An indirect comparison of HbA1c treatment effect with albiglutide and exenatide 2.0 mg QW using the Bucher method

Alan A Martin1
Daniel Parks2
1Department of Value Evidence Analytics, GlaxoSmithKline, Uxbridge, Middlesex, UK; 2Department of Value Evidence Analytics, GlaxoSmithKline, Collegeville, PA, USA

Abstract: No head-to-head comparisons exist between once-weekly (QW) glucagon-like peptide-1 receptor agonists; accordingly, this indirect comparison was conducted to evaluate the comparative efficacy of QW albiglutide vs QW exenatide. Following a systematic literature search, it was determined that HARMONY 7 and DURATION 6, Phase III trials for albiglutide and exenatide, respectively, were similar in study design and baseline characteristics and included a common comparator arm, making them suitable for an indirect comparison using the Bucher method. The primary endpoint of change from baseline in glycated hemoglobin (HbA1c) with albiglutide 50 mg QW and exenatide 2.0 mg QW was compared and tested for noninferiority. The indirect comparison showed a treatment difference of 0.0% (95% confidence interval: −0.189% to 0.189%) in mean change in HbA1c from baseline, and albiglutide 50 mg was noninferior to exenatide 2.0 mg QW at the noninferiority margin of 0.3%. In the absence of a head-to-head trial, these results can be used in pharmacoeconomic analysis and to inform health technology assessment and clinical decision making.

Keywords: albiglutide, exenatide 2.0 mg QW GLP-1 RA, Bucher method

Introduction
With the expanding number of classes and agents for the treatment of type 2 diabetes mellitus (T2DM), as well as the focus on patient-centered care,1 it is becoming increasingly important to understand the parameters for optimal clinical use of these agents. Albiglutide is a once-weekly (QW) glucagon-like peptide-1 receptor agonist (GLP-1 RA) recently approved for the treatment of T2DM. The GLP-1 RA class includes both weekly and daily administered injectable products; thus, frequency of administration may be a factor in the choice of a GLP-1 RA. If a QW GLP-1 RA is preferred, then the relative efficacy of the weekly GLP-1s currently approved for clinical use (albiglutide, exenatide 2.0 mg QW and dulaglutide2–4) is of interest. Because albiglutide was the second QW GLP-1 RA to achieve marketing authorization, comparative efficacy against exenatide 2.0 mg QW (the first weekly approved GLP-1 RA) is important for health technology appraisal.5 The outcomes of interest, in a pharmacoeconomic analysis of T2DM treatments, are long-term diabetes complications. These must be modeled from effects observed in trials; the key treatment effect being modification of blood glucose measured as glycated hemoglobin (HbA1c). In the absence of head-to-head data, the purpose of this analysis was to provide an indirect comparison of the efficacy of these two agents on HbA1c lowering.
Methods
This analysis was a single-step, indirect comparison of change from baseline in HbA1c with albiglutide 50 mg QW (the highest approved dose) vs exenatide 2.0 mg QW. Liraglutide 1.8 mg once daily (QD) served as the common comparator. The indirect comparison was made using the Bucher method and included data from two studies: HARMONY 7 and DURATION 6 (Table 1). The primary endpoints (change from baseline in HbA1c after 26 weeks, respectively, for albiglutide and exenatide 2.0 mg QW) of each study were compared.

Potential studies for inclusion in an indirect comparison were identified from a systematic literature search of EMBASE, MEDLINE, MEDLINE in Process, and Cochrane library electronic databases performed on January 8, 2013, and then updated on May 12, 2014. Phase III randomized controlled trials in adults with type 2 diabetes that lasted for 24 weeks or more, with treatment arms for albiglutide, exenatide 2.0 mg QW and other available GLP-1 RAs (exenatide BID [twice daily], liraglutide), sitagliptin, insulin lispro, insulin glargine and placebo were searched for, with no limitation on publication language or year of publication. The search was conducted by two independent analysts, and only full publications and clinical study reports for albiglutide Phase III trials were included. Following de-duplication, a total of 2,078 records were identified for abstract screening, and ultimately 51 full-text papers and eight albiglutide clinical study reports for unique studies were identified for assessment; this included eight albiglutide studies (references refer to publications subsequent to this analysis for some HARMONY studies) and eight exenatide 2.0 mg QW studies.

Studies were assessed as suitable for inclusion in an indirect comparison if they fulfilled the condition of similarity of study design and populations, and other factors which could be modifiers of relative treatment effect and thus where differences could bias results. Of the eight albiglutide studies, five (HARMONY 1–5) were substantially different in design to exenatide 2.0 mg QW studies: much longer total study duration (3 years vs 24–30 weeks for exenatide 2.0 mg QW studies), primary endpoints evaluated at 52 weeks or 104 weeks (vs 24–30 weeks for exenatide 2.0 mg QW studies), the use of optional uptitration of albiglutide from 30 mg to 50 mg in four of the studies, and an approach to subject inclusion and hyperglycemic rescue designed to be more real-world. HARMONY 6 compared albiglutide to insulin lispro, which is not suitable as a common comparator due to variability in insulin lispro formulation and regimen. HARMONY 8 was conducted in patients with renal impairment, and no exenatide 2.0 mg QW study was conducted in a similar population. Consequently, these were excluded due to study heterogeneity. HARMONY 7, however, which compared QW albiglutide to liraglutide 1.8 mg QD, was similar in design and study population to several exenatide 2.0 mg QW studies. The inclusion of only the HARMONY 7 albiglutide study meant that constructing a meaningful network was not possible; thus, studies with other potential intermediate comparators were not included. DURATION 6 was the only exenatide 2.0 mg QW study with a common comparator, liraglutide 1.8 mg QD, so HARMONY 7 and DURATION 6 were ultimately included in the indirect comparison.

With only two studies, and therefore no networks with a closed loop, the Bucher method was selected for the analysis.

Table 1. Comparative summary description of HARMONY 7 and DURATION 6

<table>
<thead>
<tr>
<th>Test</th>
<th>Noninferiority test</th>
<th>Noninferiority test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>Open-label, randomized, parallel-group Phase III study 32 weeks’ duration</td>
<td>Open-label, randomized, parallel-group Phase III study 26 weeks’ duration</td>
</tr>
<tr>
<td>Treatment arms</td>
<td>Albiglutide 30 mg QW titrated to 50 mg at week 6</td>
<td>Exenatide 2.0 mg QW</td>
</tr>
<tr>
<td></td>
<td>Liraglutide 0.6 mg once daily titrated to 1.2 mg at week 1 and 1.8 mg at week 2</td>
<td>Liraglutide 0.6 mg once daily titrated to 1.2 mg at week 1 and 1.8 mg at week 2 (could be delayed if severe nausea or vomiting)</td>
</tr>
<tr>
<td>Primary endpoint</td>
<td>Change from baseline in HbA1c at 32 weeks; ITT population, LOCF</td>
<td>Change from baseline in HbA1c at 26 weeks; ITT population, LOCF</td>
</tr>
<tr>
<td>Population</td>
<td>Uncontrolled (HbA1c ≥7.0% and ≤10.0%) on oral therapy</td>
<td>Uncontrolled (HbA1c ≥7.1% and ≤11.0%) on oral therapy</td>
</tr>
<tr>
<td></td>
<td>Background oral therapy: metformin, thiazolidinediones, sulfonylureas, or any combination of these drugs</td>
<td>Background oral therapy: metformin, sulfonylurea, metformin + sulfonylurea, or metformin + pioglitazone</td>
</tr>
<tr>
<td></td>
<td>BMI ≥20 kg/m² and ≤45 kg/m²</td>
<td>BMI ≤45 kg/m²</td>
</tr>
</tbody>
</table>

Generally similar exclusions for comorbid conditions

Abbreviations: BMI, body mass index; HbA1c, glycated hemoglobin; ITT, intent to treat; LOCF, last observation carried forward; QW, once weekly.
Calculations were carried out using SAS (SAS Institute, Inc., Cary, NC, USA). The indirect treatment difference in mean change from baseline in HbA\textsubscript{1c} was calculated; noninferiority of albiglutide 50 mg vs exenatide 2.0 mg QW was tested. The noninferiority margin of 0.3% was chosen to be consistent with the HARMONY 7 study and was based on the US Food and Drug Administration guidance for industry.\textsuperscript{23}

### Results

#### Baseline characteristics

Baseline characteristics of patients in HARMONY 7 and DURATION 6 were similar, with no obvious differences between the two studies that could bias the results of the indirect comparison (Table 2).

#### Indirect comparison of albiglutide vs exenatide 2.0 mg QW on HbA\textsubscript{1c}

Using the Bucher method, albiglutide 50 mg demonstrated noninferiority to exenatide 2.0 mg QW with regard to HbA\textsubscript{1c} lowering based on a difference of 0.0% with a 95% confidence interval of −0.189% to 0.189% at the noninferiority margin of 0.3% (\(P=0.002\); Table 3, Figure 1).

#### Discussion

Exenatide 2.0 mg QW was not yet approved when the HARMONY program was initiated. Thus, no head-to-head study comparing albiglutide and exenatide 2.0 mg QW exists. Based on the timing of approvals (exenatide 2.0 mg QW was the first QW GLP-1 RA to receive marketing authorization, and albiglutide was the second), comparative efficacy data vs the compound most likely to be displaced is an important component for reimbursement in countries where health technology appraisal is conducted. In the absence of head-to-head data, this indirect comparison was conducted using the Bucher method and demonstrated that albiglutide is noninferior to exenatide 2.0 mg QW with respect to HbA\textsubscript{1c} reduction, with a noninferiority margin of 0.3%. Even if a more stringent noninferiority margin of 0.25% had been used, noninferiority would still have been achieved.

The similarity assumption is central to indirect comparison\textsuperscript{24} and requires that the inclusion criteria and baseline characteristics be similar across the two trials, that the liraglutide titration schedule be similar, that missing data be unrelated to treatment efficacy, and that effect-modifying factors exhibit a similar distribution across studies. Based on the similarity of the inclusion criteria and baseline characteristics and the fact that the majority of patients who withdraw from GLP-1 RA studies do so for reasons unrelated to efficacy, this assumption has been met. Because only two studies were compared, the heterogeneity assumption and the consistency assumption could not be checked statistically.

### Table 2 Baseline characteristics of patients enrolled in HARMONY 7 and DURATION 6

<table>
<thead>
<tr>
<th>Baseline characteristics; data are mean (SD) or %</th>
<th>HARMONY 7\textsuperscript{7}</th>
<th>DURATION 6\textsuperscript{8}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>Albiglutide 50 mg (N=402)</td>
<td>Exenatide 2.0 mg QW (N=390)</td>
</tr>
<tr>
<td></td>
<td>55.4 (10)</td>
<td>57.9 (9.4)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>HbA\textsubscript{1c} (%)</td>
<td>8.18 (0.89)</td>
<td>8.15 (0.84)</td>
</tr>
<tr>
<td>BMI (kg/m2)</td>
<td>32.8 (6.0)</td>
<td>32.3 (5.6)</td>
</tr>
<tr>
<td>Duration of diabetes (years)</td>
<td>8.4 (6.1)</td>
<td>8 (6)</td>
</tr>
</tbody>
</table>

**Abbreviations:** BMI, body mass index; HbA\textsubscript{1c}, glycated hemoglobin; QW, once weekly; SD, standard deviation.

### Table 3 Results of indirect comparison analysis using Bucher method

<table>
<thead>
<tr>
<th></th>
<th>HARMONY 7\textsuperscript{7}</th>
<th>DURATION 6\textsuperscript{8}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Albiglutide 50 mg (N=402)</td>
<td>Exenatide 2.0 mg QW (N=390)</td>
</tr>
<tr>
<td></td>
<td>Liraglutide 1.8 mg (N=403)</td>
<td>Liraglutide 1.8 mg (N=386)</td>
</tr>
<tr>
<td>HbA\textsubscript{1c}</td>
<td>LS mean difference from BL (SE)</td>
<td>−0.78 (0.047)</td>
</tr>
<tr>
<td></td>
<td>Treatment difference (95% CI) (%)</td>
<td>−0.99 (0.046)</td>
</tr>
<tr>
<td>Noninferiority</td>
<td>Mean difference (exenatide 2.0 mg QW albiglutide): 0.0</td>
<td>0.21 (0.08–0.34)</td>
</tr>
<tr>
<td></td>
<td>SE: 0.097</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>95% CI of difference (exenatide 2.0 mg QW albiglutide): (−0.189 to 0.189)</td>
<td>(−0.189 to 0.189)</td>
</tr>
<tr>
<td></td>
<td>Albiglutide 50 mg is noninferior to exenatide 2.0 mg QW (P=0.002)</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** BL, baseline; CI, confidence interval; HbA\textsubscript{1c}, glycated hemoglobin; LS, least squares; QW, once weekly; SE, standard error.
Other factors deserve consideration as potential sources of bias: these are the increasing HbA1c trajectory observed with both exenatide and liraglutide at 26 weeks in DURATION 6, the difference in timing of endpoints (26 weeks in DURATION 6 vs 32 weeks in HARMONY 7), and the difference in HbA1c reduction observed with liraglutide 1.8 mg in DURATION 6 and HARMONY 7. Although increasing, HbA1c trajectories in DURATION 6 are increasing for both exenatide 2.0 mg QW and liraglutide, but not converging Buse et al,8 and thus treatment difference is not changing. The Bucher analysis is based on treatment difference, not on absolute effect, therefore the indirect comparison should not be affected. In HARMONY 7, the trajectory of HbA1c for albiglutide and liraglutide 26–32 weeks is stable Pratley et al,7 so comparisons at 26 weeks and 32 weeks would give similar results. The difference in HbA1c reduction observed with liraglutide in HARMONY 7 and DURATION 6 is not out of line with the range in magnitude, from 1.0% to 1.5%, seen with liraglutide 1.8 mg in other Phase III studies.25–26 Reductions in HbA1c can vary across studies for reasons other than differences in population, such as patient behavior with respect to diet and exercise and study design factors such as structure and duration of run-in periods, and also simply due to chance. Because the Bucher method adjusts for differences in absolute effect size and the two studies satisfied the similarity condition, this was not seen as a significant source of bias.

This analysis assumes that HARMONY 7 and DURATION 6 reflect the real treatment difference between albiglutide or exenatide 2.0 mg QW and liraglutide 1.8 mg, and it is not known how the inclusion of other indirect evidence would have affected the results. Scott et al, in their network meta-analysis (NMA) comparing exenatide 2.0 mg QW and liraglutide,27 also included DURATION 6. However, although their models adjust for some sources of heterogeneity, significant inconsistencies were still present between the direct and indirect evidence for the comparison of exenatide 2.0 mg QW and liraglutide 1.8 mg and liraglutide 1.8 mg and placebo. As described in the National Institute of Health and Care Excellence Decision Support Unit (NICE DSU) Technical Support Document 4,28 when inconsistency is present after adjustment for heterogeneity, it is advisable to reconsider the entire network, reviewing studies for effect modifiers that could be the source of the inconsistency and potentially excluding them. Scott et al considered that the inconsistency could be resolved by removing DURATION 6, but could not identify a reason, based on study and patient characteristics, for doing so. However, they do not report whether removing other studies with, for instance, different background therapies, different populations (eg, Asian), or different study duration also resolved the inconsistency. Judgment must be applied on what studies to include or exclude from the network, as there is no statistical method for doing so. In this situation, it is reasonable to assume

---

**Figure 1** Forest plot of albiglutide vs exenatide 2.0 mg QW.

**Abbreviations:** QW, once weekly; LCL, 95% lower confidence limit; UCL, 95% upper confidence limit.
that the direct evidence, ie, DURATION 6, reflects the real treatment difference between liraglutide 1.8 mg and exenatide 2.0 mg QW and indeed the study found that it was probable that liraglutide 1.8 mg reduces HbA\textsubscript{1c} more than exenatide 2.0 mg QW.

NMA and mixed treatment comparison (MTC) methods, with either frequentist or Bayesian approaches, are also available for conducting indirect comparisons, but were not used here. The key consideration was that only two studies (DURATION 6 and HARMONY 7) satisfied the similarity assumption (a fundamental requirement for NMA/MTC), placing constraints on methodology, as NMA and MTC are model-based approaches requiring a greater number of studies. For this reason, the Bucher method for adjusted indirect treatment comparison was chosen. Bucher is a frequentist method, and although a Bayesian approach would not normally be adopted with only two studies included, as an additional check a Bayesian analysis was also performed. In general, a frequentist (eg, Bucher) and a Bayesian approach should yield similar, although not necessarily identical, results – in this case, the results for the Bucher and Bayesian analyses were similar and conclusions were the same.

In addition to the limited number of studies suitable, based on study design, for inclusion in the indirect comparison, other limitations should be acknowledged. First, only the primary measure of efficacy (HbA\textsubscript{1c}) was compared. Other endpoints were considered for inclusion in the analysis, but because the primary goal was to provide a basis for quantitative health economic comparison, it was decided to focus on HbA\textsubscript{1c}, an approach also adopted in the Scott study. Second, the analysis did not include dulaglutide, another QW GLP-1 RA, because it was not available at the time of this analysis.

**Conclusion**

In conclusion, this indirect comparison demonstrates the noninferiority of albiglutide to exenatide 2.0 mg QW with respect to HbA\textsubscript{1c} lowering. This finding, in the absence of head-to-head data, can be used in pharmaco-economic analysis and to inform health technology assessment and clinical decision making.

**Acknowledgments**

This study was funded by GlaxoSmithKline. Editorial support was provided by Joelle Suchy, PhD, and William Ho, PhD, MediTech Media, and funded by GlaxoSmithKline. Medical and clinical advice was provided by Dr Veronica Bainbridge, an employee of GlaxoSmithKline.

**Disclosure**

Daniel Parks and Alan Martin are employees of GlaxoSmithKline. The authors report no other conflicts of interest in this work.

**References**


