Received: 24 January 2022 | Revised: 10 July 2022 | Accepted: 24 July 2022

DOI: 10.1002/psp4.12854

ARTICLE

Optimization of trial duration to predict long-term HbA1c change with therapy: A pharmacometrics simulation-based evaluation

Hanna Kunina[1] | **Alex Al-Mashat**[1] | **Jenny Y. Chien**[2] | **Parag Garhyan**[2] | **Maria C. Kjellsson**[1]

[1]Pharmacometrics Research Group, Department of Pharmacy, Uppsala University, Uppsala, Sweden

[2]Global Pharmacokinetics/Pharmacodynamics and Pharmacometrics, Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, Indiana, USA

Correspondence
Maria C. Kjellsson, Department of Pharmacy, Uppsala University, Box 580. 751 23 Uppsala, Sweden.
Email: maria.kjellsson@farmaci.uu.se

Abstract

Glycated hemoglobin (HbA1c) is the main biomarker of diabetes drug development. However, because of its delayed turnover, trial duration is rarely shorter than 12 weeks, and being able to predict long-term HbA1c with precision using data from shorter studies would be beneficial. The feasibility of reducing study duration was therefore investigated in this study, assuming a model-based analysis. The aim was to investigate the predictive performance of 24- and 52-week extrapolations using data from up to 4, 6, 8 or 12 weeks, with six previously published pharmacometric models of HbA1c. Predictive performance was assessed through simulation-based dose–response predictions and model averaging (MA) with two hypothetical drugs. Results were consistent across the methods of assessment, with MA supporting the results derived from the model-based framework. The models using mean plasma glucose (MPG) or nonlinear fasting plasma glucose (FPG) effect, driving the HbA1c formation, showed good predictive performance despite a reduced study duration. The models, using the linear effect of FPG to drive the HbA1c formation, were sensitive to the limited amount of data in the shorter studies. The MA with bootstrap demonstrated strongly that a 4-week study duration is insufficient for precise predictions of all models. Our findings suggest that if data are analyzed with a pharmacometric model with MPG or FPG with a nonlinear effect to drive HbA1c formation, a study duration of 8 weeks is sufficient with maintained accuracy and precision of dose–response predictions.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

In diabetes drug development, the study duration in phase II is typically 12 weeks and related to achieving a steady state of the biomarker glycated hemoglobin (HbA1c). Several pharmacometric HbA1c models have been published that can predict steady-state HbA1c from studies in which HbA1c has not yet achieved steady state, for example, from phase I or early proof-of-concept studies.

WHAT QUESTION DID THIS STUDY ADDRESS?

We investigated and compared the predictive performance of previously published HbA1c models, with a key focus on their extrapolation properties with study durations ranging from 4 to 12 weeks.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

In the investigated scenarios, a study duration of 8 weeks is in general sufficient for accurate and precise long-term HbA1c predictions when using pharmacometric models, which dynamically predict HbA1c using mean plasma glucose or fasting plasma glucose nonlinearly as the driver of HbA1c formation.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

The results of the current study may guide the model choice based on the available biomarkers and study duration as well as suggest the opportunity for reduction of clinical study duration leveraging a model-based approach.

INTRODUCTION

The approval of sodium-glucose cotransporter 2 inhibitors and glucagon-like peptide 1 analogs signifies a major breakthrough in diabetes drug development.[1,2] However, maintaining recommended glycemic goals over time is challenging, which necessitates continued research to identify novel antidiabetic drugs and hence diabetes drug development. A search in May 2021 of industry-funded, diabetes interventional studies in clinicaltrials.gov revealed that 19%, 36%, and 65% of the studies listed glycated hemoglobin (HbA1c) as the primary outcome measure, in phase I, phase II, and phase III, respectively,[3] and that HbA1c was also included as a secondary outcome in almost all studies. Most of the studies with HbA1c as the primary outcome were designed to compare baseline-corrected HbA1c (ΔHbA1c) between two or more arms after 12/16 weeks or 24/52 weeks of treatment for phase II and phase III, respectively. The relatively long study duration in phase II is associated with the slow turnover of red blood cells (RBC), consequently a slow effect onset of HbA1c with changes in glucose exposure.

HbA1c is formed as glucose binds to hemoglobin (Hb) in RBC and is measured as the fraction of glycated to total Hb. Considering that the RBC life span is approximately 120 days, HbA1c reflects the glucose exposure during that time period, providing a reliable long-term assessment.[4,5] Consequently, it takes roughly 12–16 weeks for treatment, with an immediate effect on glucose, to achieve steady state of HbA1c. With such a delayed pharmacodynamic effect, models may be useful for earlier assessment of the achievable HbA1c steady state.

Several models of HbA1c have been published using various drivers of HbA1c formation (fasting plasma glucose [FPG], daily mean plasma glucose [MPG], fasting serum insulin [FSI]).[6] Nathan et al.[7] presented an empirical, steady-state regression between MPG and HbA1c useful to predict HbA1c from steady-state MPG. However, this regression cannot handle dynamic predictions of HbA1c or differences in variability between glucose and HbA1c.

Pharmacometric models handle variability well and can predict HbA1c dynamically. Among the pharmacometric models, there is the empirical A Dynamic HbA1c Endpoint Prediction Tool (ADOPT) model[8] and the semiphysiological Integrated Glucose-RBC-HbA1c (IGRH) model,[9] which are driven by MPG; the semiphysiological FPG-Hb-HbA1c (FHH) model,[10] driven by FPG; and the FSI-FPG-HbA1c with a steady-state solution (FFH_{ss})[11] and FSI-FPG-HbA1c (FFH_2) models,[12] which mechanistically handle FPG and FSI, however, with a more empirical HbA1c model.

Wellhagen et al.[13] compared four of the aforementioned models (ADOPT, IGRH, FHH, and FFH_{ss}), looking at power, prognostic, and extrapolation predictive performance in relation to hypothetical drug effects. The authors of that work concluded that the best model depends on the intended use and the hypothetical drug effect. However, the extrapolation properties beyond 24 weeks of treatment remain unexplored, as does the benefit of using data from shorter trials, with or without a delayed onset of drug effect on glucose. Thus, this study aimed to investigate the predictive performance of previously published models of HbA1c with a focus on extrapolation properties from studies of 12 weeks or shorter to studies of 24 or 52 weeks with a direct and a delayed drug effect on glucose.

METHODS

Data

The data analyzed herein were created through simulations with the integrated glucose-insulin (IGI) model[14] connected to the IGRH model.[9] The collected biomarkers

were MPG (average from continuous glucose monitoring over 24 h), FPG, HbA1c, and fasting plasma insulin (FPI) extracted every second week up to 8 weeks and at 12, 24, and 52 weeks. The data from 24 and 52 weeks were only used as a reference for the "true" HbA1c.

Inclusion criteria for baseline HbA1c and FPG were 7.5%–9.8% (58.5–83.6 mmol/mol; HbA1c (mmol/mol) = 10.93 · HbA1c (%) – 23.5 mmol/mol)[15] and 7.0–13.3 mmol/L,[13] respectively. In addition, patients with glucose <4.5 mmol/L were excluded. These criteria were applied in a postprocessing step of the simulations and to the corresponding "true" underlying glucose concentrations without residual error. Thus, with the addition of residual error, the glucose could range outside the inclusion criteria, and the model used for simulations (i.e., IGRH model) would not necessarily provide the best fit.

A standardized meal plan was used with three large meals (62.5 g glucose) and three snacks (12.5 g glucose) per day, consumed at 8:00, 12:00, and 18:00 and at 10:00, 15:00, and 21:00, respectively. The baseline sample was taken before treatment initiation. Parallel, placebo-controlled studies were simulated, each with the following five arms: placebo, 500 and 1000 mg twice daily of a metformin-similar drug, and 1.2 and 1.8 mg once daily of a liraglutide-similar drug. The metformin-similar drug was simulated with the delayed inhibition of endogenous glucose production (EGP),[16] and the liraglutide-similar drug was simulated with direct stimulation of the incretin effect.[17] The incretin analog was titrated to the desired dose with 0.6 mg/week.[18] Both placebo effect and disease progression acted on glucose at steady state, a parameter of the IGI model. Placebo effect was implemented as an instantaneous decrease of 0.1 mmol/L,[19] whereas disease progression was time related with an increase of 0.2 mmol/L/year[20] The pharmacokinetics (PK) and effects on glucose of the liraglutide-similar drug were simulated using a previously published model,[17] with a linear stimulation of insulin secretion by liraglutide concentrations. To align the HbA1c observations created by the simulation with the published outcome,[17] the variability of clearance and volume of distribution were halved. Metformin PK was simulated using a previously published model.[21] The effect of metformin on glucose was implemented with a maximal fractional inhibition (I_{max}) function on EGP, and the delayed onset was handled with an effect-compartment model. The parameters of the delayed effect on EGP were adjusted to capture the magnitude and shape of the metformin response reported by Williams-Herman et al.[22] and were the following: I_{max} = 0.8, concentration giving half of I_{max} (IC_{50}) = 0.1 mg/L, and rate constant of delayed effect (k_{e0}) = 0.0003 h^{-1} (corresponding to a turnover of 20 weeks). Each study was simulated with 100 individuals per arm, and in total 100 studies were simulated and analyzed.

Investigated models

Figure 1 summarizes all investigated models.

Nathan regression

The Nathan regression is a linear regression between MPG and HbA1c at steady state[7]:

$$\mathrm{HbA1c} = \beta_1 \cdot \mathrm{MPG}_{SS,\mathrm{mmol/L}} + \beta_2 \quad (1)$$

where β_1 = 0.629 L/mol and β_2 = 1.629% (8.226 mmol/mol)[2] and are slope and intercept, respectively. In this model, MPG could be observed or modeled.

ADOPT model

The ADOPT model consists of two linked indirect response models, a model for the MPG and a model for HbA1c, where formation is driven by MPG and the intercept between MPG and HbA1c at steady state (β), where $k_{in,MPG}$ and $k_{in,HbA1c}$ are input rate of MPG and HbA1c, respectively, while $k_{out,MPG}$ and $k_{out,HbA1c}$ are output rate constants, for MPG and HbA1c, respectively.[8]

$$\frac{d\mathrm{MPG}}{dt} = k_{in,MPG} - k_{out,MPG} \cdot \mathrm{MPG}_{\mathrm{mmol/L}}$$

$$\frac{d\mathrm{HbA1c}}{dt} = \left(\mathrm{MPG}_{\mathrm{mmol/L}}(t) + \beta\right) k_{in,HbA1c} - k_{out,HbA1c} \cdot \mathrm{HbA1c} \quad (2)$$

IGRH model

The IGRH model describes HbA1c formation through the life span of RBC using 12 nonglycated ($RBC_{1...12}$) and 12 glycated ($RBC_{g,1...12}$) compartments of RBC. Each pair (i.e., RBC_x and $RBC_{g,x}$) represents various ages of RBCs.[9] The life span of RBC is inversely correlated with MPG, such that the life span is shorter with high glucose. In addition, a fraction of the precursors of RBCs are glycated at the release into the central circulation. Glycation is driven by MPG, and MPG could be modeled with an indirect response model (as ADOPT, Equation 2).

$$\mathrm{HbA1c}\left(\mathrm{MPG}_{\mathrm{mg/dL}}(t)\right) = \frac{\mathrm{RBC}_{g,1}\left(\mathrm{MPG}_{\mathrm{mg/dL}}(t)\right) + \cdots + \mathrm{RBC}_{g,12}\left(\mathrm{MPG}_{\mathrm{mg/dL}}(t)\right)}{\text{Total RBC}} \quad (3)$$

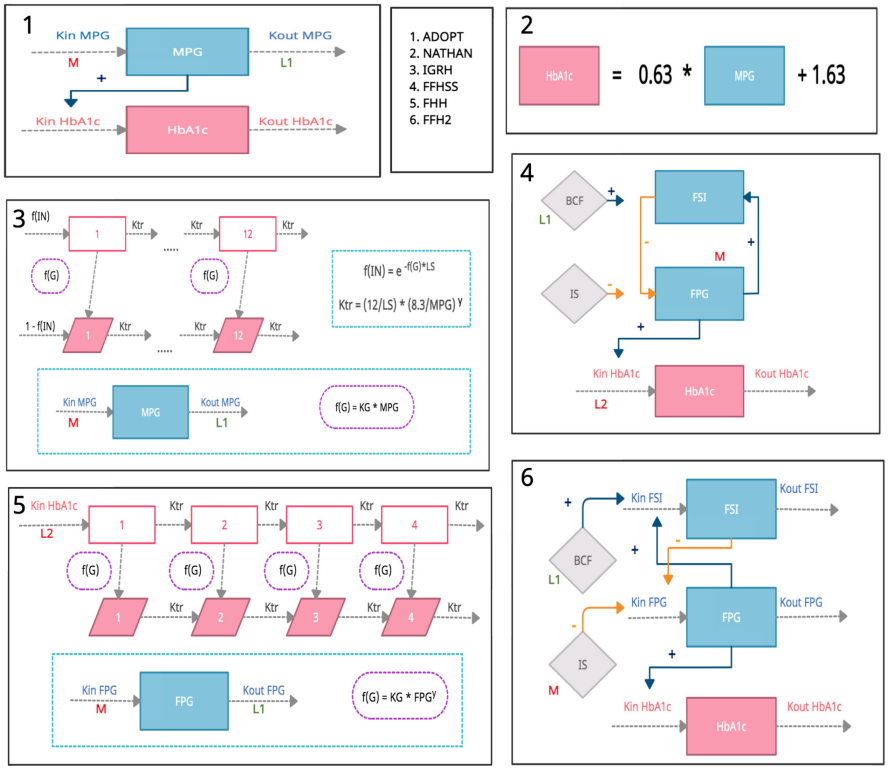

FIGURE 1 Schematic representation of the models. In the developed Nathan model, the MPG from the ADOPT model was used. Red indicates inhibition, and green indicates stimulation. ADOPT, A Dynamic HbA1c Endpoint Prediction Tool; BCF, β cell function; f(G), indicates where glucose has an effect; f(IN), glycation of precursors of hemoglobin; FFH2, FSI-FPG-HbA1c; FFHSS, FSI-FPG-HbA1c with steady-state solution; FHH, FPG-Hb-HbA1c; FPG, fasting plasma glucose; FSI, fasting serum insulin; Hb, hemoglobin; HbA1c, glycated hemoglobin; IGRH, Integrated Glucose-RBC-HbA1c; IS, insulin sensitivity effect; KG, glucose turnover; K_{in}, input rate constant; K_{out}, output rate constant; K_{tr}, transit rate constant; L1, liraglutide-similar drug effect on MPG/FPG; L2, liraglutide-similar drug effect on postprandial glucose; LS, lifespan; M, metformin-similar drug effect; MPG, mean plasma glucose; RBC, red blood cells.

FHH model

In the FHH model, HbA1c is formed by FPG in a nonlinear manner, related to the shape factor (γ).[10] FPG is modeled with an indirect response model, and glycation is driven through the life span of RBCs, similar to the IGRH model, with four nonglycated ($RBG_{1...4}$) and four glycated ($RBC_{g,1...4}$) compartments of RBCs of varying age. The life span of RBC is constant in this model, and no glycation of precursors is assumed.

$$\frac{d\mathrm{FPG}}{dt} = k_{\mathrm{in,FPG}} - k_{\mathrm{out,FPG}} \cdot \mathrm{FPG}_{\mathrm{mmol/L}}$$

$$\mathrm{HbA1c}\left(\mathrm{FPG}_{\mathrm{mmol/L}}(t)^{\gamma}\right) = \frac{\mathrm{RBC}_{g,1}\left(\mathrm{FPG}_{\mathrm{mmol/L}}(t)^{\gamma}\right) + \cdots + \mathrm{RBC}_{g,4}\left(\mathrm{FPG}_{\mathrm{mmol/L}}(t)^{\gamma}\right)}{\text{Total RBC}} \quad (4)$$

FFH_2 model

The FFH_2 model consists of three linked indirect response models for FSI, FPG, and HbA1c.[12] The IGI model generates FPI measurements, whereas the FFH_2 model was developed for FSI measurements. For simplicity, FSI was assumed to be equal to FPI.[23] FPG stimulates production of FSI, and FSI inhibits production of FPG. Baseline conditions of FPG and FSI are parameters in the model, from which β-cell function (BCF) and insulin sensitivity (IS) are derived using the homeostatic assessment (HOMA) model.[24] HbA1c is formed by FPG.

$$\frac{d\mathrm{FSI}}{dt} = \mathrm{BCF} \cdot \left(\mathrm{FPG}_{\mathrm{mmol/L}} - 3.5\right) \cdot k_{\mathrm{in,FSI}} - k_{\mathrm{out,FSI}} \cdot \mathrm{FSI}_{\mu\mathrm{U/mL}}$$

$$\frac{d\mathrm{FPG}}{dt} = \frac{k_{\mathrm{in,FPG}}}{\mathrm{IS} \cdot \mathrm{FSI}_{\mu\mathrm{U/mL}}} - k_{\mathrm{out,FPG}} \cdot \mathrm{FPG}_{\mathrm{mmol/L}}$$

$$\frac{d\mathrm{HbA1c}}{dt} = \mathrm{FPG}_{\mathrm{mmol/L}} \cdot k_{\mathrm{in,HbA1c}} - k_{\mathrm{out,HbA1c}} \cdot \mathrm{HbA1c} \quad (5)$$

FFH_{SS} model

The FFH_{SS} model is similar to the FFH_2 model except that (1) BCF and IS are parameters of the model from which

baseline FPG and FSI are derived using the HOMA model and (2) a steady-state solution is used instead of differential equations for FPG and FSI.[11] The steady-state solution is implemented to satisfy:

$$0 = \mathrm{FPG}^2 + 3.5\mathrm{FPG} - \frac{\mathrm{kG}}{\mathrm{kI} \cdot \mathrm{IS} \cdot \mathrm{BCF}}$$

in which kI and kG are $k_{in,FSI}/k_{out,FSI}$ and $k_{in,FPG}/k_{out,FPG}$, respectively. Solving the quadratic equation, keeping the positive solution, gives the following:

$$\mathrm{FPG} = -\frac{3.5}{2} + \sqrt{\left(\frac{3.5}{2}\right)^2 + \frac{\mathrm{kG}}{\mathrm{kI} \cdot \mathrm{IS} \cdot \mathrm{BCF}}}$$

$$\mathrm{FSI} = \mathrm{kI} \cdot \mathrm{BCF} \cdot (\mathrm{FPG} - 3.5) \quad (6)$$

FSI was assumed to be equal to FPI.[23]

Model building

The workflow of the study is presented in supplementary material Figure S1.

For the Nathan regression, two approaches were investigated to extrapolate 24- and 52-week HbA1c ($\mathrm{HbA1c}_{24/52}$). In Approach i, $\mathrm{HbA1c}_{24/52}$ was calculated from the observed MPG at Weeks 12, 8, 6, or 4, and in Approach ii, $\mathrm{HbA1c}_{24/52}$ was calculated from the model predictions of MPG at 24/52 weeks. Approach i used only observed glucose data at a particular study week (12, 8, 6, or 4), whereas Approach ii used all glucose data up to and including a particular study week in a model that predicted the MPG at 24/52 weeks. An indirect response model was used to model MPG data (the MPG part of Equation 2).

For the dynamic models (i.e., ADOPT, IGRH, FHH, FFH_{SS}, and FFH_2), the following two approaches to extrapolate $\mathrm{HbA1c}_{24/52}$ were investigated: (1) use a model where full model building was performed, assessing the need for structural changes between glucose and HbA1c, and (2) use the published model with priors from the original publications for all parameters unrelated to drug effect,[25] not allowing changes of the relationship between glucose and HbA1c. Model development was performed simultaneously for all studies. In Approach 1, model development of IGRH and FHH included investigations of the number of transit compartments. Full model development was performed independent of the approach for the drug effects, and the associated parameters were estimated without priors. Both approaches used biomarker data, that is, glucose, insulin (FFH_{SS} and FFH_2), and HbA1c, up to and including 12, 8, 6, or 4 weeks.

Drug effects were explored on the glucose (and insulin, when applicable) with various shapes, such as linear, E_{max}, and sigmoidal E_{max} models. In ADOPT, IGRH, and FHH, the drug effect was implemented multiplicative to glucose k_{in} and k_{out}. Delay of metformin concentrations for the effect was explored, as was drug-specific glucose k_{out}. In FFH_{SS} and FFH_2, drug effect was investigated on BCF, IS, and the turnover of FPG (k_G) for both drug arms. In addition, for FHH and FFH_{SS}, an additional FPG-dependent effect of the liraglutide-similar drug was investigated on HbA1c.

The final parameters were determined simultaneously for all studies and later used to provide extrapolations of ΔHbA1c at 24 weeks ($\Delta\mathrm{HbA1c}_{24}$) and 52 weeks ($\Delta\mathrm{HbA1c}_{52}$) from the point estimates without uncertainty.

Model discrimination, result evaluations, and software

Model development was guided by objective function value (OFV), goodness-of-fit graphs, predictive performance and changes in interindividual variability (IIV), and relative standard error. The likelihood ratio test with a p value = 0.05 and the Akaike information criterion (AIC) were used for hierarchical models and non-hierarchical models, respectively. Predictive performance was assessed using visual predictive checks (VPCs) with 1000 samples, assessing the 2.5th, 50th, and 97.5th percentiles of simulated data compared with the corresponding percentiles of observed data.

Extrapolation performance of the dynamic models was assessed both graphically and statistically by comparing the "true" $\Delta\mathrm{HbA1c}_{24/52}$ to model extrapolations. Graphically, the median of the observed $\Delta\mathrm{HbA1c}_{24/52}$ and the median and the 95% confidence interval (CI) of the model-predicted $\Delta\mathrm{HbA1c}_{24/52}$ per dose were displayed. In addition, the observed and model-predicted $\Delta\mathrm{HbA1c}_{24}$ were tested with an unpaired, two-sided t-test for each model, study duration, and dose to assess the statistical significance between model predictions and observations. To assess whether predictions at 24 and 52 weeks differed, the model-predicted $\Delta\mathrm{HbA1c}_{24}$ and $\Delta\mathrm{HbA1c}_{52}$ were tested with a paired, two-sided t-test for each model, study duration, and dose. The p-value was assessed as the median of all p values of all t-tests performed for each study (N = 100). The Nathan regression was assessed as the median response (i.e., $\Delta\mathrm{HbA1c}_{24/52}$) and 90% range of all individuals in the study, comparing the "true" response to extrapolations.

To assess the overall predictive performance of Approach 1 for both drugs simultaneously, model averaging (MA) with three different MA approaches was performed: AIC (focusing on descriptive performance), cross-validation (CV; focusing on predictive performance[26] [E. Salomonsson, 2019, unpublished data]), and bootstrap (BS; focusing on parameter uncertainty).[27] As the models

used different biomarkers, the MA assessments were performed in the following two steps: (1) parameter estimation with the required biomarkers and (2) prediction of HbA1c with fixed parameters from Step 1 to generate HbA1c-specific OFV. These HbA1c-specific OFVs for both drugs were then used to calculate the MA weights[28]:

$$\Delta I_m = I_m - I_{m,\min} \tag{7}$$

$$W_m = \frac{1}{n}\left(\frac{e^{\frac{-\Delta I_m}{2}}}{\sum_{m'=1}^{M} e^{\frac{\Delta I_{m'}}{2}}}\right) \tag{8}$$

where ΔI_m is the difference in OFV between the model with the lowest OFV ($I_{m,\min}$), that is, the best model and all other models; M is the number of compared models; and n is the number of individuals. Predictions of ΔHbA1c were generated using the MA weights to scale the predictions of ΔHbA1c from each model. The stepwise workflow of all three MA approaches is presented in the Supplementary Text.

Data management, statistical calculations, and graphical evaluation were performed using R Version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).[29] The simulation, estimation, and extrapolation steps were performed using NONMEM Version 7.4.4 with PsN Version 5.2.0 (Icon Development Solutions)[30] using Perl-speaks-NONMEM.[31]

RESULTS

Model development

With full model building, all parameters of the models were estimated except for three physiological parameters of the IGRH model, which were fixed to published values: life span of RBCs, life span of precursors, and IIV in precursor life span. Changing the number of transit compartments (IGRH and FHH models) did not improve the fit. A statistically significant improvement was seen when including a correlation between the IIV of baseline glucose and the life span of RBC (ΔOFV = −1212). All models showed good predictive performance within the study (assessed by VPCs) and satisfactory goodness of fit. Although the developed models showed overall slightly better fits than the corresponding models relying on priors (data/figures not shown), estimations showed no signs of misfit in comparison with the data used to estimate parameters.

Independent of approach, the same structural models were identified for the drug effects, that is, the description of metformin and liraglutide effects on glucose/insulin were unaffected by assumptions about the relationship between glucose and HbA1c. Metformin was best described with an I_{max} function inhibiting $k_{in,MPG/FPG}$ for Nathan, ADOPT, IGRH, and FHH and the FPG turnover for FFH_{SS} and FFH_2. An additional delay was estimated when metformin drug effects were implemented on the turnover in the FFH_{SS} and FFH_2. Liraglutide was best described with an E_{max} function stimulating $k_{out,MPG/FPG}$ for Nathan, ADOPT, IGRH, and FHH and a linear stimulation of BCF for FFH_{SS} and FFH_2. For FHH and FFH_{SS}, the liraglutide was significantly improved when estimating an additional, FPG-dependent drug effect on HbA1c (FHH and FFH_{SS}: ΔOFV = −5118 and ΔOFV = −3668, respectively, Approach 1).

$$\frac{d\text{HbA1c}}{dt} = \text{FPG}_{\text{mmol/L}} \cdot \text{L2} \cdot k_{\text{in,HbA1c}} - k_{\text{out,HbA1c}} \cdot \text{HbA1c} \tag{9}$$

The additional FPG-dependent liraglutide dose–effect relationship (L2) was implemented as the E_{max} model. Due to long runtimes, an additional FPG-dependent liraglutide effect was not investigated for FFH_2. Estimating a drug-specific MPG/FPG k_{out} significantly improved the fits for ADOPT and FFH_2.

Prediction performance

The main results are shown in Figures 2, 3, and 4 (extrapolations for 52 weeks in supplementary material, Figures S2 and S3). Model-predicted $\Delta HbA1c_{24}$ and $\Delta HbA1c_{52}$ differed for FFH_{SS} and FFH_2 according to the *t*-tests performed, and $\Delta HbA1c_{52}$ provided worse predictions than $\Delta HbA1c_{24}$. All other models showed no statistical difference between predictions. Thus, for simplicity, only the best results (i.e., $\Delta HbA1c_{24}$) are discussed in the following sections.

Nathan regression

For the Nathan regression, ΔHbA1c extrapolations using model-predicted MPG were more precise than extrapolations using observed MPG (Figure 2). The dose response was similar and overpredicted for most scenarios. It is clear from the variability of predictions that the model-predicted MPG did not propagate the glucose variability into HbA1c predictions, which the observed MPG did.

For the drug with delayed effect (i.e., metformin-similar), the best prediction was seen after 4 weeks of data. This may seem surprising at first, however, as this method overpredicted $\Delta HbA1c_{24}$ with a good prediction

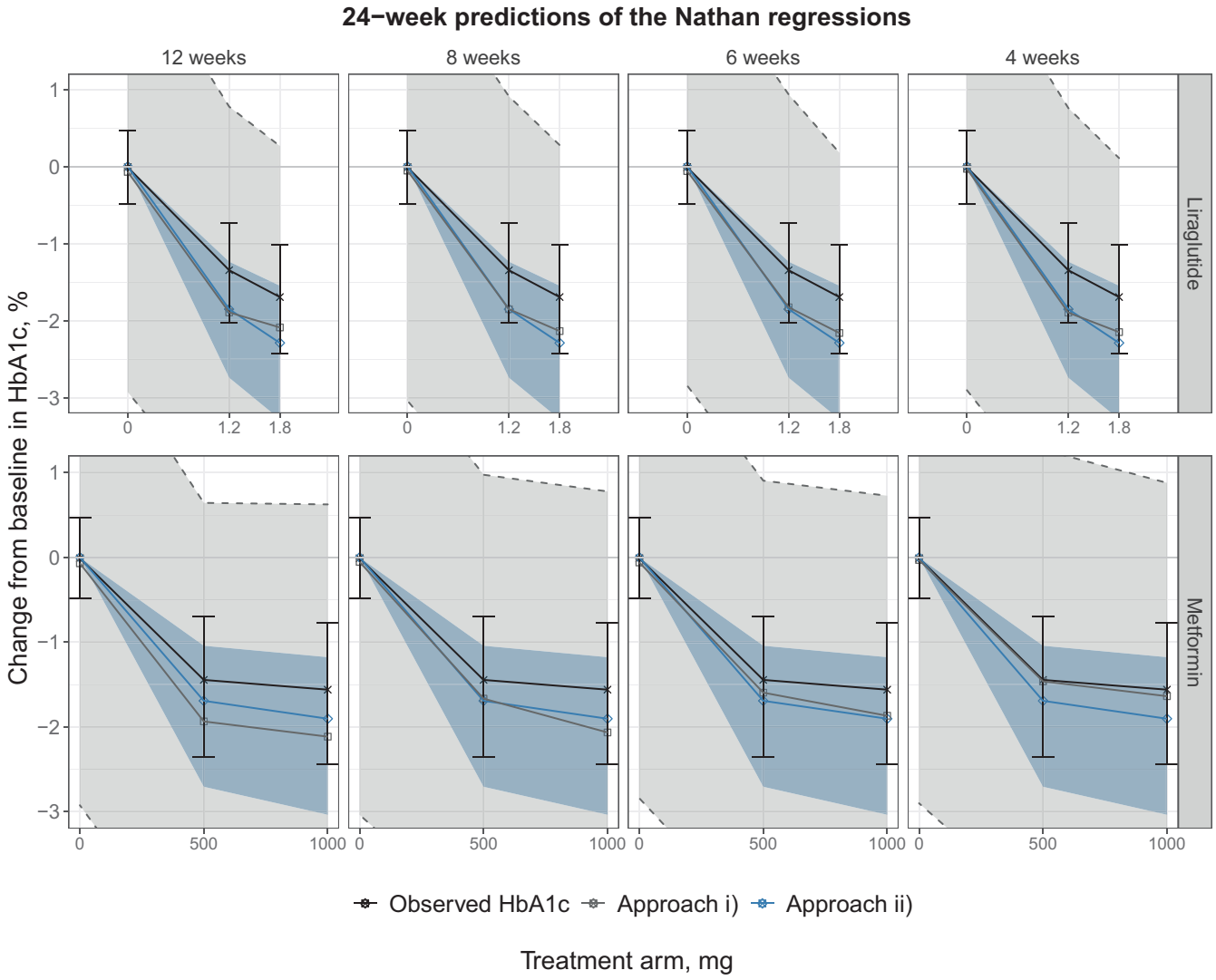

FIGURE 2 The 24-week individual extrapolations of change in HbA1c from the Nathan regression[7] for the liraglutide- (top panels) and metformin-similar drugs (bottom panels). The figure displays observed response (black) compared with extrapolated response using observed MPG (Approach i, gray) and model-predicted MPG (Approach ii, blue) for different study durations (from left to right): 12, 8, 6, and 4 weeks. The lines represent medians, whereas the shaded areas (extrapolations) and error bars (observed) display the 90% range of individuals' responses in the study. HbA1c, glycated hemoglobin; MPG, mean plasma glucose.

of glucose, an underprediction of glucose from 4 weeks of data resulted in an apparent better extrapolation. The Nathan regression with model-predicted MPG was insensitive to study duration reduction and showed the same overprediction for shorter studies as was observed for 12 weeks, indicating a robust model for MPG despite the data reduction.

Dynamic models

As shown in Figures 3 and 4, the $\Delta HbA1c_{24}$ extrapolations using 12-week data were similar between developed models and models relying on priors; however, as study duration shortened, the differences became more pronounced. In general, the developed models performed equal to or better than the models relying on priors. Notably, the IGRH model, used for data creation, showed the largest difference between the approaches, with the better performance of the developed model. Also, the FHH model showed some difference for the shorter study durations (4 and 6 weeks), with the prior model overpredicting the dose response, in particular for liraglutide (i.e., the direct drug effect on the secretion of insulin). Notable, for 4 weeks of metformin data, the prior model of FHH was better at extrapolating $\Delta HbA1c_{24}$ than the developed model. Also, the FFH_2 with priors outperformed the developed model for the 4-week study duration with metformin treatment. The glucose predictions of metformin treatment with the

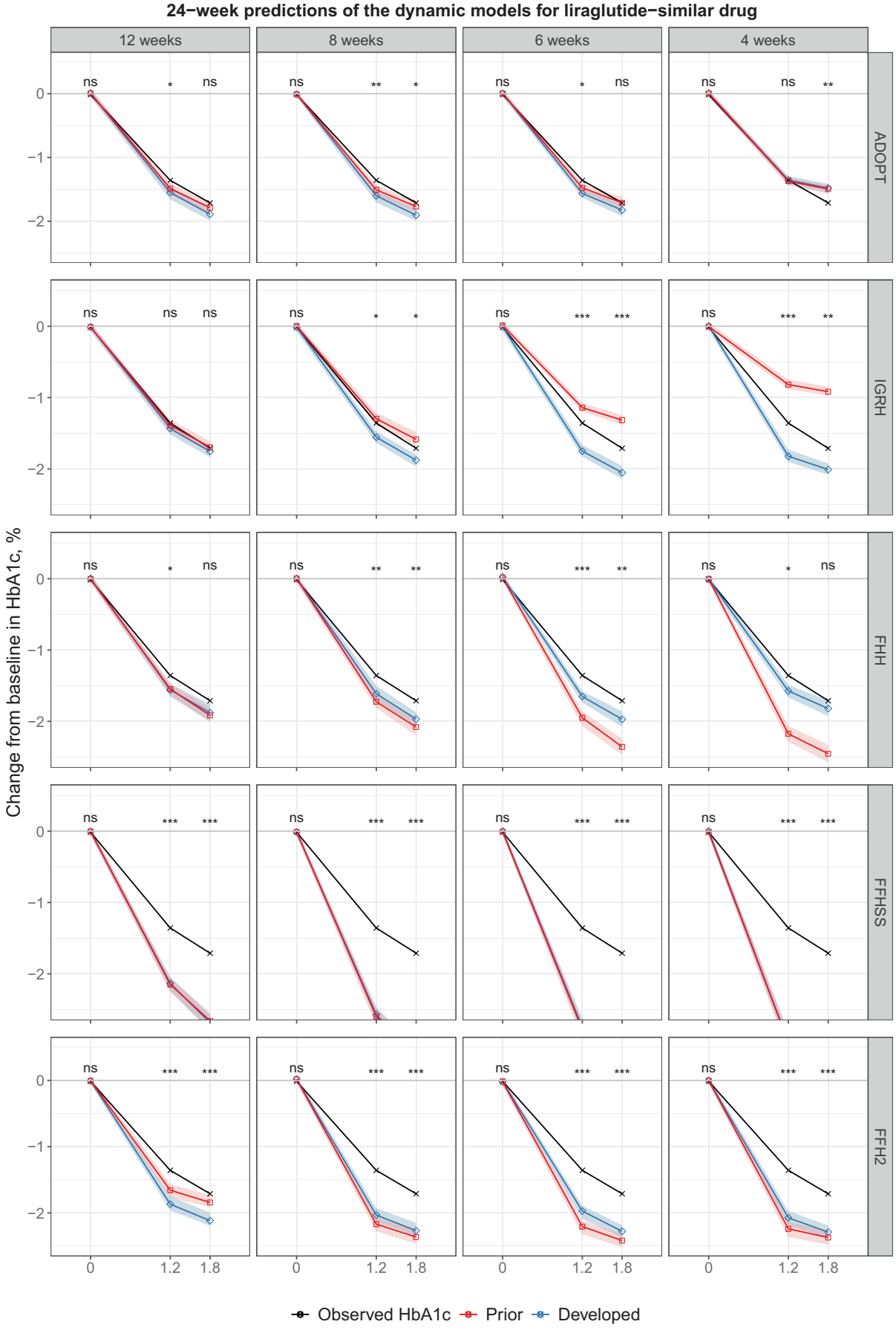
24−week predictions of the dynamic models for liraglutide−similar drug
12 weeks
8 weeks
6 weeks
4 weeks
ADOPT
IGRH
FHH
FFHSS
FFH2
Change from baseline in HbA1c, %
Observed HbA1c
Prior
Developed

FIGURE 3 The 24-week study extrapolations of change in HbA1c for different doses of the liraglutide-similar drug with the dynamic models (top to bottom): ADOPT,[8] IGRH,[9] FHH,[10] FFH_{SS},[11] FFH_2,[12] and Nathan.[7] The figure displays the observed response (black) compared with the extrapolated response from models relying on priors (red) and developed models (blue) for the study durations (left to right): 12, 8, 6, and 4 weeks. Statistically significant differences in responses between extrapolations from the developed models and observations are indicated with asterisks (* <0.05, ** <0.01, *** <0.001). The lines represent medians, whereas the shaded areas correspond to the 95% confidence intervals of the extrapolated responses in 100 studies. ADOPT, A Dynamic HbA1c Endpoint Prediction Tool; FFH2, FSI-FPG-HbA1c; FFHSS, FSI-FPG-HbA1c With Steady State; FHH, FPG-Hb-HbA1c; FPG, fasting plasma glucose; FSI, fasting serum insulin; Hb, hemoglobin; HbA1c, glycated hemoglobin; IGRH, Integrated Glucose-RBC-HbA1c; ns, not significant; RBC, red blood cells.

prior models are similar across study durations for these models (supplementary material Figure S5) and thus the difference must be in the estimate of turnover of Hb/HbA1c, which appeared sensitive to the study duration reduction.

Using 12 weeks of data, most models performed well in predicting $\Delta HbA1c_{24}$ (Figures 3 and 4). The ADOPT and FHH models overpredicted the liraglutide-similar drug, dose = 1.2 mg ($p = 0.018$ and $p = 0.032$, respectively), but other arms with these models and all arms of the IGRH model performed well. The worst performance was seen by the FFH_{SS} model (both drugs). Wellhagen et al. reported similar results.[13] The second-worst model was FFH_2, which showed misprediction in particular for liraglutide, although smaller than FFH_{SS}.

In general, changes in extrapolations of $\Delta HbA1c_{24}$ were surprisingly small as the study duration was reduced. The dynamic models displayed two different distinct behaviors: a gradual overprediction of the dose response for the liraglutide-similar drug with a shorter study duration (IGRH and FFH_{SS}) and a gradual deterioration of the dose response for the metformin-similar drug with a shorter study duration, where ADOPT and FHH gradually underpredicted dose response and FFH_2 overpredicted. These gradual changes for metformin appear to be linked to changes in glucose extrapolations (supplementary material Figure S5); however, the glucose data of liraglutide (supplementary material Figure S4) do not indicate a reason for the gradual deterioration of the IGRH model with reduced study duration, and thus the reason must be in the estimates of turnover of Hb/HbA1c.

The MA weights were similar across all approaches (AIC, CV, and BS) and in line with the results of the statistical testing and the graphical assessment (Figure 5). The FFH_{SS} model was assigned the lowest MA weight for all approaches and study durations. The MA weights were similar for all other models, with generally slightly lower weights for FFH_2 compared with ADOPT, IGRH, and FHH. The ADOPT model was weighted considerably lower for the shortest study duration, transferring weight to the IGRH model. This considerable deterioration of the ADOPT model for the shortest study duration may seem surprising as the ADOPT model appeared rather insensitive to study duration reductions for the liraglutide-similar drug. However, as the MA weight represents the average weight for both treatments, the low MA weight for the ADOPT model with a study duration of 4 weeks originates from the model's poor extrapolation of the metformin-similar drug.

Using the MA weights for extrapolations of ΔHbA1c showed that the differences between the MA-based extrapolations were similar to the best model (Figure 6; the ADOPT model for 12, 8, and 6 weeks and the IGRH model for 4 weeks of study duration) and that the predictions were better captured with the MA extrapolations, compared with the best single model, in the short-duration studies (4 weeks with both drugs, and 6 weeks with metformin).

DISCUSSION

In this work, the predictive performance for the 24- and 52-week dose response of six models of HbA1c was investigated in relation to reduced study durations and drugs with different mechanisms of action. A natural extension of our work would be to investigate the findings on study duration with clinical data from another compound in development to validate the study duration exploration.

All pharmacometric models performed reasonably well in predicting the long-term HbA1c for different drugs and doses except for the FFH models. The FFH_2 had a slightly better performance than the FFH_{SS}; however, neither of them performed on par with ADOPT, IGRH, or FHH. The FFH models use an indirect response model driven by FPG for HbA1c. The IGRH model uses a transit model driven by MPG, the ADOPT model uses an indirect response model driven by MPG, and the FHH model uses a transit model driven by FPG in a nonlinear relationship. Thus, it appears as if MPG or a nonlinear FPG driving HbA1c is crucial for a good predictive performance. In addition, the FFH models are the only models that use insulin in the assessment. Thus, alternatively, the poor extrapolation properties may be related to insulin. This explanation is, however, less likely than the earlier hypothesis because the glucose extrapolations of these models were good (Figure S4 and S5) and the main driver for the relationship with HbA1c

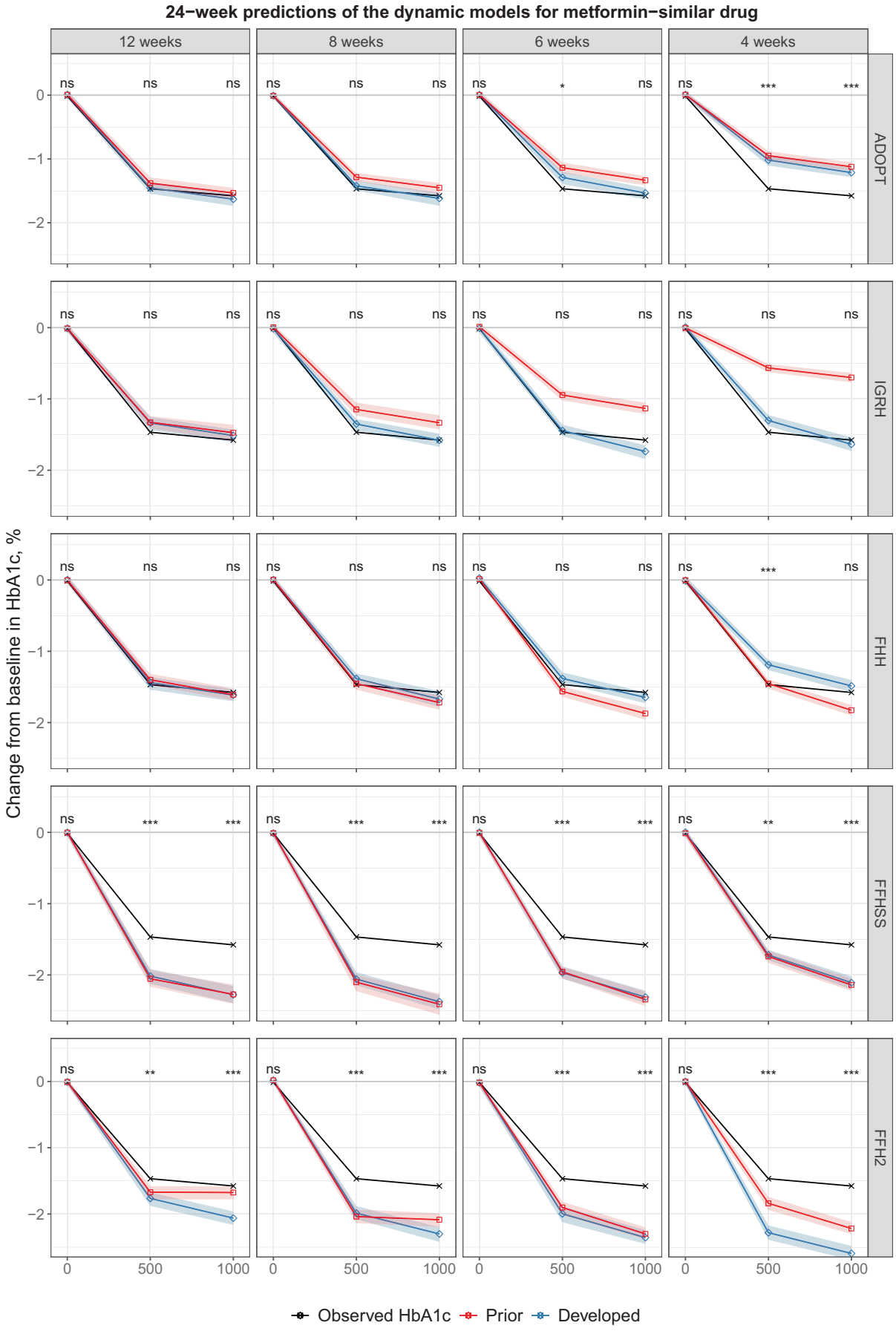
24−week predictions of the dynamic models for metformin−similar drug
12 weeks
8 weeks
6 weeks
4 weeks
ADOPT
IGRH
FHH
FFHSS
FFH2
Change from baseline in HbA1c, %
Observed HbA1c
Prior
Developed

FIGURE 4 The 24-week study extrapolations of change in HbA1c for different doses of the metformin-similar drug with the dynamic models (top to bottom): ADOPT,[8] IGRH,[9] FHH,[10] FFH_{SS},[11] FFH_2,[12] and Nathan.[7] The figure displays the observed response (black) compared with the extrapolated response from models relying on priors (red) and developed models (blue) for the study durations (left to right): 12, 8, 6, and 4 weeks. Statistically significant differences in responses between extrapolations from developed models and observations are indicated with asterisks (* <0.05, ** <0.01, *** <0.001). The lines represent medians while the shaded areas correspond to the 95% confidence intervals of the extrapolated responses in 100 studies. ADOPT, A Dynamic HbA1c Endpoint Prediction Tool; FFH2, FSI-FPG-HbA1c; FFHSS, FSI-FPG-HbA1c With Steady State; FHH, FPG-Hb-HbA1c; FPG, fasting plasma glucose; FSI, fasting serum insulin; Hb, hemoglobin; HbA1c, glycated hemoglobin; IGRH, Integrated Glucose-RBC-HbA1c; ns, not significant; RBC, red blood cells.

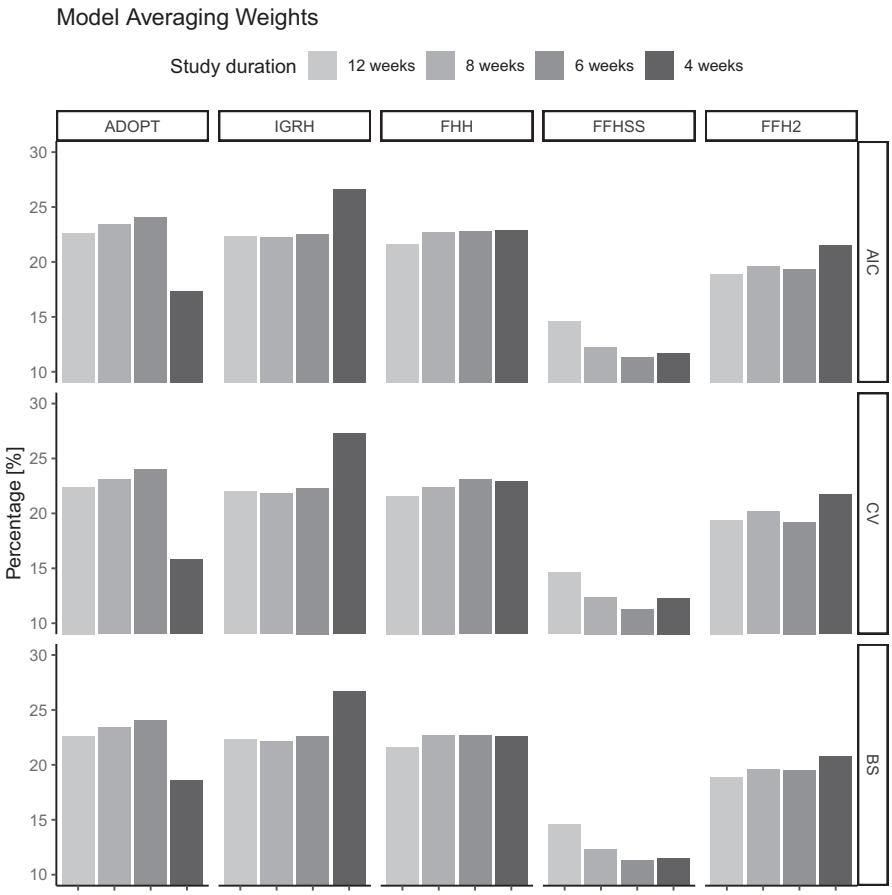

FIGURE 5 Model-averaging weights (in percentages) per model (from left to right: ADOPT,[8] IGRH,[9] FHH,[10] FHHSS,[11] FFH2[12]) from the model-averaging approaches (from top to bottom: AIC, BS, CV) for different study durations (12, 8, 6, and 4 weeks of data). ADOPT, A Dynamic HbA1c Endpoint Prediction Tool; AIC, Akaike information criterion; BS, bootstrap; CV, cross-validation; FFH2, FSI-FPG-HbA1c; FFHSS, FSI-FPG-HbA1c With Steady State; FHH, FPG-Hb-HbA1c; FPG, fasting plasma glucose; FSI, fasting serum insulin; Hb, hemoglobin; HbA1c, glycated hemoglobin; IGRH, Integrated Glucose-RBC-HbA1c; RBC, red blood cells.

is glucose. Similar studies have provided support for the superiority of MPG over FPG in predicting HbA1c.[32] A nonlinearity factor of FPG in the FFH models may improve predictive performance; however, this was not investigated in the current work.

We showed in this study that, despite ignoring the disease progression, models with reasonable extrapolations of $\Delta HbA1c_{24}$ provided equally good extrapolations of $\Delta HbA1c_{52}$, also for shorter study durations. In the current work, simulations to create data were performed with a linear disease progression of 0.2 mmol/L/year in fasting glucose, corresponding to ~0.15%/year (0.76 mmol/mol/year; HbA1c (mmol/mol) = 10.93 · HbA1c (%) – 23.5 mmol/mol) in HbA1c. This small progression originates from the placebo arm in the study by Pratley et al.[20] In a study population with a faster progression, the validity of the conclusions would be affected and the estimation model would have to include disease progression to perform equally well. A disease progression, as was used in the current work, on fasting glucose is a simplification of diabetic disease progression, which would also impact the postprandial glucose response. This simplification may have implications for the conclusions of the extrapolations, in particular for the liraglutide-similar drug, which has an additional effect on postprandial glucose.

Among the models that performed well with 12 weeks of data (ADOPT, IGRH, and FHH), shortening the study duration only marginally affected the extrapolations. The deterioration from 12 to 8 weeks of data was within acceptable limits. Changes in $\Delta HbA1c_{24}$ between 12 and 8 weeks were 0%, 0.15%, and 0.08% for the highest dose

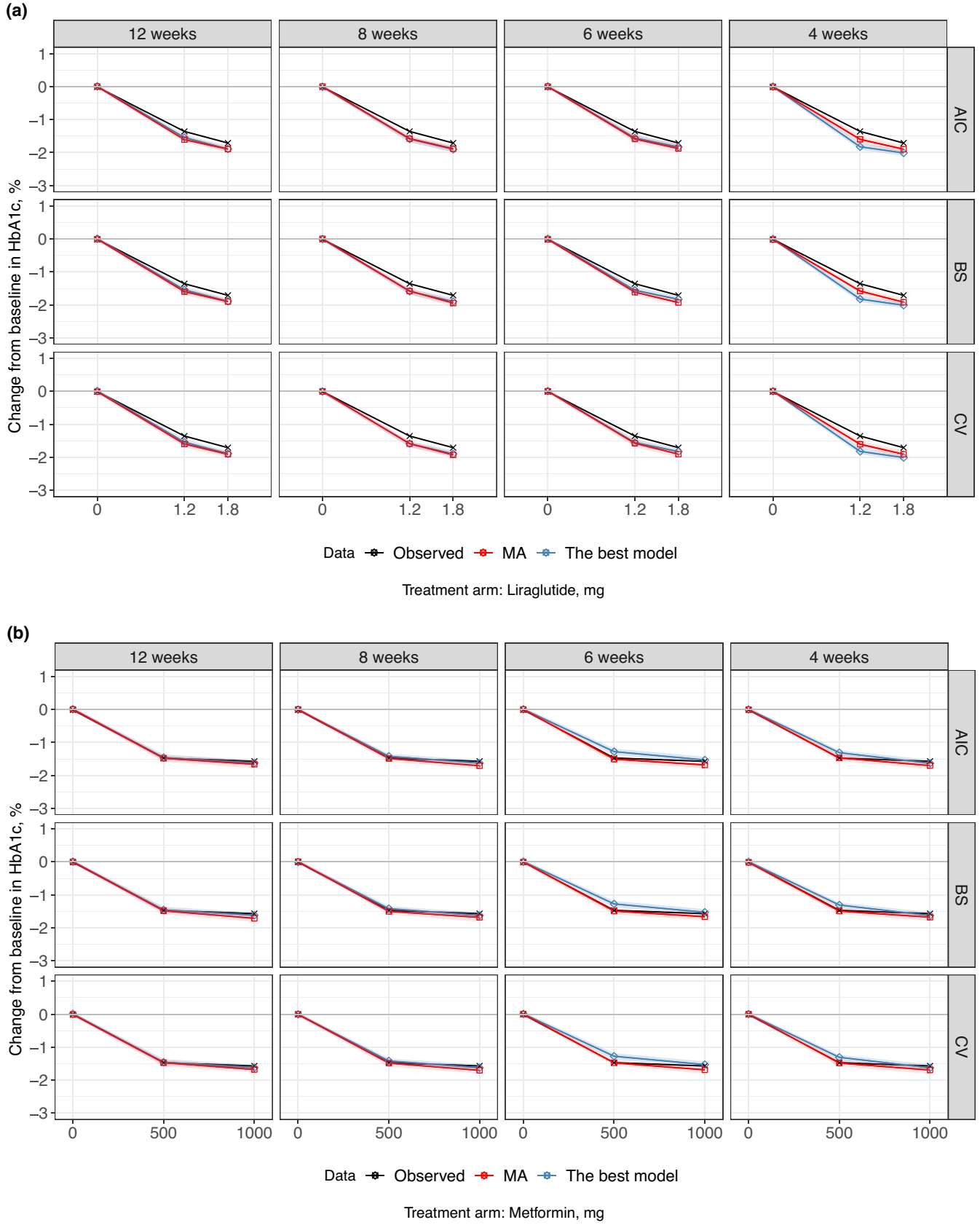

FIGURE 6 The 24-week study extrapolation of change in HbA1c for different doses of the (a) liraglutide- and (b) metformin-similar drugs comparing observed responses (black) to MA (red) and the corresponding best model (blue), that is, ADOPT[8] for 12, 8, and 6 weeks and IGRH[9] for 4 weeks. The lines correspond to the medians, whereas the shaded areas correspond to the 95% confidence intervals of the extrapolated responses in 100 studies. ADOPT, A Dynamic HbA1c Endpoint Prediction Tool; AIC, Akaike information criterion; BS, bootstrap; CV, cross-validation; HbA1c, glycated hemoglobin; IGRH, Integrated Glucose-RBC-HbA1c; MA, model averaging; RBC, red blood cells.

of the liraglutide-similar drug and 0%, 0.08%, and 0.04% for the highest dose of the metformin-similar drug for ADOPT, IGRH, and FHH, respectively. Reducing the duration further to 6 weeks resulted in considerable deterioration for the IGRH model with the liraglutide-similar drug and some deterioration for the ADOPT model with the metformin-similar drug. Thus, as general advice based on the results in this study, reducing the study duration to 8 weeks would only marginally affect the predictive performance of the good-performing models (ADOPT, IGRH, and FHH) with the investigated types of drugs.

The ADOPT model performed the best overall. This was supported by both the statistical and graphical assessments of extrapolation as well as the MA approaches. The combination of a flexible model structure and MPG-driven HbA1c seems beneficial. Møller et al.[32] reported that the ADOPT model had good predictive performance for various phase III studies, however, without comparisons with other models. Wellhagen et al.[13] concluded that the ADOPT model, despite a simple structure, performed well in comparison with other, more complex models with various hypothetical drug effects, however, only for the usage of 12-week data. We showed in the current work that the ADOPT model was rather insensitive to the reduction of study duration, except with study duration much shorter than the drug effect onset for delayed effects (4-week metformin). Thus, this model appears to be an overall good choice for predicting HbA1c from MPG. The potential challenge using MPG is that 24-h glucose sampling is commonly not done in early-phase drug development for noninsulin therapeutics, thus the appropriate biomarker is often missing.

The next best models were the IGRH and the FHH. A good performance of the IGRH model was expected, as the model was used for data creation, although the data were postprocessed with the inclusion criteria. Surprisingly though, the IGRH model relying on priors, identical to the model used for data creation, had a worse performance than the developed IGRH model. The observed misfit between prediction and "true" likely originated from the inflexibility of the IGRH model with priors, in particular, the lack of correlations between estimates of IIV. It is possible that although the simulations were performed without these correlations, they were introduced by the inclusion criteria applied.

The MA approaches provided an overall assessment for both drugs simultaneously. Based on those results, the IGRH and the FHH models were weighted equally in all but the shortest study duration. Unsurprisingly, the FHH model was struggling with extrapolating liraglutide, where the drug effect is related to postprandial glucose. The model performance of the FPG-driven models was improved with an additional FPG-dependent effect, allowing an additional degree of freedom for the relationship between FPG and HbA1c. However, despite this addition, the model could not fully capture the glucose-HbA1c relationship.

In the IGRH, FHH, and Nathan models, the relationship between MPG and HbA1c is less flexible than in the ADOPT model. In the IGRH and FHH models, the life span of RBC is less flexible than the turnover of HbA1c in the ADOPT model, as it is kept at physiologically reasonable estimates. Thus, the misfits of glucose cannot be compensated with changes in the relationship between MPG and HbA1c, with these less-flexible models. The 12-week data provide an excellent fit of glucose (see Figures S4 and S5), whereas the change in HbA1c is slightly overpredicted for these inflexible models. Consequently, when the drug effect is delayed, the glucose is naturally underpredicted with a reduced study duration, and thus, the apparent predictions of HbA1c improve with a shorter study duration.

The Nathan regression overpredicted the extrapolation of dose response, although with a better representation of variability when using model-predicted MPG. As the Nathan regression propagates the measurement error of MPG to HbA1c predictions, we recommend that when using the Nathan regression, MPG should be modeled. In addition, although model-based MPG improved the predictions, this approach assumes a relationship between steady-state MPG and HbA1c without IIV, an issue that has been discussed previously.[33] As such, the predictions of HbA1c using pharmacometric models are preferred as they account more accurately for IIV in the relationship between glucose and HbA1c. In addition, some of the dynamic models provide better extrapolation than the Nathan regression.

Supporting the parameter estimation using information from prior analyses may be beneficial, especially when information in the current analysis is sparse.[22,34] However, in the current study, the performance of these prior models was worse than for the developed models, sometimes even with the sparsest information (i.e., 4-week study duration). As the dose–response relationship was the same between the prior and developed models, the differences (Figures 3 and 4) originate from parameter estimates. If the estimations had been initialized at values further away from the priors, the results might have been different.

In addition to the dose–response extrapolation, we used MA to quantify the differences between the performance of the models. Standard MA could not be performed with these models as they use different biomarkers to predict HbA1c. Thus, we developed an approach enabling MA for models using different biomarkers. Further investigations of this approach are needed to explore the general applicability; however, for the current application, it was feasible and provided

consistent support for the results of the extrapolations through the MA weights (Figure 5). The trends of the MA weight seem to be unaffected by the approach used to derive them (i.e., AIC, XV and BS). Thus the ranking of models was similar independent if derived based on overall fit (i.e., AIC), predictive performance (i.e., XV) or uncertainty assessment (i.e., BS); the ADOPT model was weighted the highest, except with 4-week study duration, where the IGRH model was the best model. For a short study duration (4 and 6 weeks), the MA approaches (any of the investigated) seem to provide an alternative in terms of more accurate extrapolations (Figure 6).

CONCLUSION

In conclusion, the prediction performance of previously published models of HbA1c with the key focus on their extrapolation properties was analyzed and compared. Overall, the ADOPT model showed the best prediction performance for drugs with direct or delayed drug effects with the short study duration, closely followed by the IGRH and FHH models. The Nathan regression can be used for a quick assessment but does result in a bias. In addition, if the Nathan regression is used, the MPG should be predicted with a model. Overall, a study duration of at least 8 weeks is needed for accurate long-term HbA1c predictions with the investigated models. For shorter study durations, weighting the model extrapolations with MA provides a better prediction. Thus, there appears to be room for a reduction in study duration in trials of HbA1c.

AUTHOR CONTRIBUTIONS

H.K., A.A-M., J.Y.C., P.G, and M.C.K. wrote the manuscript. H.K., M.C.K., J.Y.C., and P.G. designed the research. H.K., M.C.K., and A.A-M. performed the research and analyzed the data.

FUNDING INFORMATION

Associate Professor Kjellsson and Mrs. Kunina received grants from Eli Lilly & Company during the conduct of the study.

CONFLICT OF INTEREST

Associate Professor Kjellsson and Mrs. Kunina received grants from Eli Lilly & Company during the conduct of the study. Dr. Chien and Dr. Garhyan are employees and stockholders of Eli Lilly & Company. All other authors declared no competing interests for this work.

ORCID

Jenny Y. Chien https://orcid.org/0000-0001-9303-5947
Maria C. Kjellsson https://orcid.org/0000-0003-3531-9452

REFERENCES

1. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373:2117-2128.
2. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2016;375:311-322.
3. ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biomedical Information. [updated 2021 May]. Available: http://clinicaltrials.gov/. Accessed 13 Jul 2021.
4. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2009;32(Suppl 1):S62-S67.
5. Cersosimo E, Triplitt C, Solis-Herrera C, et al. Pathogenesis of type 2 diabetes mellitus. [Updated 2018 Feb 27]. In: Feingold KR, Anawalt B, Boyce A, et al., eds. *Endotext [Internet]*. MDText.com, Inc.; 2000.
6. Ajmera I, Swat M, Laibe C, Le Novère N, Chelliah V. The impact of mathematical modeling on the understanding of diabetes and related complications. *CPT Pharmacometrics Syst Pharmacol*. 2013;10:e54.
7. Nathan DM, Kuenen J, Borg R, Zheng H, Schoenfeld D, Heine RJ. A1c-derived average glucose study group. Translating the A1C assay into estimated average glucose values. *Diabetes Care*. 2008;31(8):1473-1478.
8. Møller JB, Overgaard RV, Kjellsson MC, et al. Longitudinal modeling of the relationship between mean plasma glucose and HbA1c following antidiabetic treatments. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e82.
9. Lledo-García R, Mazer NA, Karlsson MO. A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects. *J Pharmacokinet Pharmacodyn*. 2013;40:129-142.
10. Hamrén B, Björk E, Sunzel M, Karlsson M. Models for plasma glucose, HbA1c, and hemoglobin interrelationships in patients with type 2 diabetes following Tesaglitazar treatment. *Clin Pharmacol Ther*. 2008;84(2):8-235.
11. Choy S, Kjellsson MC, Karlsson MO, de Winter W. Weight–HbA1c–insulin–glucose model for describing disease progression of type 2 diabetes. *CPT Pharmacometrics Syst Pharmacol*. 2016;5(1):11-19.
12. Gaitonde P, Hurtado FK, Garhyan P, Chien JY, Schmidt S. Development and quantification of a drug-disease modeling platform to characterize clinically relevant endpoints in type 2 diabetes trials. *Clin Pharmacol Ther*. 2018;104(4):699-708.
13. Wellhagen GJ, Karlsson MO, Kjellsson MC. Comparison of power, prognosis and extrapolation properties of four population pharmacodynamic models of HbA1c for type 2 diabetes. *CPT Pharmacometrics Sys Pharmacol*. 2018;7(5):331-341.
14. Jauslin PM, Frey N, Karlsson MO. Modeling of 24-hour glucose and insulin profiles of patients with type 2 diabetes. *J Clin Pharmacol*. 2011;51(2):153-164.
15. Ostawal A, Mocevic E, Kragh N, Xu W. Clinical effectiveness of liraglutide in type 2 diabetes treatment in the real-world setting: a systematic literature review. *Diabetes Ther*. 2016;7:411-438.
16. Stumvoll M, Nurjhan N, Perriello G, Dailey G, Gerich JE. Metabolic effects of metformin in non-insulin-dependent diabetes mellitus. *N Engl J Med*. 1995;333:550-554.

17. Røge PM et al. The effects of a GLP-1 analog on glucose homeostasis. *CPT Pharmacometrics Syst Pharmacol*. 2015;4:28-36.
18. RxList. Victoza adult dosage [Internet]. Available at: https://www.rxlist.com/victoza-drug.htm#indications. Accessed 12 March 2021.
19. Kjellsson MC, Cosson VF, Mazer NA, Frey N, Karlsson MO. A modelbased approach to predict longitudinal HbA1c, using early phase glucose data from type 2 diabetes mellitus patients after anti-diabetic treatment. *J Clin Pharmacol*. 2013;53:589-600.
20. Pratley R, Nauck M, Bailey T, et al. One year of liraglutide treatment offers sustained and more effective glycaemic control and weight reduction compared with sitagliptin, both in combination with metformin, in patients with type 2 diabetes: a randomised, parallel-group, open-label trial. *Int J Clin Pract*. 2011;65(4):397-407. doi:10.1111/j.1742-1241.2011.02656.x
21. Hong Y, Rohatagi S, Habtemariam B, Walker JR, Schwartz SL, Mager DE. Population exposure-response modeling of metformin in patients with type 2 diabetes mellitus. *J Clin Pharmacol*. 2008;48:696-707.
22. Williams-Herman D, Johnson J, Teng R, et al. Efficacy and safety of sitagliptin and metformin as initial combination therapy and as monotherapy over 2 years in patients with type 2 diabetes. *Diabetes Obes Metab*. 2010;12:442-451.
23. Feldman JM, Chapman BA. Radioimmunoassay of insulin in serum and plasma. *Clin Chem*. 1973;19(11):1250-1254. doi:10.1093/clinchem/19.11.1250
24. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 1985;28(7):412-419.
25. Bauer RJ. NONMEM tutorial part II: estimation methods and advanced examples. *CPT Pharmacometrics Syst Pharmacol*. 2019;8(8):538-556.
26. Erik Salomonsson. *Model-averaging using a cross-validation weighting method*. Unpublished MS-thesis, Uppsala University; 2019.
27. Aoki Y, Röshammar D, Hamrén B, Hooker AC. Model selection and averaging of nonlinear mixed-effect models for robust phase III dose selection. *J Pharmacokinet Pharmacodyn*. 2017;44(6):581-597.
28. Buatois S, Ueckert S, Frey N, Retout S, Mentré F. Comparison of model averaging and model selection in dose finding trials analyzed by nonlinear mixed effect models. *AAPS J*. 2018;20(3):56.
29. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2022. http://www.R-project.org/
30. Beal SL, Sheiner LB, Broeckmann A, Bauer RJ. *NONMEM User's Guides*. Icon Development Solutions; 2009.
31. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN) – a Perl module for NONMEM related programming. *Comput Methods Programs Biomed*. 2004;75:85-94.
32. Møller JB, Kristensen NR, Klim S, Karlsson MO, Ingwersen SH, Kjellsson MC. Methods for predicting diabetes phase III efficacy outcome from early data: superior performance obtained using longitudinal approaches. *CPT Pharmacometrics Syst Pharmacol*. 2014;3(7):e122.
33. Chalew S, Hempe JM. Caveats regarding the use of HbA1c for prediction of mean blood glucose. *Diabetologica*. 2008;51:903-904.
34. AHP CK, EAM C, Fabre D, Gattacceca F, Khier S. Prior information for population pharmacokinetic and pharmacokinetic/pharmacodynamic analysis: overview and guidance with a focus on the NONMEM PRIOR subroutine. *J Pharmacokinet Pharmacodyn*. 2020;47:431-446.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kunina H, Al-Mashat A, Chien JY, Garhyan P, Kjellsson MC. Optimization of trial duration to predict long-term HbA1c change with therapy: A pharmacometrics simulation-based evaluation. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:1443-1457. doi:10.1002/psp4.12854