Published in final edited form as:

Sci Transl Med. 2016 October 05; 8(359): 359ra130. doi:10.1126/scitranslmed.aaf9304.

Mechanistic Modeling of Hemoglobin Glycation and Red Blood Cell Kinetics Enables Personalized Diabetes Monitoring

Roy Malka^{1,2}, David M. Nathan^{3,4}, and John M. Higgins^{1,2,*}

¹Center for Systems Biology and Department of Pathology, Massachusetts General Hospital, Boston, MA, USA

²Department of Systems Biology, Harvard Medical School, Boston, MA, USA

³Diabetes Center, Massachusetts General Hospital, Boston, MA, USA

⁴Department of Medicine, Harvard Medical School, Boston, MA, USA

Abstract

The glycated hemoglobin assay (HbA1c) is essential for the diagnosis and management of diabetes because it provides the best estimate of a patient's average blood glucose (AG) over the preceding 2–3 months and is the best predictor of disease complications. However, there is substantial unexplained glucose-independent variation in HbA1c that makes AG estimation inaccurate and limits the precision of medical care for diabetics. The true AG of a non-diabetic and a poorly-controlled diabetic may differ by less than 15 mg/dL, but patients with identical HbA1c and thus identical HbA1c-based estimates of AG may have true AG that differs by more than 60 mg/dL. We combine a mechanistic mathematical model of hemoglobin glycation and red blood cell flux with large sets of intra-patient glucose measurements to derive patient-specific estimates of non-glycemic determinants of HbA1c including mean red blood cell age (M_{RBC}). We find that interpatient variation in derived M_{RBC} explains all glucose-independent variation in HbA1c. We then use our model to personalize prospective estimates of AG and reduce errors by more than 50% in four independent sets of more than 200 patients. The current standard of care provided AG estimates with errors > 15 mg/dL for 1 in 3 patients. Our patient-specific method reduced this error rate to 1 in 10. This personalized approach to estimating AG from HbA1c should improve medical care for diabetes using existing clinical measurements.

Introduction

Diabetes mellitus is a growing global health burden affecting about 400 million people worldwide (5). A person's glycated hemoglobin fraction (HbA1c) reflects the average concentration of glucose in the blood (AG) over the past 2–3 months and is the gold standard measure for establishing risk for diabetes-related complications in patients with type 1 or type 2 diabetes (6–8). An HbA1c greater than or equal to 6.5% is diagnostic for

*To whom correspondence should be addressed: john_higgins@hms.harvard.edu.

Author Contributions: RM, DMN, and JMH designed the study, performed the study, and wrote the paper.

Competing Interests: The authors are listed as inventors on a patent application related to this work submitted by Partners Healthcare.

diabetes, and the treatment goal for most people with diabetes is an HbA1c less than 7% (9). HbA1c is used to infer AG because continuous glucose measurements (CGM) are not routinely available (10). Glycation of hemoglobin occurs in two-step process including the condensation of glucose with the N-terminal amino group of the hemoglobin beta chain to form a Schiff base and the rearrangement of the aldimine linkage to a stable ketoamine (11). The kinetics of this slow non-enzymatic post-translational modification are thought to depend largely on the concentration of glucose, with previous studies establishing first-order kinetics (1, 11,12) and irreversibility of HbA1c formation (2,13). Other glycosylated forms are generated, but HbA1c is the clinically relevant glycation product and is therefore the focus of this analysis. Hemoglobin in older RBCs has had more time to become glycosylated, and older RBCs therefore have higher glycosylated fractions. HbA1c is measured as an average over RBCs of all ages in the circulation and therefore depends on both AG and M_{RBC} . Other factors may also be involved, including glucose gradients across the RBC membrane, intracellular pH, and glycation rate constants.

Here we study the glycemic and non-glycemic determinants of HbA1c. First, we dissect the contributions of glycemic and non-glycemic factors by deriving a mechanistic mathematical model quantifying the dependence of HbA1c on the chemical kinetics of hemoglobin glycation in a population of RBCs in dynamic equilibrium. Second, we personalize the model parameters for individual patients using existing CGM data. Third, we validate the personalized model's utility estimating future AG accurately from future HbA1c for each patient, and we compare the accuracy of the patient-specific model estimates of AG with those made using the current standard regression method.

Derivation of a mechanistic model of hemoglobin glycation and RBC flux

The process of HbA1c formation inside a single RBC can be described by the irreversible chemical reaction of hemoglobin (Hb) with glucose to form glycosylated hemoglobin (gHb) with rate k_g :

The rate of change in gHb can be modeled with a differential equation:

$$\frac{d}{dt}gHb(t) = k_g \cdot AG \cdot (tHb - gHb(t)) \quad (2)$$

tHb is the concentration of total hemoglobin in the RBC. The variable t is the time for the glycation reaction and is equivalent to the RBC's age. This model of glycation kinetics in general has been reported previously (11,12,14,15). We use it here to describe glycation in a single RBC. Equation (2) can be solved analytically and scaled by tHb to yield the *HbA1c* in an RBC of age t .

$$HbA1c(t) = \frac{gHb(t)}{tHb} = 1 - e^{(-k_g \cdot AG \cdot t)} + \frac{gHb(0)}{tHb} e^{(-k_g \cdot AG \cdot t)} \quad (3)$$

We use AG instead of a time-varying glucose to simplify this initial derivation, and we account for the effects of time-varying glucose below beginning in “Results: Measured variation in M_{RBC} is sufficient to explain all non-glycemic variation in HbA1c.” $gHb(0)$ is the concentration of glycated hemoglobin in the RBC when it is a reticulocyte and has just entered the circulation. Because $e^x \approx 1 + x$ when x is small, we can approximate Equation (3) with a linear function. By linearizing the exponential in this way, we can then average over the roughly uniformly-distributed ages of RBCs (t) in a patient’s circulation (3, 4,16) to provide the clinically measured HbA1c:

$$HbA1c = \underbrace{HbA1c(0)}_{\text{Intercept}} + \underbrace{[1 - HbA1c(0)] \cdot k_g \cdot M_{RBC}}_{\text{Slope}} \cdot AG \quad (4)$$

This linear relationship between AG and HbA1c has been reported in several studies such as the ADAG (1) (see Figure 1). These studies also show that direct estimation of AG based solely on HbA1c may be inaccurate, in part because of the imprecision and inaccuracies of the component measurements. In addition, significant glucose-independent variation has been described (17–20), including a linear relationship between M_{RBC} and HbA1c (2, 3). Regardless of the cause, an AG of 150 mg/dL may be associated with HbA1c anywhere between 5.5% and 8.0%, and HbA1c of 6.5% may reflect AG anywhere between 125 mg/dL and 175 mg/dL. See “Derivation of the AG-HbA1c linear regression from the physiological model of glycation” and “Synopsis of prior models of hemoglobin glycation” in Supplementary Methods for more detail.

Patient-specific differences in the AG-HbA1c relationship are caused by slope variation and not intercept variation

The scatter of data points away from the regression line as shown in Figure 1 represents patient-specific deviation from the regression model (Equation (4)) in terms of intercept or slope. Published estimates of the intercept $HbA1c(0)$ are small (~0.3%). It is difficult to measure accurately *in vivo*. One previous study measured HbA1c in transferrin-receptor-positive RBCs, which are typically reticulocytes and found results consistent with glycation in the bone marrow proceeding at a rate similar to that in the peripheral circulation (3). The range of potential interpatient variation in $HbA1c(0)$ (< 0.6%) is too limited to cause significant inter-patient variation in HbA1c., though it is difficult to measure accurately *in vivo*.

Analysis of HbA1c variance as a function of AG suggests that inter-patient variation in the slope rather than the intercept is a much more significant cause of glucose-independent variation in HbA1c. The amount of variation in HbA1c in Figure 1 is different at different AG levels, with apparently less variation in HbA1c at lower AG. Inter-patient variation in

slope will have a different effect from variation in intercept as illustrated in Figure 2A and Figure 2B. We analyze this relationship in more detail by calculating the HbA1c variance within 10 mg/dL intervals of AG. We first assess the possibility that there is of significant inter-individual variability in the intercept as illustrated in Figure 2A. Figure 2C simulates the effect of increasing variability in the intercept ($HbA1c(0)$ or reticulocyte HbA1c) when the slope is fixed. This hypothesized model of inter-patient differences in intercept (black line in Figure 2C) generates data (blue points in Figure 2C) that do not agree with the experimental data (red dots). Thus, inter-patient differences in reticulocyte HbA1c are unlikely to be responsible for glucose-independent variation seen in HbA1c. Figure 2B depicts the effect of inter-patient differences in the slope of the regression line. In Figure 2D, the correlation between the simulated and actual data in this case is very high. Overall, it is much more likely that inter-individual differences in slope are responsible for non-glycemic variation in HbA1c. See “Conditional variance of HbA1c controlling for AG” in Methods for more detail.

Measured variation in M_{RBC} is sufficient to explain all non-glycemic variation in HbA1c

We therefore focus on inter-patient variation in the slope: $\theta = [1 - HbA1c(0)] \cdot k_g \cdot M_{RBC}$. The first component ($[1 - HbA1c(0)]$), as discussed above, varies too little overall ($\sim 0.994 - 1.00$) to be a significant cause of glucose-independent variation in HbA1c. The second component (k_g) is not currently possible to measure directly *in vivo*, but it does not appear to vary between patients (21), and there is no reason to expect that a first-order chemical reaction rate would vary systematically between patients. The third component (M_{RBC}) has been measured with demanding and sophisticated labeling methods and has a mean of about 58 days and a standard deviation of 4.5–6.5 days, for a coefficient of variation

$$\left[CV(M_{RBC}) = \frac{std(M_{RBC})}{mean(M_{RBC})} \right] \text{ between 7.8\% and 11.2\% (3, 4).}$$

We can calculate a patient-specific corrected slope $\hat{\theta} = \frac{HbA1c - HbA1c(0)}{AG}$ at the time of a specific HbA1c measurement using AG determined from large intra-patient continuous glucose monitoring (CGM) data sets. (We use the symbol $\hat{\theta}$ to represent an estimate of the true patient-specific slope θ which cannot be measured directly.) AG is calculated from CGM data using a weighted average of individual glucose measurements because glucose levels in the blood immediately prior to the HbA1c measurement influence the glycation levels in RBCs of all ages, while more distant glucose levels influence only those RBCs which are old enough to have been in the circulation at that time (22). See “Calculation of AG and coverage from continuous glucose monitoring (CGM) data” in Methods for more detail.

We calculated $\hat{\theta}$ for 36 distinct patients at our hospital and find $CV(\hat{\theta}) = 10.8\%$, within the range of variation that can be explained entirely by inter-patient variation in M_{RBC} . In three additional independent sets of 339 patients we find $CV(\hat{\theta})$ equal to 8.8% (30 patients), 9.4% (234 patients), and 9.9% (75 patients). Analysis of all four populations suggest that glucose-

independent variation in HbA1c can be explained entirely by variation in M_{RBC} . (See “Patient populations” in Methods for more detail.) Figure 3 further shows that if $[1 - HbA1c(0)]$ and k_g are constant, all measured glucose-independent variation in HbA1c in the ADAG (1) study data can be accounted for by simulating variation in M_{RBC} with magnitude equivalent to that previously measured (3, 4). If either or both of the other two slope components ($[1 - HbA1c(0)]$ and k_g) vary significantly, they must be strongly negatively correlated with M_{RBC} , or else $CV(\hat{\theta})$ would be much greater than $CV(M_{RBC})$.

Personalizing the model for increased accuracy of prospectively-estimated AG

Because $CV(M_{RBC}) \approx CV(\hat{\theta})$, we can estimate a patient's M_{RBC} using published estimates

of $[1 - HbA1c(0)]$ and k_g : $\widehat{M_{RBC}} = \frac{\hat{\theta}}{[1 - HbA1c(0)] \cdot k_g}$. (We use the symbol $\widehat{M_{RBC}}$ to represent an estimate of the patient's true M_{RBC} .) Recent work directly measuring (4) and modeling (23–25) M_{RBC} suggests it is tightly regulated within individuals, and we therefore hypothesized that we can derive a patient-specific $\widehat{M_{RBC}}$ at one point in time and use it prospectively in Equation (4) to improve the accuracy of future AG estimates made from

future HbA1c: $AG = \frac{HbA1c - HbA1c(0)}{[1 - HbA1c(0)] \cdot k_g \cdot \widehat{M_{RBC}}}$. See Figure 4 for two examples. Note that while we present a linearized model and analysis here for clarity, we obtain very similar results with an exact numerical solution. See “Numerical solution” in Methods for more detail.

We evaluated AG estimates for 16 HbA1c measurements from 9 distinct adult patients at Massachusetts General Hospital. The patient-specific model reduced the median absolute error in estimated AG from more than 15 mg/dL to less than 5 mg/dL, an error reduction of more than 66%. It is most informative to compare AG predictions where the model-based method differs from the current standard regression-based approach. Figure 5 compares errors in predicted AG when the two methods differ by at least 10 mg/dL and confirms the superior accuracy of model-based AG prediction in 3 additional independent patient populations totaling more than 300 individuals. (See “Patient Populations” in Methods for further detail.) Figure 5 shows that in each of these patient populations, the model-based approach reduced the median absolute error in estimated AG by at least 50%. The substantial improvement in accuracy achieved by the model is highlighted by the fact that for all 4 independent study groups, the 75th percentile of the model-based estimation error is less than the median error for the current regression-based prediction.

The difference in AG between a non-diabetic ($HbA1c < 6.5\%$) and a diabetic with sub-optimal disease control ($HbA1c > 7.0\%$) can be ~15 mg/dL (9). Thus, errors of 15 mg/dL or less in estimated AG could mislead clinicians and patients and compromise patient care and optimal management of long-term risk of complications. Across our 4 sets of patients, the current regression method generated AG estimation errors greater than 15 mg/dL for about 1 patient in 3 (31.4%), while the patient-specific model produced errors this large for only 1

patient in 10 (9.6%). An error in estimated AG of 28.7 mg/dL is equivalent to an error of ~1.0% in HbA1c. The current regression method generated AG estimation errors at least this large for 1 patient in 13, and the patient-specific method for only 1 patient in 220.

Real-time estimates of HbA1c for patients with CGM

A method to estimate HbA1c from CGM in real-time would provide useful feedback for patients trying to optimize glucose management between clinic visits. Patients are already accustomed to thinking about the quality of their glucose control in terms of HbA1c. Previous studies have developed sophisticated methods to estimate HbA1c by combining prior HbA1c levels with multipoint profiles of self-monitored glucose (26). These methods have generated impressive results with sparse measurements of glucose, achieving correlation between estimated and measured HbA1c as high as 0.76, with estimates of HbA1c deviating from measured HbA1c by an average of as little as 0.5%. For example, if the measured HbA1c was 7.0%, this method would typically estimate an HbA1c between 6.5% and 7.5%. The patient-specific model presented here has two advantages over these other approaches in that it controls for patient-specific variation in non-glycemic factors influencing HbA1c, and it also takes advantage of the vastly richer glucose characterization provided by CGM. It is therefore not surprising that our patient-specific method estimated HbA1c with significantly higher accuracy. We estimated HbA1c for 200 patients in our study populations and found a correlation of 0.90 and an average deviation from measured HbA1c of 0.3%, meaning for example that if the measured HbA1c was 7.0%, our method would typically estimate an HbA1c between 6.7% and 7.3%. Given that analytic variation in HbA1c assays would be expected to generate an uncertainty range of at least 6.9% – 7.1% (27), the patient-specific model thus makes a significant advance toward optimal estimation.

Discussion

We have developed a model of glycation kinetics and derived a patient-specific correction factor ($\widehat{M}_{RBC}$) to improve the accuracy of AG estimation from HbA1c. That $\widehat{M}_{RBC}$ improves the accuracy of HbA1c-derived AG is not entirely unexpected; however, the prospective utility of $\widehat{M}_{RBC}$ to improve accuracy suggests that it is consistent in individuals over time. Optimal diagnosis and management of diabetes requires an accurate estimate of AG. The improvement in AG calculation afforded by our model should improve medical care and provide for a personalized approach to determining AG from HbA1c. The model would require one pair of CGM-measured AG and an HbA1c measurement that would be used to determine the patient's $\widehat{M}_{RBC}$. $\widehat{M}_{RBC}$ would then be used going forward to refine the future AG calculated based on HbA1c.

Our study follows a rich history of mathematical modeling in diabetes which has revealed important pathophysiologic insights with great potential to inform early diagnosis and effective treatment (28–32), as well as more recent studies modeling other aspects of diabetes, including models classifying diabetes sub-types by integrating medical record data (33), predicting near-term glucose based on dietary intake (34, 35), identifying patients at

high short-term risk of diabetes (36), controlling for non-biologic measurement errors (37), and optimizing treatment strategies using fasting plasma glucose measurements (38).

Future work is needed to define the duration of CGM required for sufficient calibration of $\widehat{M}_{RBC}$. Our analysis of these 4 data sets suggests that no more than 30 days is required, and we find statistically significant improvement with as little as 21. If a patient's monthly glucose averages are stable, then the prior one month would be sufficient, and if the patient's weekly glucose averages are stable then even one week of CGM might be sufficient. The patients in our four study populations all received regular routine medical care and were generally healthy. Follow-up study is necessary to assess model accuracy in the setting of more acute and serious co-morbid disease, including conditions known to affect RBC turnover. We note that our patient-specific model may be particularly helpful in situations where plasma glucose is likely to deviate significantly from the longer-term average reflected in HbA1c, such as optimization of treatment for a patient recently-diagnosed with diabetes (38). By controlling for patient-specific non-glycemic factors, the model should improve the clinical utility of HbA1c to provide more information regarding average glucose levels.

Our patient-specific model provides a substantial improvement in the accuracy of AG estimates, but its estimates are not perfect. When used to estimate AG from HbA1c, the model's sensitivity to variation in true AG will depend on the accuracy of the input HbA1c and CGM. HbA1c is typically rounded to multiples of 0.1%, meaning the model is theoretically sensitive to changes of 2–3 mg/dL in AG, and higher resolution HbA1c measurements would increase the model's sensitivity. Analytic variation in current HbA1c measurements is reported to be ~3% (27), and this variation alone would be expected to generate AG estimation errors of ~7 mg/dL. The median error in the model-based estimate of AG may thus be as low as is possible given current HbA1c measurement methods, but errors for some individual patients are higher, and the source of those errors warrants further investigation. Individual CGM measurements have a reported error of about 10% (39), but because AG is an average over thousands of separate CGM measurements with frequent calibration, the expected error in AG is about 0.1%. Systematic bias in CGM measurement or calibration would reduce the accuracy of AG estimation, and advances in CGM technology to minimize bias would increase model sensitivity. Other potential sources of error beyond the model include incomplete CGM data and fluctuations in M_{RBC} within an individual. Given the small median estimation errors we find, the magnitude of variation in those quantities must be small on average in all four groups of study patients, but it will be important and informative to explore those possible explanations for the few patients with much larger estimation errors.

Although direct measurement of M_{RBC} was not carried out, the inter-patient variation in $\widehat{M}_{RBC}$ and the intra-patient stability of $\widehat{M}_{RBC}$ are consistent with what has been shown for M_{RBC} in other studies, both those directly measuring M_{RBC} (3, 4) and those providing model-based estimates (24, 40). Moreover, the number of factors that might be involved in the differences between measured and calculated AG is limited, and factors such as glycation rates or intracellular pH would not be practical to measure. Future studies that

directly measure M_{RBC} will be required to increase confidence that inter-individual variability in M_{RBC} is the operant factor and improvements in methods to measure M_{RBC} directly will be important to determine more precisely how much non-glycemic variation in HbA1c remains unexplained. In the meantime, the correction factor we have identified appears to be sufficient to improve the accuracy of the AG estimation from HbA1c. More generally, our study demonstrates how clinical accuracy can be enhanced in a patient-specific manner by combining large intra-patient data sets with mechanistic dynamic models of physiology.

Materials and Methods

Study design

Our goal was to develop a more accurate method for estimating AG from HbA1c by adjusting for inter-patient variation in non-glycemic factors that help determine HbA1c. Our analysis required 3 steps.

1. We first quantified the factors determining AG-independent variation in HbA1c by developing a mechanistic mathematical model describing how HbA1c depends on the chemical kinetics of hemoglobin glycation in a population of RBCs at dynamic equilibrium.
2. We then combined the model with CGM measurements to personalize the model for each patient.
3. Using the patient-specific model in combination with one set of CGM and HbA1c, we derived a patient's $\widehat{M}_{RBC}$ and used it prospectively to estimate AG from future HbA1c. We then compared the accuracy of patient-specific model estimates of AG with those made using the current standard regression method.

Conditional variance of HbA1c controlling for AG

The amount of variation in HbA1c in Figure 1 increases at higher AG levels, with apparently less variation in HbA1c at lower AG. We now analyze this relationship in more detail. We calculate the HbA1c variance in the ADAG data conditioned on AG. This conditional variance calculation is similar to conditional expectation calculations. Both involve averaging over all measurements that have corresponding AG levels within intervals (of 10 mg/dL in this case). Instead of averaging HbA1c itself as in conditional expectation, we now average the squared deviation of each HbA1c measurement from the mean:

$$(HbA1c - \{HbA1c\})^2.$$

The scatter of data around the AG-HbA1c linear regression line may reflect inter-individual variation in the slope or the intercept the regression model, or both. Figure 2 illustrates the different effects the intercept and slope variation would be expected to have on the HbA1c conditional variance. We first assess the possibility of significant inter-individual variability in the intercept as illustrated in Panel A of Figure 2. Figure 2C shows a simulation of the effect of increasing variability in reticulocyte HbA1c (equivalent to variation in the intercept β) when M_{RBC} is fixed. This hypothesized model (black line in Figure 2C) of variation in

the intercept β generates data (blue points) that do not agree with the experimental data (correlation coefficient of $r^2 = -0.05$). Thus, variation in reticulocyte HbA1c is unlikely to explain the observed scatter of HbA1c around the regression line.

Figure 2B illustrates the expected effect of variation in the slope of the regression line. The correlation between the conditional variance and AG calculated from simulated HbA1c and the ADAG data is $r_s^2 = 0.94$. Similarly, the correlation in the ADAG data is $r_d^2 = 0.65$ (Figure 2D). Note that in the ADAG data, out of 507 samples, there are 2 outliers both with AG in the range of 110–120 mg/dL, creating a single bin for the calculation of conditional variance. In Figure 2D, we remove these two samples, increasing r_d^2 to 0.80. Overall, it is much more likely that inter-individual variation in the regression slope is responsible for variation observed in the AG-HbA1c relationship. Because the model is linear, we can calculate the variance around the regression line analytically (Equation (5) below and black dotted line in Figure 2 Panels C and D).

In the following variance calculations, conditional expectation is taken with respect to the RBC age across the population of RBCs in one patient's circulation, as well as with respect to the M_{RBC} for an individual patient across the population of individuals. We assume the

initial glycation fraction $\frac{gHb(0)}{tHb}$ is a random variable. To simplify the expression, we treat the glycation rate as a constant.

$$V\{HbA1c|AG\} = V\left\{\frac{gHb(0)}{tHb}\right\} \cdot [1 - 2\{M_{RBC}\} \cdot k_g \cdot AG] + AG^2 \cdot k_g^2 \cdot \{M_{RBC}\}^2 \cdot V\left\{\frac{gHb(0)}{tHb}\right\} + \left\{\frac{gHb(0)}{tHb}\right\}^2 \cdot V\{M_{RBC}\} + V\left\{\frac{gHb(0)}{tHb}\right\} \cdot V\{M_{RBC}\} \quad (5)$$

Note that the contribution of increased variability in the reticulocyte HbA1c has the approximate effect of an 'additive noise' on the total variance, because in terms of numerical values, it has little dependence on AG. However, when we consider the theoretical case of no variability in M_{RBC} , a negative slope emerges as seen in Figure 2C, contradicting the empirical data.

Calculation of AG and coverage from continuous glucose monitoring (CGM) data

The AG that determines HbA1c is a weighted average of glucose levels prior to the HbA1c measurement (22). As discussed above, the clinically measured HbA1c is an average of single-RBC HbA1c over the ages of RBCs in a patient's blood sample. The RBC ages are assumed to be uniformly distributed between 0 and $2 \cdot M_{RBC}$. The blood glucose level on the day prior to the HbA1c measurement affects the HbA1c of almost every RBC in the blood sample. The blood glucose levels measured much earlier and closer to $2 \cdot M_{RBC}$ days prior to the HbA1c measurement will only affect the small fraction of the very oldest RBCs

still in circulation. The AG from CGM for the linearized model (Equation (4)) can be calculated using the following equation:

$$AG = \frac{1}{2 \cdot M_{RBC}} \int_0^{2 \cdot M_{RBC}} \left(\frac{1}{t} \int_{-t}^0 glucose(\tau) d\tau \right) dt \quad (6)$$

When full CGM data is not available, it is therefore more valuable to have recent CGM measurements. Defining $I(t)$ as 1 if there is CGM data within the 5 minutes prior to t and 0 if not, we can calculate the fractional coverage of CGM data during the desired time period using a related equation:

$$Coverage = \frac{1}{2 \cdot M_{RBC}} \int_0^{2 \cdot M_{RBC}} \left(\frac{1}{t} \int_{-t}^0 I(\tau) d\tau \right) dt \quad (7)$$

See below for numerical calculation of AG from CGM without assuming the linear approximation.

Model simulation

For the simulation in Figure 3, we assume that RBC lifespan is normally distributed among different individuals with mean and variance estimated from prior publications. Note that a specific assumption on the parametric distribution of M_{RBC} among individuals is necessary only in the simulation and is not required for the analytic calculations. We take M_{RBC} to be normally distributed across individuals but find that a gamma distribution yields similar results. The age distribution of RBCs within an individual is assumed to be uniform, with cell ages between 0 and $2 \cdot M_{RBC}$. We also assume that the glycation rate is essentially constant as has been demonstrated previously (3,21). See “Glycation rate and M_{RBC} ” in Supplementary Methods for more detail.

We then use the fitted average parameter values (slope and intercept) obtained from the corresponding linear regression line, and the model reconstructs the scatter of data points around the regression line adding variability in M_{RBC} equivalent to that previously measured (41). In the simulations, we use a value of 0.001 for the standard deviation of reticulocyte HbA1c for the ADAG data. These values were adopted from the measurements of variation of HbA1c in reticulocytes (42). For k_g we allow a CV of 1%, though a CV of 5% with constant M_{RBC} will reconstruct the variation around the regression line, as expected from the functional form of the model.

We first assume that AG is estimated with high accuracy as a result of the large number of measurements included in the average. Indeed, in the ADAG study (1) each AG value is calculated using more than 250 samples over the course of 3 months. The standard error

(SE) is $SE < \frac{SD}{\sqrt{250}} \approx \frac{SD}{15}$. Thus, even if the level of variability in a single glucose measurement is extremely high, for example: $SD = 30$ (mg/dL), the resulting coefficient of variation will be less than 3% for all AG values in the ADAG data. The SD for the full

ADAG data set is 39 (mg/dL), and 8 (mg/dL) when restricting to the non-diabetic patients, and thus the uncertainty in AG is expected to be less than 1 (mg/dL).

Numerical solution

The physiologic model for glycation can be solved numerically without making a linear approximation. We start with the differential equation model including a time-varying glucose concentration ($G(t)$):

$$\frac{d}{dt}gHb(t) = k_g \cdot G(t) \cdot (tHb - gHb(t)) \quad (8)$$

This equation can be integrated numerically to provide the HbA1c in an RBC of age t :

$$HbA1c(t) = HbA1c(0) + \frac{k_g}{tHb} \cdot \int_0^t G(\tau) \cdot (tHb - gHb(\tau)) d\tau \quad (9)$$

The clinical HbA1c measurement is the average over a uniform distribution of RBC ages ranging between 0 and $2 \cdot M_{RBC}$:

$$HbA1c = \frac{1}{2 \cdot M_{RBC}} \int_0^{2 \cdot M_{RBC}} \left[HbA1c(0) + \frac{k_g}{tHb} \cdot \int_0^t G(\tau) \cdot (tHb - gHb(\tau)) d\tau \right] dt \quad (10)$$

Given sufficient CGM data to define $G(\tau)$ and a concurrent HbA1c measurement, the above equation can be solved numerically for M_{RBC} to provide a patient-specific $\widehat{M_{RBC}}$. For the model-based prediction of AG from HbA1c, the patient's $\widehat{M_{RBC}}$ is used, and the following equation is solved numerically for AG:

$$HbA1c = \frac{1}{2 \cdot \widehat{M_{RBC}}} \int_0^{2 \cdot \widehat{M_{RBC}}} \left[HbA1c(0) + AG \cdot \frac{k_g}{tHb} \cdot \int_0^t (tHb - gHb(\tau)) d\tau \right] dt \quad (11)$$

Results shown in the paper use the linear approximation. We replicated all analysis with numerical solutions and reach very similar conclusions.

Patient populations

Patient set #1—We analyzed existing continuous glucose monitoring (CGM) data from 36 adult patients at Massachusetts General Hospital (MGH) under a research protocol approved by the Partners Healthcare Institutional Review Board. CGM measurements were made with Dexcom G4 continuous glucose monitors (Dexcom, Inc., San Diego, CA). HbA1c was measured either on a Roche COBAS instrument (Roche Diagnostics, Indianapolis, Indiana) or a BIO-RAD Variant II Turbo (BIORAD, Hercules, California). 36 patients had at least one HbA1c measurement with concurrent CGM covering a period of time equivalent to the most recent 30 days prior to the HbA1c. See “Calculation of AG and coverage from

continuous glucose monitoring (CGM) Data” above for more detail. 9 of those 36 individuals had a total of 16 additional future HbA1c measurements with concurrent CGM covering a period of time equivalent to the most recent 30 days prior to the HbA1c. Those 16 future HbA1c measurements were used to validate the accuracy of the model-based AG estimation.

Patient set #2—Data for the second, third, and fourth patient populations was made available by the Jaeb Center for Health Research, a coordinating center for multi-center clinical trials and epidemiologic research. Their studies of diabetic control reported CGM and HbA1c measurements in patients and generously included raw data, enabling us to test our model and hypothesis in three additional independent data sets. Our “Patient Set #2” comes from a study entitled “Effect of Metabolic Control at Onset of Diabetes on Progression of Type 1 Diabetes” (<http://direcnet.jaeb.org/Studies.aspx?RecID=165>). The original purpose of this study was to investigate the impact of intensive metabolic control from the onset of diabetes on preservation of C-peptide secretion. This study, conducted between November 2008, and October 2013, included patients aged 6–46. 30 patients had at least one HbA1c measurement with concurrent CGM covering a period of time equivalent to the most recent 45 days prior to the HbA1c. 23 of those 30 individuals had a total of 79 additional future HbA1c measurements with concurrent CGM covering a period of time equivalent to the most recent 45 days prior to the HbA1c. Those 79 future HbA1c measurements and corresponding CGM were used to validate the accuracy of the model-based AG estimation.

Patient set #3—The data for this third patient population comes from a study entitled, “A Randomized Clinical Trial to Assess the Efficacy of Real-Time Continuous Glucose Monitoring in the Management of Type 1 Diabetes” (<http://diabetes.jaeb.org/RTCGMRCTProtocol.aspx>). This study was designed to compare continuous versus standard intensive glucose monitoring in three age groups (>25, 15–24, 8–14), of intensively-treated type 1 diabetics having high glycated hemoglobin (HbA1c) 7.0%–10.0%. 234 patients had at least one HbA1c measurement with concurrent CGM covering a period of time equivalent to the most recent 45 days prior to the HbA1c. 155 of those 234 individuals had a total of 276 additional future HbA1c measurements with concurrent CGM covering a period of time equivalent to the most recent 45 days prior to the HbA1c. Those 276 future HbA1c measurements and corresponding CGM were used to validate the accuracy of the model-based AG estimation.

Patient set #4—The data for this fourth patient population comes from a study entitled, “A Randomized Clinical Trial to Assess the Efficacy and Safety of Real-Time Continuous Glucose Monitoring in the Management of Type 1 Diabetes in Young Children (4 to <10 Year Olds)” (<http://direcnet.jaeb.org/Studies.aspx?RecID=162>). This study was designed to assess the efficacy of CGM in young children (4–10 years old) in terms of tolerability, safety, and effect on quality of life with type 1 diabetes. 37 patients had at least one HbA1c measurement with concurrent CGM covering a period of time equivalent to the most recent 45 days prior to the HbA1c. See “Calculation of AG and Coverage from Continuous Glucose Monitoring (CGM) Data” above for more detail. 31 of those 37 individuals had a

total of 69 additional future HbA1c measurements with concurrent CGM covering a period of time equivalent to the most recent 45 days prior to the HbA1c. Those 69 future HbA1c measurements and corresponding CGM were used to validate the accuracy of the model-based AG estimation.

Computational modeling and statistical analysis

Modeling and statistical analysis were performed in MATLAB (MathWorks, Inc., Natick, MA).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Carlo Brugnara, Frank Bunn, Norman Mazer, Lorette Noiret, and Anwesha Chaudhury for useful discussions and four anonymous reviewers for very helpful suggestions. We thank Curt Rohlfing for help with the DCCT data; Denise Richards and Tiffany Soper for help with MGH CGM data, and the JAEB for making their data (used as patient sets 2, 3, and 4) publically available. The ADAG study was supported by research grants from the American Diabetes Association and European Association for the Study of Diabetes, and by other financial support from Abbott Diabetes Care, Bayer Healthcare, GlaxoSmithKline, sanofi aventis Netherlands, Merck, Lifescan, and Medtronic Minimed, and with supplies and equipment provided by Medtronic Minimed, Lifescan, and Hemocue.

Funding:

JMH and RM were supported by an NIH Director's New Innovator Award (DP2DK098087) and by a research grant from Abbott Diagnostics. None of the funding agencies had any input on study design or decision to publish.

Abbreviations

AG	average glucose
CGM	continuous glucose monitoring
HbA1c	clinical glycated hemoglobin assay
M_{RBC}	mean (average) red blood cell age
RBC	red blood cell

References

1. Nathan DM, et al. Translating the A1C assay into estimated average glucose values. *Diabetes care*. 2008; 31:1473. [PubMed: 18540046]
2. Higgins PJ, Bunn HF. Kinetic-Analysis of the Non-Enzymatic Glycosylation of Hemoglobin. *Journal of Biological Chemistry*. 1981; 256:5204. [PubMed: 7228877]
3. Cohen RM, et al. Red cell life span heterogeneity in hematologically normal people is sufficient to alter HbA1c. *Blood*. Nov.2008 112:4284. [PubMed: 18694998]
4. Khera PK, et al. Use of an oral stable isotope label to confirm variation in red blood cell mean age that influences HbA1c interpretation. *Am J Hematol*. 2015; 90:50. [PubMed: 25293624]
5. International Diabetes Foundation. IDF Diabetes Atlas. 2015. (<http://www.idf.org/diabetesatlas>)
6. DCCT Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med*. 1993; 329:977. [PubMed: 8366922]

7. U. P. D. S. G. (UKPDS). et al. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *The Lancet*. 1998; 352:837.
8. Sacks DB. Hemoglobin A1c in diabetes: panacea or pointless? *Diabetes*. 2013; 62:41. [PubMed: 23258914]
9. American Diabetes Association. Standards of medical care in diabetes—2010. *Diabetes care*. 2010; 33:S11. [PubMed: 20042772]
10. Crane PK, et al. Glucose levels and risk of dementia. *New England Journal of Medicine*. 2013; 369:540. [PubMed: 23924004]
11. Bunn HF, Haney DN, Kamin S, Gabbay K, Gallop P. The biosynthesis of human hemoglobin A1c. Slow glycosylation of hemoglobin in vivo. *Journal of Clinical Investigation*. 1976; 57:1652. [PubMed: 932199]
12. Beach KW. A theoretical model to predict the behavior of glycosylated hemoglobin levels. *Journal of theoretical biology*. 1979; 81:547. [PubMed: 537384]
13. Ladyzynski P, et al. Validation of hemoglobin glycation models using glycemia monitoring in vivo and culturing of erythrocytes in vitro. *Annals of biomedical engineering*. 2008; 36:1188. [PubMed: 18449641]
14. Lledó-García R, Mazer NA, Karlsson MO. A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects. *Journal of pharmacokinetics and pharmacodynamics*. 2013; 40:129. [PubMed: 23307170]
15. Osterman-Golkar SM, Vesper HW. Assessment of the relationship between glucose and A1c using kinetic modeling. *Journal of Diabetes and its Complications*. 2006; 20:285. [PubMed: 16949515]
16. Franco RS. The measurement and importance of red cell survival. *Am J Hematol*. Feb.2009 84:109. [PubMed: 19025796]
17. Rohlfing C, et al. Biological variation of glycohemoglobin. *Clinical Chemistry*. Jul.2002 48:1116. [PubMed: 12089188]
18. Gould BJ, Davie SJ, Yudkin JS. Investigation of the mechanism underlying the variability of glycated haemoglobin in non-diabetic subjects not related to glycaemia. *Clin Chim Acta*. Apr.1997 260:49. [PubMed: 9101100]
19. Cohen RM, Holmes YR, Chenier TC, Joiner CH. Discordance between HbA(1c) and fructosamine – Evidence for a glycosylation gap and its relation to diabetic nephropathy. *Diabetes Care*. Jan. 2003 26:163. [PubMed: 12502674]
20. Yudkin JS, et al. Unexplained Variability of Glycated Hemoglobin in Nondiabetic Subjects Not Related to Glycemia. *Diabetologia*. Apr.1990 33:208. [PubMed: 2347434]
21. Ladyzynski P, et al. Hemoglobin glycation rate constant in non-diabetic individuals. *Annals of biomedical engineering*. 2011; 39:2721. [PubMed: 21796500]
22. Tahara Y, Shima K. KINETICS OF HBA(1C), GLYCATED ALBUMIN, AND FRUCTOSAMINE AND ANALYSIS OF THEIR WEIGHT-FUNCTIONS AGAINST PRECEDING PLASMA-GLUCOSE LEVEL. *Diabetes Care*. Apr.1995 18:440. [PubMed: 7497851]
23. Higgins JM, Mahadevan L. Physiological and pathological population dynamics of circulating human red blood cells. *Proceedings of the National Academy of Sciences*. 2010; 107:20587.
24. Patel HH, Patel HR, Higgins JM. Modulation of red blood cell population dynamics is a fundamental homeostatic response to disease. *American journal of hematology*. 2015; 90:422. [PubMed: 25691355]
25. Golub MS, Hogrefe CE, Malka R, Higgins JM. Developmental plasticity of red blood cell homeostasis. *Am J Hematol*. May.2014 89:459. [PubMed: 24415575]
26. Kovatchev BP, Flacke F, Sieber J, Breton MD. Accuracy and Robustness of Dynamical Tracking of Average Glycemia (A1c) to Provide Real-Time Estimation of Hemoglobin A1c Using Routine Self-Monitored Blood Glucose Data. *Diabetes Technol Ther*. May.2014 16:303. [PubMed: 24299302]
27. Lenters-Westra E, Slingerland RJ. Six of Eight Hemoglobin A(1c) Point-of-Care Instruments Do Not Meet the General Accepted Analytical Performance Criteria. *Clinical chemistry*. Jan.2010 56:44. [PubMed: 19926777]

28. Matthews DR, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. Jul.1985 28:412. [PubMed: 3899825]
29. El-Khatib FH, Russell SJ, Nathan DM, Sutherland RG, Damiano ER. A Bihormonal Closed-Loop Artificial Pancreas for Type 1 Diabetes. *Science Translational Medicine*. Apr.2010 2
30. Wang YQ, Dassau E, Doyle FJ. Closed-Loop Control of Artificial Pancreatic beta-Cell in Type 1 Diabetes Mellitus Using Model Predictive Iterative Learning Control. *IEEE Trans Biomed Eng*. Feb.2010 57:211. [PubMed: 19527957]
31. Bergman RN, Phillips LS, Cobelli C. Physiologic Evaluation Of Factors Controlling Glucose-Tolerance in Man – Measurement of Insulin Sensitivity and Beta-Cell Glucose Sensitivity from the Response to Intravenous Glucose. *Journal of Clinical Investigation*. 1981; 68:1456. [PubMed: 7033284]
32. Bergman RN. Toward Physiological Understanding of Glucose-Tolerance – Minimal-Model Approach. *Diabetes*. Dec.1989 38:1512. [PubMed: 2684710]
33. Ali R, Hussain J, Siddiqi MH, Hussain M, Lee S. H2RM: A Hybrid Rough Set Reasoning Model for Prediction and Management of Diabetes Mellitus. *Sensors*. Jul.2015 15:15921. [PubMed: 26151207]
34. Georga EI, Protopappas VC, Polyzos D, Fotiadis DI. Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models. *Medical & Biological Engineering & Computing*. Dec.2015 53:1305. [PubMed: 25773366]
35. Zecchin, C., Facchinetti, A., Sparacino, G., Cobelli, C. *Artificial Neural Networks*. 2nd. Cartwright, H., editor. Vol. 1260. 2015. p. 245-259.
36. Huang CL, et al. Using Hemoglobin A1C as a Predicting Model for Time Interval from Pre-Diabetes Progressing to Diabetes. *Plos One*. Aug.2014 9
37. Given JE, Bunting BP, Coates VE, O’Kane MJ. Measurement error in estimated average glucose: a novel approach. *Clinical Chemistry and Laboratory Medicine*. Jul.2014 52:E147. [PubMed: 24558229]
38. Barua A, Acharya J, Ghaskadbi S, Goel P. The relationship between fasting plasma glucose and HbA(1c) during intensive periods of glucose control in antidiabetic therapy. *J Theor Biol*. Dec. 2014 363:158. [PubMed: 25158164]
39. Damiano ER, et al. A comparative effectiveness analysis of three continuous glucose monitors: the Navigator, G4 Platinum, and Enlite. *Journal of diabetes science and technology*. Jul.2014 8:699. 2014. [PubMed: 24876423]
40. Higgins JM, Mahadevan L. Physiological and pathological population dynamics of circulating human red blood cells. *Proc Natl Acad Sci U S A*. Nov.2010 107:20587. [PubMed: 21059904]
41. Hoelzel W, et al. IFCC reference system for measurement of hemoglobin A1c in human blood and the national standardization schemes in the United States, Japan, and Sweden: a method-comparison study. *Clinical chemistry*. 2004; 50:166. [PubMed: 14709644]
42. Cohen RM, Franco RS, Joiner CH. Is poor glycemic control associated with reduced red blood cell lifespan? *Diabetes care*. 2004; 27:1013. [PubMed: 15047667]
43. Hempe JM, Gomez R, McCarter RJ, Chalew SA. High and low hemoglobin glycation phenotypes in type 1 diabetes: a challenge for interpretation of glycemic control. *Journal of diabetes and its complications*. 2002; 16:313. [PubMed: 12200073]
44. Higgins PJ, Bunn HF. Kinetic analysis of the nonenzymatic glycosylation of hemoglobin. *Journal of Biological Chemistry*. 1981; 256:5204. [PubMed: 7228877]
45. Mortensen HB, Vølund A, Christophersen C. Glucosylation of human haemoglobin a. dynamic variation in HbA 1c described by a biokinetic model. *Clinica chimica acta*. 1984; 136:75.
46. Jansen H, et al. Determinants of HbA1c in nondiabetic Dutch adults: genetic loci and clinical and lifestyle parameters, and their interactions in the lifelines cohort study. *Journal of internal medicine*. 2013; 273:283. [PubMed: 23121487]
47. Malka R, Delgado FF, Manalis SR, Higgins JM. In vivo volume and hemoglobin dynamics of human red blood cells. *PLoS Comput Biol*. 2014; 10

48. Piva E, Brugnara C, Chiandetti L, Plebani M. Automated reticulocyte counting: state of the art and clinical applications in the evaluation of erythropoiesis. *Clinical Chemistry and Laboratory Medicine*. Oct.2010 48:1369. [PubMed: 20666695]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Summary

Effective medical care for diabetes crucially depends on accurate estimates of a patient's recent average blood glucose. We developed a mathematical model integrating known mechanisms of hemoglobin glycation and RBC flux and combined it with existing routine clinical measurements to make personalized estimates of average glucose that reduce diagnostic errors by more than 50% compared to the current standard method. This method should personalize and improve care for diabetes.

Figure 1. Linear relationship between AG and HbA1c in diabetic and non-diabetic subjects
 Data from the ADAG study (1): 507 subjects with type 1 (blue dots, n=268) and type 2 (green dots, n=159) diabetes, as well as non-diabetic subjects (red dots, n=80). The lower dashed line is the 6.5% threshold for initial diagnosis of diabetes, and the upper dashed line is the 7% treatment target.

Figure 2. The variance of HbA1c increases with AG and suggests that inter-patient differences in slope are more important than differences in intercept for determining non-glycemic variation in HbA1c

A deviation from the regression line in Figure 1 can be explained by a patient-specific line that has a different intercept, or a different slope, or both. **Panel A** shows that variation in the intercept (reticulocyte glycation fraction) alone will lead to fixed deviations from the regression line that are independent of the AG. **Panel B** shows in contrast that variation in the slope will lead to increased variance as AG increases. We can use the model to simulate the effect of each type of inter-patient difference and compare the simulation results with actual data. **Panel C** shows the effect of simulated inter-patient differences in the intercept (i.e., the reticulocyte glycation fraction). r_d^2 is the rank correlation coefficient for the raw ADAG data (hence the “d” in r_d^2) shown as red dots in both (C) and (D) and its correlation is the same in each panel because it is independent of a model of the source of variation. r_s^2 (“s” for simulation) in panel (C) is the correlation for the blue dots in (C) which show the simulated effect of inter-patient differences in intercept on variance in AG. Their correlation (0.13) is quite different from r_d^2 (0.8). The simulations show that inter-patient differences in intercept would generate data (blue dots) whose trend (black line) is quite different from that actually seen in the real data (red dots). **Panel D** shows the effect of inter-patient differences in slope on AG variance. The raw data (red dots) is repeated from Panel C for comparison and has the same correlation coefficient (r_d^2). r_s^2 (“s” for simulation) in panel (D) is the correlation for the blue dots in (D) which show the simulated effect of inter-patient differences in slope on variance in AG. This simulation shows that inter-patient differences in slope would generate data (blue dots) whose trend (black line) is consistent with that actually seen in the real data (red dots), and r_s^2 is similar to r_d^2 .

Figure 3. Inter-patient variation in M_{RBC} is sufficient to explain all glucose-independent variation in HbA1c

We used a simulation to test the hypothesis that measured variation in M_{RBC} can generate all observed glucose-independent variation in HbA1c as a function of AG. The AG values from the ADAG(1) study were used as input to Equation (4) along with constant k_g , constant $HbA1c(0)$, and an M_{RBC} randomly-sampled from a normal distribution with mean and standard deviation as measured in reference #(4). The medians are indistinguishable (p is the significance of a Kruskal-Wallis test of equal medians). (See “Model simulation” in Methods for more detail.)

Figure 4. Modeling M_{RBC} reduces errors in estimated AG

(Top panel) One patient's modeled M_{RBC} was 45 days in the fall of 2014. The blue line (#1) shows the M_{RBC} -adjusted AG-HbA1c relationship personalized for this patient, in contrast to the red line showing the current standard AG-HbA1c formula. One year after the M_{RBC} estimation, the patient visited the clinic and had an HbA1c of 8.1% (gray horizontal line, #2). The current standard method predicted an AG of 186 mg/dL (red "X"). The model predicted 209 mg/dL (blue "X"). This patient had CGM data available providing a direct and independent measurement of AG equal to 210 mg/dL (green checkmark). This patient's personalized AG-HbA1c model reduced the error in AG estimation from 24 mg/dL to 1 mg/dL. (Bottom panel) A second patient had a model-estimated M_{RBC} of 60 days in the spring of 2015, yielding a personalized AG-HbA1c relationship corresponding to the blue line (#1, bottom panel) in contrast to the red line showing the current standard formula. About 6 months later in the fall of 2015, the patient returned to the clinic and had an HbA1c of 10.5% (gray horizontal line, #2). The current standard method predicted an AG of 255 mg/dL (red "X"). The model predicted 205 mg/dL (blue "X"). This patient had CGM data available providing a direct and independent measurement of AG equal to 207 mg/dL (green

checkmark). This patient's personalized AG-HbA1c model reduced the error in AG estimation from 48 mg/dL to 2 mg/dL. These two examples highlight the fact that with current methods, a patient with lower AG (bottom) may actually have a significantly higher HbA1c than a patient with a higher AG (top), potentially compromising disease diagnosis and management.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5. Model-based inference of AG from HbA1c reduces estimation errors by about 50%
 Top row shows histograms of errors in AG estimation for 4 different sets of patients using the current standard regression-based formula. Second row shows histograms of errors using model-based estimation of AG. Histograms include predictions where estimation methods differ by at least 10 mg/dL. Errors for model-based predictions are significantly more tightly clustered around zero. The bottom panel compares boxplots of median absolute error and shows that the model reduces error by at least 50% in each of the 4 independent sets of patients. The model-based estimates are superior to the standard method in all four cases with $p < 0.001$. (See “Patient sets” in Methods and Supplementary Results for more detail.)