</th>
<th>Dose (mg/day)</th>
<th>Treatment duration (weeks)</th>
<th>Patients (n)</th>
<th>Baseline age (years)</th>
<th>Female (%)</th>
<th>Baseline HbA1c (%)</th>
<th>Baseline BMI (kg/m²)</th>
<th>Washout duration (weeks)</th>
<th>Concomitant medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aschner et al</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>244</td>
<td>54.3</td>
<td>48.6</td>
<td>8.03</td>
<td>30.8</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>24</td>
<td>229</td>
<td>53.4</td>
<td>42.9</td>
<td>8.01</td>
<td>30.3</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>24</td>
<td>238</td>
<td>54.9</td>
<td>53.2</td>
<td>8.08</td>
<td>30.3</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td>Bergenstal et al</td>
<td>Sitagliptin</td>
<td>100</td>
<td>26</td>
<td>166</td>
<td>52.0</td>
<td>48.0</td>
<td>8.50</td>
<td>32.0</td>
<td>0</td>
<td>Metformin</td>
</tr>
<tr>
<td>Charbonnel et al</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>224</td>
<td>54.7</td>
<td>40.5</td>
<td>8.03</td>
<td>31.5</td>
<td>18</td>
<td>Metformin</td>
</tr>
<tr>
<td>Goldstein et al</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>165</td>
<td>53.3</td>
<td>47.2</td>
<td>8.68</td>
<td>32.5</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td>Hanefeld et al</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>175</td>
<td>53.6</td>
<td>48.0</td>
<td>8.87</td>
<td>31.2</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>12</td>
<td>107</td>
<td>55.9</td>
<td>36.9</td>
<td>7.59</td>
<td>31.4</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>12</td>
<td>107</td>
<td>55.1</td>
<td>48.6</td>
<td>7.71</td>
<td>31.9</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>12</td>
<td>107</td>
<td>55.3</td>
<td>54.5</td>
<td>7.60</td>
<td>31.6</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>12</td>
<td>108</td>
<td>55.2</td>
<td>55.9</td>
<td>7.79</td>
<td>32.7</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>12</td>
<td>106</td>
<td>56</td>
<td>44.5</td>
<td>7.78</td>
<td>31.6</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Hermansen et al</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>106</td>
<td>55.2</td>
<td>45.3</td>
<td>8.43</td>
<td>30.7</td>
<td>16</td>
<td>Gilimepiride</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>24</td>
<td>106</td>
<td>55.4</td>
<td>47.2</td>
<td>8.42</td>
<td>31.0</td>
<td>16</td>
<td>Gilimepiride</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>113</td>
<td>57.7</td>
<td>47.8</td>
<td>8.26</td>
<td>30.7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>24</td>
<td>116</td>
<td>56.6</td>
<td>47.4</td>
<td>8.27</td>
<td>31.3</td>
<td>16</td>
<td>Gilimepiride + metformin</td>
</tr>
<tr>
<td>Iwamoto et al</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>73</td>
<td>60.2</td>
<td>31.5</td>
<td>7.74</td>
<td>24.1</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>25</td>
<td>12</td>
<td>80</td>
<td>59.9</td>
<td>36.3</td>
<td>7.49</td>
<td>25.0</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>12</td>
<td>72</td>
<td>60.2</td>
<td>34.7</td>
<td>7.57</td>
<td>24.5</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>12</td>
<td>70</td>
<td>58.3</td>
<td>48.6</td>
<td>7.56</td>
<td>24.2</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Mohan et al</td>
<td>Placebo</td>
<td>NA</td>
<td>18</td>
<td>169</td>
<td>50.9</td>
<td>40.0</td>
<td>8.70</td>
<td>24.9</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>18</td>
<td>339</td>
<td>50.9</td>
<td>43.0</td>
<td>8.70</td>
<td>25.1</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Nonaka et al</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>75</td>
<td>55.0</td>
<td>34.0</td>
<td>7.69</td>
<td>25.1</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>10</td>
<td>12</td>
<td>75</td>
<td>55.6</td>
<td>40.0</td>
<td>7.54</td>
<td>25.2</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Raz et al</td>
<td>Placebo</td>
<td>NA</td>
<td>18</td>
<td>103</td>
<td>55.5</td>
<td>37.3</td>
<td>8.05</td>
<td>32.5</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>18</td>
<td>193</td>
<td>54.5</td>
<td>46.3</td>
<td>8.04</td>
<td>31.8</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td>Rosenstock et al</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>174</td>
<td>56.9</td>
<td>46.9</td>
<td>8.00</td>
<td>31.0</td>
<td>18</td>
<td>Pioglitazone</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>100</td>
<td>24</td>
<td>163</td>
<td>55.6</td>
<td>42.1</td>
<td>8.05</td>
<td>32.0</td>
<td>18</td>
<td>Pioglitazone</td>
</tr>
<tr>
<td>Scott et al</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>121</td>
<td>55.3</td>
<td>37.6</td>
<td>7.88</td>
<td>31.6</td>
<td>10</td>
<td>Metformin</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>10</td>
<td>12</td>
<td>122</td>
<td>55.1</td>
<td>50.4</td>
<td>7.89</td>
<td>30.8</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Seck et al</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>122</td>
<td>56.2</td>
<td>52</td>
<td>7.85</td>
<td>30.5</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>25</td>
<td>12</td>
<td>122</td>
<td>55.9</td>
<td>42.3</td>
<td>7.89</td>
<td>31.4</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>12</td>
<td>120</td>
<td>55.6</td>
<td>42.3</td>
<td>7.89</td>
<td>31.4</td>
<td>8</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 1  Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Drug</th>
<th>Dose (mg/day)</th>
<th>Treatment duration (weeks)</th>
<th>Patients (n)</th>
<th>Baseline age (years)</th>
<th>Female (%)</th>
<th>Baseline HbA1c (%)</th>
<th>Baseline BMI (kg/m²)</th>
<th>Washout duration (weeks)</th>
<th>Concomitant medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scott et al 42</td>
<td>Placebo</td>
<td>100</td>
<td>12</td>
<td>121</td>
<td>55.1</td>
<td>47.6</td>
<td>7.96</td>
<td>30.4</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>18</td>
<td>12</td>
<td>88</td>
<td>55.3</td>
<td>41.0</td>
<td>7.68</td>
<td>30.0</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Boehringer Ingelheim</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>91</td>
<td>55.2</td>
<td>45.0</td>
<td>7.75</td>
<td>30.3</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>Study 1218.5 83</td>
<td>Linagliptin</td>
<td>100</td>
<td>12</td>
<td>63</td>
<td>59.0</td>
<td>49.2</td>
<td>8.27</td>
<td>30.9</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>12</td>
<td>57</td>
<td>58.0</td>
<td>22.8</td>
<td>8.24</td>
<td>31.0</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.5</td>
<td>12</td>
<td>55</td>
<td>60.0</td>
<td>52.7</td>
<td>8.38</td>
<td>31.5</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>12</td>
<td>54</td>
<td>56.0</td>
<td>42.6</td>
<td>8.38</td>
<td>31.2</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Sitagliptin</td>
<td>18</td>
<td>91</td>
<td>55.2</td>
<td>45.0</td>
<td>7.75</td>
<td>8.24</td>
<td>30.3</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>Forst et al 19</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>70</td>
<td>60.0</td>
<td>38.6</td>
<td>8.37</td>
<td>32.2</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>1</td>
<td>12</td>
<td>64</td>
<td>59.0</td>
<td>43.8</td>
<td>8.24</td>
<td>32.2</td>
<td>6</td>
<td>Metformin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>12</td>
<td>62</td>
<td>60.0</td>
<td>46.8</td>
<td>8.46</td>
<td>31.6</td>
<td>6</td>
<td>Metformin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>12</td>
<td>66</td>
<td>62.0</td>
<td>47.0</td>
<td>8.35</td>
<td>31.7</td>
<td>6</td>
<td>Metformin</td>
</tr>
<tr>
<td>Del Prato et al 18</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>163</td>
<td>55.0</td>
<td>54.0</td>
<td>8.00</td>
<td>29.2</td>
<td>6</td>
<td>Metformin</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>24</td>
<td>333</td>
<td>56.0</td>
<td>51.4</td>
<td>8.00</td>
<td>29.0</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td>Taskinen et al 21</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>175</td>
<td>57.0</td>
<td>42.3</td>
<td>8.02</td>
<td>30.1</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>24</td>
<td>513</td>
<td>57.0</td>
<td>46.8</td>
<td>8.09</td>
<td>29.8</td>
<td>6</td>
<td>Metformin</td>
</tr>
<tr>
<td>Owens et al 20</td>
<td>Placebo</td>
<td>NA</td>
<td>24</td>
<td>262</td>
<td>58.0</td>
<td>53.2</td>
<td>8.14</td>
<td>28.2</td>
<td>6</td>
<td>Metformin</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>24</td>
<td>778</td>
<td>58.0</td>
<td>51.5</td>
<td>8.15</td>
<td>28.4</td>
<td>0</td>
<td>Metformin+SU</td>
</tr>
<tr>
<td>Gallwitz et al 22</td>
<td>Linagliptin</td>
<td>5</td>
<td>52</td>
<td>776</td>
<td>60.0</td>
<td>40.7</td>
<td>7.69</td>
<td>30.2</td>
<td>0</td>
<td>Metformin+SU</td>
</tr>
<tr>
<td>Araki et al 24</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>80</td>
<td>60.0</td>
<td>28.6</td>
<td>7.95</td>
<td>24.3</td>
<td>8</td>
<td>Metformin</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>12</td>
<td>159</td>
<td>60.0</td>
<td>30.2</td>
<td>8.07</td>
<td>24.6</td>
<td>4</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>12</td>
<td>160</td>
<td>61.0</td>
<td>30.0</td>
<td>7.98</td>
<td>25.0</td>
<td>4</td>
<td>NA</td>
</tr>
<tr>
<td>Lewin et al 45</td>
<td>Placebo</td>
<td>NA</td>
<td>18</td>
<td>82</td>
<td>56.0</td>
<td>39.0</td>
<td>8.60</td>
<td>28.1</td>
<td>4</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>18</td>
<td>158</td>
<td>57.0</td>
<td>52.5</td>
<td>8.61</td>
<td>28.3</td>
<td>6</td>
<td>SU</td>
</tr>
<tr>
<td>Patel et al 23</td>
<td>Placebo</td>
<td>NA</td>
<td>18</td>
<td>73</td>
<td>56.0</td>
<td>57.5</td>
<td>8.06</td>
<td>30.0</td>
<td>6</td>
<td>SU</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>18</td>
<td>147</td>
<td>57.0</td>
<td>64.0</td>
<td>8.11</td>
<td>29.0</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td>Rafeiro et al 24</td>
<td>Placebo</td>
<td>NA</td>
<td>12</td>
<td>43</td>
<td>59.0</td>
<td>51.2</td>
<td>7.92</td>
<td>28.6</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Linagliptin</td>
<td>5</td>
<td>12</td>
<td>435</td>
<td>58.0</td>
<td>42.3</td>
<td>7.97</td>
<td>29.7</td>
<td>6</td>
<td>Metformin</td>
</tr>
</tbody>
</table>

BMI, body mass index; HbA1c, glycated haemoglobin; NA, not applicable; SU, sulfonylurea.
Figure 2A,B depict the application of the statistical model to each individual study, demonstrating that the observed data from the studies fall mostly within the 90% prediction interval (between 5% and 95% prediction bounds), with no overall systematic overprediction or underprediction. Both change from baseline and placebo-corrected change from baseline HbA1c percentage points are presented to demonstrate longitudinal model performance for each therapy. Similarly, figure 3A–D show the 90% credible intervals at the endpoint for the linagliptin and sitagliptin change from baseline and placebo-corrected change from baseline, demonstrating accurate prediction of the effect, on average.

The simulations performed using the model show that both linagliptin 5 mg and sitagliptin 100 mg reduce HbA1c levels by 0.81% (placebo-adjusted), at week 24, when administered to patients with T2DM for 24 weeks (figure 4A,B). Credible intervals for participants without washout were 0.88 to 0.75 (linagliptin) and 0.89 to 0.73 (sitagliptin). For those who underwent washout of previous anti-hyperglycaemic therapy, the credible intervals were −0.91 to −0.76 (linagliptin) and −0.91 to −0.75 (sitagliptin). Figure 5 shows simulated differences in the true effect at 24 weeks between linagliptin 5 mg and sitagliptin 100 mg with no washout, demonstrating that the model predicted difference lies almost entirely within 0.2 percentage points, less than the previously suggested margins for non-inferiority of 0.3–0.4 percentage points.46 47

As a post hoc assessment, a t test was used to compare the HbA1c difference from placebo residuals (unexplained variations after fitting of the model) for linagliptin and sitagliptin. A p value of 0.14 was generated, suggesting no evidence of a systematic bias in favour of linagliptin by conventional thresholds (p<0.05).

**DISCUSSION**

The model developed in this study incorporates Bayesian methodology and provides a tangible approach to indirectly estimating the comparative efficacy of two compounds. The findings presented suggest that the model developed in this study provides a valid alternative approach to indirect drug comparisons. The findings of this MBMA show that linagliptin is as effective as sitagliptin in the reduction of HbA1c levels, both showing a mean, placebo-adjusted reduction of approximately 0.81% after 24-weeks’ treatment of patients with T2DM. In this study, evidence was gathered from the results of randomised, double-blind trials of sitagliptin and linagliptin. The sensitivity analyses performed in this study, using various prior distributions, support the robustness of the model. Although the use of MBMA is relatively new in the field of diabetes therapy, this method is nonetheless being increasingly recognised as an important tool in the evaluation of pharmaceutical therapies.48 49

There might be some limitations in applying the findings of the present analysis to the general population of patients with T2DM because of the selectively treated patient populations in the included trials, which included mostly white, middle-aged patients with mean baseline HbA1c <9%. The participants in the analysed trials would have been further restricted during pretrial run-in periods, which would exclude those with poor treatment adherence. Furthermore, the analysis was performed...
retrospectively, using data from different trials. As with all meta-analyses based on published data, there is a potential for publication bias. In the context of the present analysis, this potential bias pertains only to our estimates of the effects of sitagliptin, as our linagliptin data sources were not subject to publication selection. However, this is unlikely to have a substantial impact on the findings for sitagliptin, as current practice in clinical research mandates that all clinical trials are published regardless of their results and several sources were searched, including trial registries and documents used in the regulatory process.

The model includes the assumption that HbA1c levels are maintained after the full effect of treatment has been reached. This is based on observations in previous 24-week trials, where HbA1c levels have been shown to be maintained for this period,\(^{19-21}\) and the known pharmacological properties of DPP-4 inhibitors.\(^4\) The final model was adjusted for baseline HbA1c, ethnic origin and washout duration. Other covariates (concurrent medications, fraction of patients on previous oral antidiabetic drugs, BMI, age, gender, duration of T2DM) were not included in the final model because they did not show significant impact on the model parameters. Reasons for this might be either that only mean covariate values were available, or that some covariates are confounded (eg, BMI was shown to vary as a function of ethnic origin, making it difficult to isolate the independent effects of these covariates). It is important to recognise that these covariates might be of clinical importance, and their exclusion from the model could simply reflect an inability to reliably estimate the independent effect of these factors with the data available.

To date, four standard meta-analyses of the DPP-4 inhibitor class have been published, none of which has provided any results on the comparative efficacies of linagliptin and sitagliptin.\(^{16}\) These analyses confirm

---

**Figure 3** Drug effects (as glycated haemoglobin (HbA1c) percentage points) of the relevant studies at their respective endpoints. Filled dots represent observed data; horizontal lines show the 90% unconditional prediction intervals and also represent the median predicted value. (A) Linagliptin change from baseline. (B) Sitagliptin change from baseline. (C) Linagliptin difference from placebo. (D) Sitagliptin difference from placebo.
the efficacy of DPP-4 inhibitors, in terms of HbA1c reduction, and their tolerability, in particular resulting from the absence of weight gain and low risk of hypoglycaemia associated with monotherapy. The findings also indicate that therapy with DPP-4 inhibitors reduces HbA1c reductions to a similar extent to comparator drugs. Several of the limitations associated with traditional meta-analysis arise from the fact that only study end-point data are used in these analyses. For example, difficulties in selecting an appropriate summary statistic are often encountered because the treatment effect of interest varies as a function of the duration of treatment. Similarly, it might be difficult to appropriately adjust for the effect of covariates on treatment response when response is assessed at different time-points in different studies. To address the limitations of traditional meta-analysis, a general methodology has recently been proposed for the statistically valid use of MBMA. The advantage of this approach, also used in the present study, is that it enabled the synthesis of longitudinal data from multiple studies with different durations and different sampling schedules, resulting in analyses that are both more comprehensive (including a greater number of studies) and more efficient (incorporating more of the relevant data within each study) than previous methods. The unique MBMA approach in the current study also allows adjustment for covariates (eg, differences in the use of washout or racial composition in individual trials) to allow comparison of treatment response in comparable patients under similar conditions. One limitation of the study by Gibbs et al was that the MBMA used did not account for correlations across time points within treatment arms, which could lead to an overestimation of the intertrial variability in drug effect. In contrast, our approach takes account of longitudinal correlations, in accordance with previously published methods, which is a prerequisite to the correct characterisation of uncertainty in the estimation of drug effects.

Figure 4  (A) Estimated drug effects on glycated haemoglobin (HbA1c) for reference population, with no pretreatment washout, over 24 weeks (difference from placebo). (B) Estimated drug effects on HbA1c for reference population, with 4-week washout plus 2-week placebo run-in period, over 24 weeks (difference from placebo). Reference population of 1000 participants, baseline HbA1c: 8%, racial composition: 61.5% White, 1.5% Black, 37% Asian.

Figure 5  Posterior distribution for the difference in effect estimates between linagliptin (5 mg) and sitagliptin (100 mg) at 24 weeks. Reference population of 1000 participants (therefore involving 10^6 simulated patients), baseline glycated haemoglobin (HbA1c): 8%, racial composition: 61.5% White, 1.5% Black, 37% Asian.
As the clinical use of DPP-4 inhibitors increases, patients, prescribers and payers will require information on the relative benefits of the individual drugs within this class. Based on the model developed in this study, it is apparent that the efficacy of the two DPP-4 inhibitors, sitagliptin and linagliptin, is virtually indistinguishable, in terms of changes in mean HbA1c levels, in patients with T2DM treated with a range of background antihyperglycaemic therapies. Both linagliptin and sitagliptin act by inhibiting the DPP-4 enzyme that rapidly inactivates the intestinal hormone, glucagon-like-peptide (GLP)-1. GLP-1 stimulates insulin secretion in a glucose-dependent manner. Sitagliptin is largely excreted via the kidneys, with the major portion of the oral dose (87%) being excreted in the urine. Unlike sitagliptin and other DPP-4 inhibitors, linagliptin has a largely non-renal route of excretion (only ~5% excreted renally), with the majority being eliminated via the bile and gut; it therefore does not require dose adjustment in patients with declining renal function. In view of the similar efficacy of these two drugs, treatment choices might, therefore, be made on the basis of other differences between the drugs and consideration of patient characteristics, such as the patient’s renal function.

Broadening the use of MBMA has the potential to improve the comparison of individual drug therapies, compared with older statistical methods, and could provide a new way of generating results for populations that have not yet been studied.

Author affiliations
1Endocrine Division, Department of Internal Medicine, Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil
2Metrum Research Group, Tariffville, Connecticut, USA
3Boehringer Ingelheim, Biberach, Germany
4Boehringer Ingelheim, Ingelheim, Germany
5Boehringer Ingelheim, Bracknell, Berkshire, UK

Contributors All authors were fully responsible for all content and editorial decisions. They were involved at all stages of manuscript development, including reviewing and revising the manuscript for scientific content, and have approved the final version. In addition: JG contributed to the data analysis and interpretation of the findings. JR, DP and WG shared primary responsibilities for developing the statistical analysis plan, executed all statistical analyses (including model development, model selection and model summary) and interpreted the findings. SP monitored data collection, and contributed to data selection, the statistical analysis plan and interpretation of the results. CF contributed to the analysis concept, the statistical analysis plan and interpretation of the findings. BM contributed to data collection and the statistical analysis plan, and interpreted the findings. YG contributed to the interpretation of the findings. AS contributed to the analysis strategy, the statistical analysis plan and interpretation of the results. SR contributed to analysis strategy, the statistical analysis plan, data collection and interpretation of the results.

Funding Medical writing assistance, supported financially by Boehringer Ingelheim, was provided by Jennifer Edwards, MBBS, of Envision Scientific Solutions during the preparation of this article.

Competing interests All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: JG has received fees for Board membership from Boehringer Ingelheim, Novo Nordisk and Eli Lilly, and he has also received research grants from Boehringer Ingelheim, Eli Lilly, GlaxoSmithKline and Janssen. JR, DP and WG have received fees for participation in review activities, and for manuscript writing and reviewing from Boehringer Ingelheim. CF, YG, BM, SP, AS and SR are employees of Boehringer Ingelheim, the manufacturer of linagliptin.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Previous presentations Abstracts based on this study have been presented as posters at the 72nd Scientific Sessions of the American Diabetes Association, 8–12 June 2012, Philadelphia, Pennsylvania, USA, and at the Population Approach Group Europe conference, Venice, Italy, 5–8 June 2012.

REFERENCES


27. Gastonguay MR. Full covariate models as an alternative to methods relying on statistical significance for inferences about covariate precision. AAPS J 2004;6(S1):Abstract W4354.


Technical Appendix to:
A novel model-based meta-analysis to indirectly estimate the comparative efficacy of two medications: an example using DPP-4 inhibitors, sitagliptin and linagliptin, in treatment of type 2 diabetes mellitus

December 3, 2012

Jorge L Gross¹, James Rogers², Dan Polhamus², William Gillespie², Christian Friedrich³, Yan Gong⁴, Brigitta Monz⁴, Sanjay Patel⁵, Alexander Staab³, Silke Retlich³

(1) Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
(2) Metrum Research Group, Tariffville, Connecticut, USA
(3) Boehringer Ingelheim, Biberach, Germany
(4) Boehringer Ingelheim, Ingelheim, Germany
(5) Boehringer Ingelheim, Bracknell, Berkshire, UK
This appendix is intended primarily to provide a mathematical/statistical specification of the final model employed in our analysis. We additionally mention several variants of the model that were considered during model development.

1 Base model: structural form

The structural form of the model (i.e., the parametric form relating the central tendency of predictions to time, dose, and covariates) remained largely unchanged across all model-fitting iterations described in the present report. This structural form may conceptualized in terms of a latent (random-effect) baseline HbA1c ($HbA1c_{base}$) that is modified as a function of dose ($D$) and time ($t$) by multiplicative terms for the effect of washout, placebo intervention, and drug effect (where applicable).

$$HbA1c(t, D) = HbA1c_{base} \times I_{washout} \left( \frac{1 + \Delta W_{\infty} \left( 1 - e^{-k_w(t + t_{wash})} \right)}{1 + \Delta W_{\infty} \left( 1 - e^{-k_w t_{wash}} \right)} \right) \times \left( 1 + \Delta P_{\infty} \left( f^p(t) \right) \right) \times \left( 1 - E_{\text{max,drug}} D \left( 1 - e^{-k_{drug} t} \right) \right)$$

We now consider each of the model terms in slightly more detail:

- **Washout term**

  $$\left( \frac{1 + \Delta W_{\infty} \left( 1 - e^{-k_w(t + t_{wash})} \right)}{1 + \Delta W_{\infty} \left( 1 - e^{-k_w t_{wash}} \right)} \right)$$

  Although response observations from the (pre-treatment) washout period were not modeled, the duration of pre-treatment washout $t_{wash}$ is used to determine how much of the post-baseline trend can be attributed to washout. One may better understand the washout term in terms of the following two extreme scenarios:

  - The washout term for a group with an extremely long pre-treatment washout ($t_{wash} \approx \infty$) would be essentially 1, i.e., no adjustment.

  - The washout term for a group discontinuing prior medication only at the time of the first dose of the randomized treatment ($t_{wash} \approx 0$) would be \( (1 + \Delta W_{\infty} \left( 1 - e^{-k_w t} \right) ) \), a term that rises exponentially from 1 (no adjustment at baseline) toward a horizontal asymptote of \( (1 + \Delta W_{\infty} ) \) as $t$ becomes very large.

As a conceptual and notational device, we also define HbA1c$_{prior}$, the inferred HbA1c level at the beginning of the washout period:

$$HbA1c_{prior} = \frac{HbA1c_{base}}{1 + \Delta W_{\infty} \left( 1 - e^{-k_w t_{wash}} \right)}$$
Technical Appendix

• Placebo term \(1 + \Delta P_{\infty}\{f^P(t)\}\). Both exponential and a bi-exponential variants were considered for \(f^P\):

\[
\begin{align*}
\text{exponential' variant} & \quad f^P(t) = 1 - e^{-k^P t} \\
\text{bi-exponential' variant} & \quad f^P(t) = e^{-k^P t} - e^{-k^P t}
\end{align*}
\]

The placebo term is intended to characterize all aspects of the longitudinal profile that are common to all placebo-treated patients. In principal at least, this may include more than a mere “placebo effect”. For example, it might also partially reflect longitudinal changes that would occur even in the absence of placebo intervention. In contrast to the washout term, the placebo term does not vary systematically as a function of washout duration (in particular, in contrast to the washout term, the placebo “effect” does not approach a limit of zero as washout duration approaches infinity), so that the two effects are distinctly identifiable.

The exponential and bi-exponential variants of the placebo function have different implications with respect to the limiting placebo effect \(P_{\infty}\) at large time values. The bi-exponential variant, sometimes referred to as a “Bateman” function is a typical choice in the pharmacology domain for modeling placebo effects, and is applicable when placebo effects are expected to return to zero at some duration [4]. By contrast, our “exponential” variant implies a non-zero limiting effect due to placebo. In practice, the predictive implications for the fitted model may be similar for both functions over a finite duration of interest, since the estimated parameters in the Bateman function may (if supported by the data) imply a positive first derivative over most or all of that duration (in our case, over a 24 week duration). To the extent that data do not support any return to zero for the placebo effects over the duration of interest, both convergence diagnostics and model selection criteria will tend to favor the simpler “exponential” placebo function. In our application the bi-exponential / Bateman functional form appeared to be reasonably well estimated and so was therefore employed in the final model.

As a conceptual and notational device, we also define \(HbA1c_{\text{placebo}}(t)\), the expected time course for an individual randomized to placebo:

\[
HbA1c_{\text{placebo}}(t) = HbA1c_{\text{base}} \times I_{\text{washout}} \times \left(1 + \Delta P_{\infty} \left(1 + \Delta W_{\infty} \left(1 - e^{-k^W(t+twash)}\right)\right) \right) \times \left(1 + \Delta P_{\infty} \left(f^P(t)\right)\right)
\]

Further, we define \(HbA1c_{\infty}\) to be the limiting value of \(HbA1c_{\text{placebo}}(t)\) as \(t\) becomes very large.

• Drug effect term

\[
(1 - E_{\text{drug}}) \equiv \left(1 - \frac{E_{\text{max,drug}} D}{ED_{50,\text{drug}} + D} \left(1 - e^{-k_{\text{drug}} t}\right)\right)
\]

This term varies as a function of both dose (\(D\)) and time (\(t\)). The term equals 1 (no adjustment) when \(t = 0\) and/or \(D = 0\), and approaches a horizontal asymptote of \((1 - E_{\text{max,drug}})\) as both \(D\) and \(t\) become very large. The parameter \(k_{\text{drug}}\), which describes the onset of drug effects, was provisionally assumed...
to take the same value for both linagliptin and sitagliptin, consistent with the expected pharmacology of the two DPP4 inhibitors (empirical support for this assumption is provided by the posterior predictive checks included in the primary manuscript).

The dose-response component of our model may be described as a simplified $E_{\text{max}}$ model, a more general expression of which is:

$$E = \frac{E_{\text{max}}D^\alpha}{ED_{50}^\alpha + D^\alpha}.$$  

This functional form, also known as the Hill Equation, is widely used in the pharmacology domain to describe both concentration-response and dose-response relationships [5]. Originally proposed on the basis of drug receptor theory to describe concentration-response analyses, its use in describing dose-response relationships may also be theoretically justified when linear pharmacokinetics apply and dose is therefore an appropriate proxy for concentration. Theoretical support is more tenuous when nonlinear pharmacokinetics apply (as is the case for linagliptin), but the model remains a reasonable initial default. The parameter $\alpha$, sometimes referred to as the Hill coefficient or the sigmoidicity parameter, has been set to a value of 1 in our implementation. This simplification was initially introduced on a tentative basis, in consideration of the limited degree of dose-response information in the data, and based on prior experiences of the modeling team suggesting that, empirically, the Hill coefficient for dose-response relationship is generally close to one. This simplification does imply a dose-response relationship that has a non-zero gradient at $D = 0$, however this implication is consistent with observed dose-response relationships (anecdotally) for most drugs. Final acceptance of this aspect of the model was based on the observed predictive performance of our model at all studied dose levels.

We may now use the terms introduced above to define the conditional expectation for the mean HbA1c on the $i^{th}$ occasion in the $j^{th}$ group and $k^{th}$ study:

$$\overline{\text{HbA1c}}_{ijk} = \text{HbA1c}_{\text{prior},jk} \times \left(1 + \Delta W_{\infty,ijk} \left(1 - e^{-k_{w}(t_{ij}+t_{\text{wash},ijk})}\right)\right) \times \left(1 + \Delta P_{\infty,ijk} \left(f^{p}(t_{ij})\right)\right) \times \left(1 - \frac{E_{\text{max,drug},jk}D_{jk}}{ED_{50,\text{drug},jk} + D_{jk}} \left(1 - e^{-k_{\text{aug},t_{ij}}}\right)\right)$$

Note that the circumflex, or “hat”, has not been used here to indicate an estimate (as would be common in the statistical literature), but refers rather to an expected value (as is common in the pharmacometric literature).

## 2 Base model: stochastic structure

Random effects associated with washout and placebo effects were implemented using a re-parameterization of the model. The parameters $\Delta W_{\infty}$ and $\Delta P_{\infty}$ represent conceptual steady state asymptotes for very large
time values \((t \rightarrow \infty)\). Since it is not clear that direct estimation of such parameters is supported by the data, the model was implemented using truncated parameterizations that reformulate the effect parameter as an effect at a time \(t^*\) that is richly represented in the data set. For this purpose we define \(\Delta W^*\) and \(\Delta P^*\) as the fractional changes at reference time \(t^* = 24\) weeks due to washout and placebo, respectively. That is,

\[
\Delta W^* = \frac{(1 + \Delta W_{\infty} \left(1 - e^{-k_W (t^* + t_{wash})}\right))}{1 + \Delta W_{\infty} (1 - e^{-k_W t_{wash}})} - 1
\]

\[
\Delta P^* = \Delta P_{\infty} f^P(t^*).
\]

The random-effect structure for the \(i^{th}\) visit for the \(j^{th}\) arm (or group) in the \(k^{th}\) study may then be described as:

**Inter-study variation:**

\[
\log \left(1 + \Delta P_{\text{study}, k}^*\right) \sim N \left(\log \left(1 + \Delta P^*\right), \psi_{\Delta p}^2\right)
\]

Study-level random effects are additionally considered for mean baseline HbA1c and for the washout magnitude \(\Delta W^*\), but convergence diagnostics suggested that this level of variation in these parameters was not identifiable based on the available data. All study-level random effects are assumed to be independent.

**Inter-arm variation:**

\[
\log \left(1 + \Delta P_{jk}^*\right) \sim N \left(\log \left(1 + \Delta P_{\text{study}, k}^*\right), \omega_{\Delta p}^2 / n_{1jk}\right)
\]

\[
\log \left(\text{HbA1c}_{\text{base},jk}\right) \sim N \left(\log \left(\text{HbA1c}_{\text{base-study}, k}\right), \omega_{\text{HbA1c}_{\text{base}}}^2 / n_{1jk}\right)
\]

All arm-level random effects are assumed to be conditionally independent, given study-level random effects. Our scaling of random effect variances according to sample size follows a recently published rationale \([1]\), and is similar to an approach used in model-based meta-analysis of Alzheimer’s Disease progression \([3]\). Arm-level random effects were additionally considered for the washout magnitude \(\Delta W^*\), but convergence diagnostics suggested that this level of variation was also not identifiable based on the available data.

Modeling baseline HbA1c values using Lognormal distributions constrains these to positive values. Modeling the \(\Delta\) parameters using shifted Lognormal distributions allows the placebo and washout factors in the model to take positive and negative (up to negative one) values. Taken together with constraints on the drug effects (ensured via the prior for the drug \(E_{\text{max}}\) values), and the multiplicative construction of the model, the preceding constraints ensure that all HbA1c predicted values are positive.

**Residual variation:**

\[
\text{HbA1c}_{ijk} \sim N \left(\overline{\text{HbA1c}_{ijk}}, \frac{\sigma^2}{n_{ijk}}\right).
\]
3 Covariates

Baseline HbA1c
Baseline HbA1c is implicitly included as a covariate in the model in the sense that the washout, placebo, and drug effect terms all operate multiplicatively on baseline HbA1c. In general, larger changes from baseline are implied by larger baseline values.

Washout
The inclusion of the washout term in the base model implies that washout status and washout duration act as covariates in the sense that they modify an individual's predicted change from baseline.

Other covariates

Whereas the covariate effects of baseline HbA1c and washout duration are implied by the longitudinal structure of the base model, the remaining candidate covariates were considered as potential modifiers of drug effect in a manner dependent on the covariate variable's distributional properties.

- For univariate covariates taking discrete values at the arm-level (as was the case with background medication status), the covariate effect was introduced via an exponentiated linear predictor so that, with \( x_{jk} \) denoting the dichotomous variable level on arm \( j \) in study \( k \), the following substitution is made:

\[
E_{\text{max,drug}}^{\text{drug,}jk} \leftarrow E_{\text{max,drug}}^{\text{drug,}jk} \times \exp \{ \beta_{\text{drug,}jk} x_{jk} \}
\]

(This approach is equivalent, modulo re-parameterization, to simply using a different \( E_{\text{max,drug}} \) for each level of the covariate.)

- For univariate covariates taking continuous values at the arm-level including age, body mass index (BMI), gender (recall that at the arm-level, gender is represented as a proportion), and duration of diabetes, the candidate covariate was centered at its average value and an exponentiated linear predictor was again used. In this case, letting \( x_{jk} \) be the continuous covariate, the following substitution is made:

\[
E_{\text{max,drug}}^{\text{drug,}jk} \leftarrow E_{\text{max,drug}}^{\text{drug,}jk} \times \exp \{ \beta_{\text{drug,}jk} (x_{jk} - \bar{x}) \}
\]

- Race, like gender, is represented at the arm-level using proportions. However, unlike gender, the racial composition of a each study arm must be represented with multiple proportions (specifically, proportion white, proportion black, proportion asian, and, in certain cases, proportion “other”), precluding any interpretable use of the continuous covariate parameterization that was used for gender. Instead, a separate \( E_{\text{max,drug}}^{r,\text{drug,}jk} \) was defined for each race \( r \) (separately for each drug). Race-specific terms in the model then imply race-specific conditional expectations \( \overline{\text{HbA1C}}_{ijk}^r \), and the aggregate conditional expectation is computed using a weighted average:

\[
\overline{\text{HbA1C}}_{ijk} = \sum_{r \in \text{white, black, asian, other}} F_{ijk}^r \overline{\text{HbA1C}}_{ijk}^r
\]

where \( F^r \) is the fraction of the arm identifying with race \( r \).

In order to improve parameter estimation for race-based covariate effects, racially subsetted data were used where possible (this was possible in all cases for linagliptin records).
As noted in the primary manuscript, the effects of covariates other than race were not sufficiently well estimated to justify inclusion in the model. Race was therefore the only explicit in the final model (in addition to the implicit covariates, baseline HbA1c and washout duration).

4 Priors

The probabilistic specification of the prior is provided in Table 1. With regard to the primary research questions of interest, the most consequential components of the prior are the distributional statements relating to the magnitudes of drug effects, $E_{\text{max, linagliptin}}$ and $E_{\text{max, sitagliptin}}^*$. Since those parameters represent asymptotic drug effects (at theoretically infinite times), we re-parameterize in terms of effects at 24 weeks, $E_{\text{linagliptin}}^*$ and $E_{\text{sitagliptin}}^*$, using the same reasoning as was discussed for placebo and washout effects. Since these drug effect parameters represent fractional reductions from a hypothetical untreated state, it is natural that they should be bounded between zero and one, implying that both drugs have some beneficial effect (a defensible assumption for marketed drugs), neither or which may reduce HbA1c levels below zero (patently true). The use of Uniform (flat) densities between these two extremes implies that all intermediate values are considered (a priori) equally likely. Perhaps most importantly, the distributions for $E_{\text{linagliptin}}^*$ and $E_{\text{sitagliptin}}^*$ are specified as being independent. In combination with the use of Uniform densities, the prior independence of these two parameters allows for potential findings of either substantial similarity or substantial difference between the two drugs, as determined by the data.

Priors for ancillary (non-drug-effect) parameters were chosen largely for analytical convenience, but were verified to be diffuse in comparison to their corresponding posterior distributions, suggesting that these elements of the prior were sufficiently non-informative. Additionally, alternative distributions were evaluated for a number of elements of the prior (e.g. for variance components), and did not result in any substantial differences in conclusions.
Table 1: Prior distribution for parameters in base model

<table>
<thead>
<tr>
<th>parameter</th>
<th>prior</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_W$ (w$^{-1}$)</td>
<td>Unif(0.001,7)</td>
</tr>
<tr>
<td>$k_{p0}$ (w$^{-1}$)</td>
<td>Unif(0.001,7)</td>
</tr>
<tr>
<td>$k_P$ (w$^{-1}$)</td>
<td>Unif(0.001,7)</td>
</tr>
<tr>
<td>$k_{drug}$ (w$^{-1}$)</td>
<td>Unif(0.001,7)</td>
</tr>
<tr>
<td>HbA1c$_{base}$</td>
<td>Log Normal(0, 100)</td>
</tr>
<tr>
<td>$\Delta W$</td>
<td>Unif(−1, 10)</td>
</tr>
<tr>
<td>$\log \left( 1 + \Delta P^p \right)$</td>
<td>Unif(−10,5)</td>
</tr>
<tr>
<td>$\Psi^p$</td>
<td>Unif(0,10)</td>
</tr>
<tr>
<td>$\omega_{HbA1c_0}$</td>
<td>Unif(0,5)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Unif(0,100)</td>
</tr>
<tr>
<td>$b_{\text{linagliptin}}$</td>
<td>Unif(0,100)</td>
</tr>
<tr>
<td>$b_{\text{sitagliptin}}$</td>
<td>Unif(0,100)</td>
</tr>
<tr>
<td>$E^*_\text{linagliptin}$</td>
<td>Unif(0,1)</td>
</tr>
<tr>
<td>$E^*_\text{sitagliptin}$</td>
<td>Unif(0,1)</td>
</tr>
<tr>
<td>$E_{\text{max,linagliptin}}$</td>
<td>$(\frac{1}{b_{\text{linagliptin}}} + 1) \times E^*_\text{linagliptin}$</td>
</tr>
<tr>
<td>$E_{\text{max,sitagliptin}}$</td>
<td>$(\frac{1}{b_{\text{sitagliptin}}} + 1) \times E^*_\text{sitagliptin}$</td>
</tr>
<tr>
<td>$ED_{50,\text{linagliptin}}$</td>
<td>$\frac{b_{\text{linagliptin}}}{E_{\text{max,linagliptin}}} \times 200$</td>
</tr>
<tr>
<td>$ED_{50,\text{sitagliptin}}$</td>
<td>$\frac{b_{\text{sitagliptin}}}{E_{\text{max,sitagliptin}}}$</td>
</tr>
</tbody>
</table>

References


**Sitagliptin** records identified through search of FDA Drug Approval Package, Cochrane Review, clinical trials.gov registry, and manual searching (n = 48)

**Linagliptin** records identified in sponsor’s library (n = 10)

**Sitagliptin and linagliptin** records identified in Embase searching, Cochrane library, sponsor’s library, clinical trials.gov registry, Australian and New Zealand Clinical Trial registry, and manual searching (n = 1008)

Records after duplicates removed (n = 41)

Records excluded (n = 13)

Records screened (n = 45)

Records excluded (n = 992)

Records screened (n = 16)

Full-text articles assessed for eligibility (n = 41)

Full-text articles excluded, with reasons (n = 16)

Studies included in qualitative synthesis (n = 25)

Studies included in quantitative synthesis (meta-analysis) (n = 25)
Suggested Additional Information

Table 1 Details of search strategies

Table 2 Summary of excluded references and reasons for their exclusion
## Supplementary Table 1

### Details of search strategies

<table>
<thead>
<tr>
<th>Database (search date)</th>
<th>Search string</th>
<th>Citations identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embase (8 November 2010)</td>
<td>1. linagliptin OR 'ondero'/exp OR ondero OR 'bi-1356'/exp OR 'bi- 1356' OR 'bi1356'/exp OR bi1356 OR 'bi 1356'/exp OR 'bi 1356'</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>2. 'sitagliptin'/exp OR sitagliptin OR 'januvia'/exp OR januvia OR 'sitagliptin'/exp OR sitagliptine OR 'mk 0431'/exp OR 'mk 0431' OR 'km0431'/exp OR mk0431 OR 'mk431'/exp OR mk431 OR 'mk 431'/exp OR 'mk 431'</td>
<td>1545</td>
</tr>
<tr>
<td></td>
<td>3. Search 1 OR 2</td>
<td>1582</td>
</tr>
<tr>
<td></td>
<td>4. 'diabetes'/exp OR diabetes OR 'diabetic'/exp OR diabetic</td>
<td>526 269</td>
</tr>
<tr>
<td></td>
<td>5. Search 3 AND 4</td>
<td>1450</td>
</tr>
<tr>
<td></td>
<td>6. 'comparative study'/exp OR 'comparative study' OR 'clinical trial'/exp OR 'clinical trial' OR 'randomised controlled trial'/exp OR 'randomisation'/exp OR 'single blind procedure'/exp OR 'single blind procedure' OR 'double blind procedure'/exp OR 'double blind procedure' OR 'triple blind procedure' OR 'crossover procedure'/exp OR 'crossover procedure' OR 'placebo'/exp OR 'placebo' OR 'random' OR rct OR 'single blind' OR 'single blinded' OR 'double blind' OR 'double blinded' OR 'treble blind' OR 'treble blinded' OR 'triple blind' OR 'triple blinded' OR 'prospective study'/exp OR 'prospective study'</td>
<td>2 215 299</td>
</tr>
<tr>
<td>Search</td>
<td>Query</td>
<td>Results</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>7.</td>
<td>Search 5 AND 6</td>
<td>874</td>
</tr>
<tr>
<td>8.</td>
<td>Search 7 AND [humans]/lim</td>
<td>807</td>
</tr>
<tr>
<td>Cochrane library&lt;br&gt;(8 November 2010)</td>
<td>(linagliptin OR ondero OR sitagliptin OR januvia) AND diabetes</td>
<td>48</td>
</tr>
<tr>
<td>IDEA&lt;br&gt;(10 November 2010)</td>
<td>Linagliptin OR ondero OR sitagliptin OR januvia</td>
<td>23</td>
</tr>
<tr>
<td>Manual searching</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1008</td>
</tr>
</tbody>
</table>
Supplementary Table 2

Summary of excluded references and reasons for their exclusion

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaboe 2010</td>
<td>&lt;20 patients per arm and/or crossover design</td>
</tr>
<tr>
<td>Bragz 2007</td>
<td>&lt;20 patients per arm and/or crossover design</td>
</tr>
<tr>
<td>Chan 2008</td>
<td>Renal insufficiency population</td>
</tr>
<tr>
<td>Herman 2006</td>
<td>&lt;20 patients per arm and/or crossover design</td>
</tr>
<tr>
<td>Nauck 2007</td>
<td>Seck et al. 2010 is extension study to Nauck et al. 2007, as the Seck et al. 2010 article reports the results of the full analysis dataset (in addition to those of the per-protocol dataset) whereas in Nauck et al. 2007 only those of the per-protocol group are given, only data that referred to the full analysis dataset reported in Seck et al. 2010 were used</td>
</tr>
<tr>
<td>Nonaka 2009</td>
<td>Only 4-week treatment duration</td>
</tr>
<tr>
<td>Prately 2010</td>
<td>Open-label design</td>
</tr>
<tr>
<td>Raz 2008</td>
<td>Phase IV study in poorly controlled subjects</td>
</tr>
<tr>
<td>Retnakararn 2010</td>
<td>&lt;20 patients per arm and/or crossover design</td>
</tr>
<tr>
<td>Rigby 2010</td>
<td>Open-label design</td>
</tr>
<tr>
<td>Williams-Herman 2009</td>
<td>These extension studies included only those patients from the previous study that had not required rescue medication (introducing a likely selection bias)</td>
</tr>
<tr>
<td>Williams-Herman 2010</td>
<td></td>
</tr>
<tr>
<td>Merck Study Code P015</td>
<td>Lack of suitable data</td>
</tr>
</tbody>
</table>
References


http://www.accessdata.fda.gov/drugsatfda_docs/nda/2006/021995s000TOC.cfm