Personalized Diabetes Management Using Electronic Medical Records

Dimitris Bertsimas, Nathan Kallus, Alexander M. Weinstein, and Ying Daisy Zhuo

Diabetes Care 2017;40:210–217 | DOI: 10.2337/dc16-0826

OBJECTIVE

Current clinical guidelines for managing type 2 diabetes do not differentiate based on patient-specific factors. We present a data-driven algorithm for personalized diabetes management that improves health outcomes relative to the standard of care.

RESEARCH DESIGN AND METHODS

We modeled outcomes under 13 pharmacological therapies based on electronic medical records from 1999 to 2014 for 10,806 patients with type 2 diabetes from Boston Medical Center. For each patient visit, we analyzed the range of outcomes under alternative care using a *k*-nearest neighbor approach. The neighbors were chosen to maximize similarity on individual patient characteristics and medical history that were most predictive of health outcomes. The recommendation algorithm prescribes the regimen with best predicted outcome if the expected improvement from switching regimens exceeds a threshold. We evaluated the effect of recommendations on matched patient outcomes from unseen data.

RESULTS

Among the 48,140 patient visits in the test set, the algorithm's recommendation mirrored the observed standard of care in 68.2% of visits. For patient visits in which the algorithmic recommendation differed from the standard of care, the mean posttreatment glycated hemoglobin A_{1c} (HbA_{1c}) under the algorithm was lower than standard of care by 0.44 ± 0.03% (4.8 ± 0.3 mmol/mol) ($P < 0.001$), from 8.37% under the standard of care to 7.93% under our algorithm (68.0 to 63.2 mmol/mol).

CONCLUSIONS

A personalized approach to diabetes management yielded substantial improvements in HbA_{1c} outcomes relative to the standard of care. Our prototyped dashboard visualizing the recommendation algorithm can be used by providers to inform diabetes care and improve outcomes.

Type 2 diabetes is typically managed through healthy eating, physical activity, oral medication, and/or insulin injections. Although there are evidence-based clinical guidelines for glycemic control (1), how to choose among pharmacological therapies to maximize effectiveness for a given patient is not well understood. There has been growing interest in using clinical evidence to understand the effects of treatments in different populations with type 2 diabetes. In a joint statement from 2012, the American Diabetes Association and the European Association for the Study of

Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA

Corresponding author: Dimitris Bertsimas, dbertsim@mit.edu.

Received 15 April 2016 and accepted 6 November 2016.

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc16-0826/-/DC1.

N.K. is currently affiliated with the School of Operations Research and Information Engineering, Cornell University, Ithaca, NY and Cornell Tech, New York, NY.

Diabetes highlighted the need for a patient-centered approach to diabetes management (2). The need for an individualized approach is especially pressing given the variety of disease symptoms, comorbid conditions, pharmacological treatments, individual treatment histories, and other individual characteristics that may inform treatment (3).

Evidence suggests that the response to blood glucose regulation agents can differ among population subgroups. A post hoc secondary analysis found that African American adults with prediabetes responded better to metformin than Caucasian adults with prediabetes (4). Another study recommended less aggressive treatments for older patients, as they were more likely to experience severe consequences from hypoglycemia (5). These studies each provide valuable insights with respect to a single subgroup or treatment, but do not offer a decision rule for the general population that providers can easily apply in practice.

Tailoring glycemic management for specific subpopulations can be critical. Among patients with chronic kidney disease, contraindication to metformin needs to be taken into consideration when prescribing medication (6). Separate glycated hemoglobin A_{1c} (HbA_{1c}) goals may be needed for subgroups or individuals differentiated by age, comorbidities, and other clinical characteristics (3). A personalized treatment recommendation using a quantitative approach could readily incorporate different glycemic targets and contraindications and thus allow for more systematic management of subgroups.

We provide an algorithm that generates a personalized type 2 diabetes treatment recommendation for any given patient based on evidence from historical outcomes of similar patients drawn from an electronic medical records (EMR) database. EMR analysis allows for pinpoint comparisons of effectiveness because of the abundance of clinical evidence from multiple treatment options administered to a diverse population over long-term patient clinical histories. EMR data combine the large sample sizes found in some insurance claims databases with the depth of longitudinal clinical evidence typically found in clinical trials. One caveat is that EMR data are not controlled via randomization.

Our methodological approach applies machine learning techniques and causal inference to make personalized recommendations based on comparative effectiveness among subpopulations in the EMR database. Machine learning techniques have been increasingly adopted in health care, along with many other fields (7–9). Our novel approach leverages the power of analytics and abundant data in the EMR system to improve quality of care.

The recommendations are personalized by patient characteristics, including age, sex, race, BMI, treatment history, and diabetes progression. We evaluate the effectiveness of the personalized treatment recommendations against the current standard of care by estimating patients' counterfactual outcomes from historical outcomes of similar patients in the EMR database. We develop a prototype clinical support dashboard that provides evidence for the algorithm's recommendations and could guide providers in caring for patients with type 2 diabetes in a personalized manner.

RESEARCH DESIGN AND METHODS

Analytic Overview

We modeled outcomes for patients with type 2 diabetes based on EMR data. We divided each patient's medical history into distinct lines of therapy, each characterized by a particular drug monotherapy or combination therapy. Within each line of therapy, we considered patient visits occurring every 100 days. At each visit, the provider decides whether to proceed with the patient's current line of therapy or to recommend an alternative regimen. We developed a nonparametric prescriptive algorithm that provides personalized treatment recommendations. For each patient visit, we used *k*-nearest neighbor (*k*NN) regression (10) to predict the potential HbA_{1c} outcome under each treatment alternative. The nearest neighbors were chosen to control for confounding that may be present in nonrandomized data (11) and to maximize similarity on the patient characteristics that were most predictive of outcomes. The algorithm then prescribed the regimen with best predicted outcome, provided the predicted improvement relative to the patient's current regimen exceeded a confidence threshold. The outcome metric was the average HbA_{1c} measurement 75 to 200 days after the visit date. The effect of the prescriptive algorithm was evaluated by comparing the expected HbA_{1c} outcome under our recommended therapy to the observed outcome under the standard of care (ground-truth) therapy, according to a commonly used matching approach (12). We conducted additional simulations to ensure that the results were robust to training models on different datasets and using alternative predictive modeling techniques.

Data

Through a partnership with Boston Medical Center (BMC), an academic medical center in Boston, MA, we obtained EMR for >1.1 million patients from 1999 to 2014. In this dataset, 10,806 patients met all of the following inclusion criteria:

- Were present in the system for an observation period of at least 1 year;
- Received a prescription for at least one blood glucose regulation agent, including insulin, metformin, sulfonylureas, or one of the other blood glucose regulation agents listed below, and had at least one medical record 100 days prior to the date of this prescription;
- Had at least three recorded laboratory measurements of HbA_{1c}; and,
- Did not have a recorded diagnosis of type 1 diabetes, as defined by the presence of ICD-9 diagnosis code 250.x1 or 250.x3 combined with the absence of any subsequent prescriptions for oral blood glucose regulation agents. (If the patient received oral blood glucose regulation agents subsequent to one of these diagnosis codes, we assumed the diagnosis record was an error.)

For each patient, we had access to demographic data, including date of birth, sex, and race/ethnicity, and to all BMC EMR data, including a history of drug prescriptions and measurements of height, weight, BMI, and HbA_{1c}, as well as creatinine levels (Table 1). Neither the size of the population nor the proportion with good glycemic control changed substantially over the course of the study.

Interpreting Individual Medical Histories

We divided each patient's medical history into distinct lines of therapy, each characterized by a particular drug regimen (Supplementary Fig. 1). Within each line of therapy, we considered patient visits occurring every 100 days, corresponding to the life cycle of a red blood cell (13). These patient visits

Table 1—Demographics, medical history, and treatment history of patients (*N* = 10,806)

Feature	
Age (years)*	59.7 (13.6)
Percent male	42.4
Percent black	58.5
Percent Hispanic	15.1
Percent white	16.6
BMI (kg/m^2)*	33.1 (8.1)
HbA_{1c} (%)*	7.9 (1.8)
HbA_{1c} (mmol/mol)*	62.8 (19.7)
Percent with good glycemic control (i.e., $HbA_{1c} \leq 7.0\%$ [53 mmol/mol])*	37.7
Years since first treatment in EMR*	3.52 (3.66)
Percent with current prescription for metformin*†	45.6
Percent with current prescription for insulin*†	30.2
Percent with contraindication to metformin*‡	17.4
Number of patients with first visit prior to 2007 (%)	6,175 (57.1)
Observed standard of care regimen (abbreviation)	Number of patient visits (*N* = 48,140)
No regimen prescribed, new patient (NEWPT)	5,449
No regimen prescribed, existing patient (NORX)	2,137
Metformin monotherapy (MET0)	9,649
Insulin monotherapy (INS0)	7,539
Other blood glucose regulation agent monotherapy (OTHER0)	4,671
Metformin combined with one other noninsulin agent (MET1)	6,959
Metformin combined with insulin (METINS0)	3,977
Insulin combined with one nonmetformin agent (INS1)	2,139
Combination of two nonmetformin, noninsulin agents (OTHER1)	1,047
Metformin combined with two other noninsulin agents (MET2)	1,749
Metformin combined with insulin and one other agent (METINS1)	2,005
Insulin combined with two nonmetformin agents (INS2)	249
All other multidrug ($\geq$3) combinations (MULTI)	570

Data are mean (SD) unless otherwise indicated. *Sample statistics are calculated across all patient visits. Individual patients with longer medical histories may be overrepresented in the sample. †Individuals may have a current prescription for both metformin and insulin. ‡A patient was considered to be contraindicated to metformin when current serum level of creatinine was $>$1.5 mg/dL.

provided the basis for our definition of patient outcomes.

Lines of Therapy

We developed an algorithm to define precisely when each line of therapy ends and the next line begins according to when the combination of drugs prescribed to the patient changes in the EMR data. Each line of therapy was characterized by a unique drug regimen, defined to include all blood glucose regulation agents prescribed to the patient within the first 6 months after starting that line of therapy.

Regimens were defined as combinations of drugs from one or more drug classes. The drug classes we considered were metformin, insulin, and other blood glucose regulation agents; the other agents included sulfonylureas, thiazolidinediones, dipeptidyl peptidase 4 inhibitors, meglitinides, α-glucosidase inhibitors, glucagon-like peptide 1 agonists, and other antihyperglycemic agents. If a sufficient number of HbA_{1c} observations existed during a period in which no drugs were prescribed, we defined the patient's line of therapy as "NoRx." We considered 13 possible regimen types (Table 1). A combination of drug classes was included as a regimen type if it was observed in a sufficient number of patient visits.

We determined that a patient's current line of therapy had ended whenever the patient's drug regimen changed in some way, such as when one or more drugs were added to or removed from the drug regimen, or when the phase was interrupted by a period of at least 500 days with no new prescriptions, in which case the subsequent line of therapy was determined to be NoRx. This definition of end date for each line of therapy intends to capture the period when the patient was experiencing the effect of the drug regimen.

Patient Visits

Within each line of therapy, we considered patient visits occurring every 100 days, beginning with the visit at which that regimen was initiated and continuing until no later than 80 days prior to the start of the subsequent regimen. There were 48,140 unique patient visits in our dataset (Table 1). At each visit, we defined a set of visit-specific patient characteristics, including the current line of therapy (i.e., therapy given during the 100 days immediately preceding the current visit) and recent HbA_{1c} and BMI history. The outcome was measured as average HbA_{1c} 75 to 200 days after the visit. This effect period was chosen to allow for a complete red blood cell life cycle to elapse before measuring the effect of a drug therapy.

We defined the standard of care for each visit as the drug regimen that was administered. For 16.3% of visits, the provider prescribed an adjustment to the current line of therapy; in the other 83.7%, the provider's prescription was to continue the current regimen.

Prescriptive Algorithm

Our novel prescriptive algorithm considers a menu of available treatment options, including the patient's current treatment; uses *k*NN regression models to predict potential outcomes under each option; rejects any noncurrent treatment option with predicted outcome above a prespecified HbA_{1c} threshold; and chooses the remaining option with best predicted outcome. The menu of options for a given patient could be determined by the provider, accounting for contraindications and other preferences, such as not using intensive control for elderly patients or patients with a history of severe hypoglycemia.

For the purposes of this analysis, the menu of options for each patient was chosen relative to the intensity and composition of the patient's current treatment regimen. Specifically, the algorithm considered only regimens that represented an incremental addition or subtraction of a drug, or substitution of a drug of comparable intensity; metformin and insulin were considered to be of the lowest and highest intensity,

respectively. Patients with serum creatinine levels >1.5 mg/dL (6), a sign of kidney disease, were not offered metformin-based regimens. The menu options used in our analysis, differentiated by current treatment, are depicted in Fig. 1; by definition, the algorithm never recommended metformin-based therapies for patients with the contraindication described above.

For each patient visit, the outcomes predicted by *k*NN under each treatment were compared. Our algorithm selected the treatment with the best predicted HbA_{1c} outcome subject to the condition that this best predicted outcome improve upon the predicted outcome under the patient's current treatment by at least some threshold δ. We chose the optimal threshold value of 0.8% by testing the algorithm on a single test set, using values of δ ranging from 0 to 1.5%. Increasing the threshold δ causes the algorithm to recommend switching for fewer patients, but the mean benefit among those who switch increases. Above a certain threshold, the recommendation fits to noise in the training data and does not provide better mean benefits in the testing set. The optimal threshold balances these concerns.

*k*NN regression is a nonparametric, instance-based algorithm that makes predictions by averaging the outcomes for the subset of observations most similar to the target as defined by some distance metric (10). To predict potential outcomes under each regimen, we used a *k*NN regression based on a treatment-specific weighted Euclidean distance across normalized patient and visit-specific factors. The weights were derived by training a separate ordinary least squares linear regression model for each treatment regimen and using the magnitudes of the regression coefficients (Supplementary Table 1). This weighted distance improves upon classical *k*NN by selecting neighbors based on the factors most predictive of HbA_{1c} outcome, rather than weighting all factors equally.

We considered factors from the following categories: demographic information, medical history, and treatment history. Specifically, the demographic factors used in the model were age, sex, and race. The medical history factors were days since first diabetes diagnosis, the

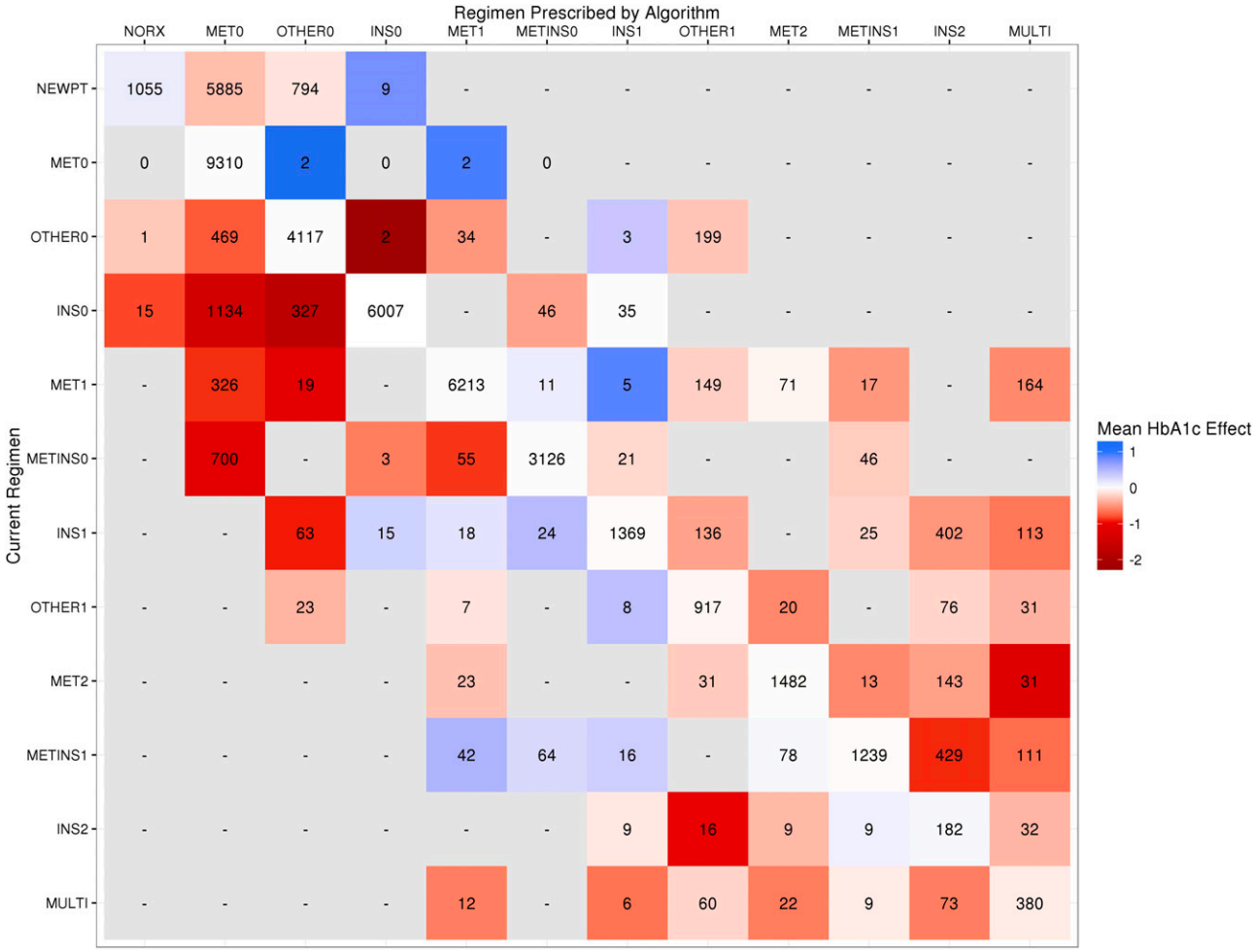

Figure 1—HbA_{1c} benefit of prescriptive algorithm for patients switching regimens. Each cell in the figure represents patients for whom the prescriptive algorithm recommended switching from the regimen on the vertical axis to the regimen on the horizontal axis. The color in each cell indicates the mean HbA_{1c} benefit (%) of the prescriptive algorithm for patients in that cell, with red indicating benefits of the algorithm and blue indicating worsening relative to standard of care. Each cell is labeled with the number of patients who made that switch; cells labeled with a dash were not on the menu of options provided to patients currently on a given regimen. Patients with serum creatinine levels >1.5 mg/dL were not considered for metformin-based regimens and therefore are never assigned by the algorithm to columns with metformin-based regimens.

patient's average serum creatinine level in the previous year, the patient's past two HbA_{1c} and most recent BMI observations up to and including the current visit, the patient's average, median, 25th percentile, and 75th percentile HbA_{1c} and BMI in the 1,000-day period up to and including the current visit, and the patient's frequency of HbA_{1c} measurements. The treatment history factors were the number of regimens the patient had tried, the number of visits since starting the current regimen, whether or not the patient had been previously prescribed metformin, and the patient's current regimen.

The prediction step of our algorithm is best illustrated through an example. Suppose we would like to estimate a patient's potential outcome under metformin monotherapy. To identify the importance of each factor in predicting outcomes, we used patient visits in which metformin monotherapy was prescribed to train an ordinary least squares regression on normalized values of each patient factor listed above. The most predictive factors were: the patient's most recent HbA_{1c} measurement (regression coefficient magnitude 0.22), whether the patient was currently prescribed insulin (0.11), the patient's mean BMI over the past 1,000 days (0.11), and several other HbA_{1c} and BMI measurements (coefficient magnitudes ranging from 0.03 to 0.10) (see Supplementary Table 1 for full details). To estimate the patient's potential outcome, we used the coefficient magnitudes to weight the Euclidean distance between this patient visit and each patient visit in which metformin monotherapy was prescribed. Thus, for any choice of k, we could rank the k closest neighbors from this treatment group. This procedure was repeated for each therapy in the patient's menu of treatment options.

Intuitively, the number of neighbors k used to estimate posttreatment HbA_{1c} levels should increase with the size of the dataset. For each treatment t, we found the value k_t^* that minimized the root mean square error of the kNN predictions on a subset of the data not used to evaluate the algorithm. We regressed k_t^* on $\sqrt{n_t}$ and thus derived the dependence function $k_t^* = 0.34\sqrt{n_t}$, which was used to select k in the prescriptive algorithm.

To verify the accuracy of the kNN HbA_{1c} predictions, we evaluated the R^2 metric. Positive values of R^2 suggest patient characteristics are predictive of future HbA_{1c}. For comparison, we evaluated the predictive accuracy of least absolute shrinkage and selection operator (LASSO) regression (14) and random forest (15), two state-of-the-art machine-learning methods used widely because of their high prediction accuracy. We used the predictions from these models in two alternative prescriptive algorithms.

Model Evaluation

To evaluate the performance of the kNN-based prescriptive model, we tested the algorithm's recommendations on a set of patient data that had not been used when training the models.

Because counterfactual treatment effects are not observable, we used the weighted matching approach embedded in the kNN regression to impute potential outcomes, an approach commonly used for causal inference in observational studies when randomization is unavailable (12). For each visit, we applied our prescriptive algorithm to recommend a therapy. If that recommendation matched the prescribed standard of care therapy, we observed the true effect from the therapy. Otherwise, the outcome was imputed by averaging the outcomes of the most similar patient visits at which the recommended therapy was administered; these similar visits were chosen from a test set not used for training, and the number of neighbors k_t^* was selected to fit the size of the test set. This estimated outcome was compared with the true outcome under standard of care at the given patient visit.

Our hypothesis was that the average predicted HbA_{1c} outcome after applying our prescriptive algorithm would be less than that observed from administering standard of care, resulting in a net average improvement in outcomes.

Sensitivity Analysis

To ensure the evaluation of our algorithm was not sensitive to the particular random split of the database into training and test data, we evaluated the effectiveness of our algorithm (with fixed threshold $\delta = 0.8$) under additional random splittings of the data.

Software

All analyses were performed in R 3.3.0 (16).

RESULTS

The R^2 of the kNN predictions on unseen data ranged from 0.20 to 0.54 depending on the regimen (Supplementary Table 2). The strongest models were for insulin monotherapy, metformin monotherapy, metformin plus insulin, and multidrug ($\geq$3) therapies. The R^2 values from the LASSO and random forest models ranged from 0.24 to 0.53. The predictive power was similar across the three methods.

The performance of the prescriptive algorithm is summarized in Table 2. The mean HbA_{1c} outcome after treatment was 0.14% lower under the prescriptive algorithm than under the standard of care

Table 2—Performance of prescriptive algorithms

	*k*NN	LASSO	Random forest
All patient visits (N = 48,140)			
Mean HbA_{1c} benefit relative to standard of care (SE)			
Percent	−0.14 (0.01)*	−0.13 (0.01)*	−0.07 (0.01)*
mmol/mol	−1.5 (0.1)*	−1.4 (0.1)*	−0.8 (0.1)*
Visits for which algorithm's recommendation differed from observed standard of care			
Number of visits (%)	15,323 (31.8)	12,684 (26.3)	14,302 (29.7)
Mean HbA_{1c} benefit relative to standard of care (SE)			
Percent	−0.44 (0.03)*	−0.45 (0.03)*	−0.26 (0.03)*
mmol/mol	−4.8 (0.3)*	−4.9 (0.3)*	−2.8 (0.3)*

*$P < 0.001$.

treatment, with SE 0.01% and significance level $P < 0.001$; equivalently, mean HbA_{1c} was 1.5 ± 0.1 mmol/mol lower under the algorithm. Of the 48,140 patient visits in our dataset, the algorithm differed from the standard of care for 15,323 visits, 31.8% of all visits. For this subset of visits, the mean HbA_{1c} outcome under the algorithm was lower by $0.44 \pm 0.03\%$ (4.8 ± 0.3 mmol/mol) compared with standard of care, with $P < 0.001$, a reduction from 8.37% under the standard of care to 7.93% under our algorithm or, equivalently, from 68.0 to 63.2 mmol/mol. The median outcome for these visits was 0.21% (2.3 mmol/mol) lower under the prescriptive algorithm compared with standard of care. For comparison, the median difference for all visits was zero because, for 68.2% of visits, there was no difference between the algorithm's recommendation and the standard of care.

In our analysis, the mean difference in HbA_{1c} was more negative than the median because of a left-skewed distribution. Some patients received particularly large benefits from using the prescriptive algorithm, which had an outsize effect on the mean but did not affect the median.

Fig. 1 depicts the number of patients for whom the prescriptive algorithm recommended switching from a given current line of therapy to a given new line of therapy, along with the mean reduction in HbA_{1c} for patient visits in each category. Among trajectories with at least 300 patients, the largest benefit of the algorithm was achieved through personalized recommendations for 7,564 patients currently on insulin monotherapy to switch to monotherapy with metformin or another blood glucose regulation agent. However, for the vast majority of patients currently on insulin-based regimens, the algorithm recommends that those patients continue with that therapy. Among the 7,564 patient visits, those who were recommended to switch from insulin were on average younger (mean 52.9 vs. 61.4 years of age) and had substantially higher average HbA_{1c} (11.0 vs. 8.0% or 97 vs. 64 mmol/mol).

The performance of the prescriptive algorithm in specific patient subgroups is summarized in Supplementary Table 3. The overall mean HbA_{1c} outcome using the prescriptive algorithm was 0.14% (1.5 mmol/mol) lower than standard of care for both male and female patients. The benefit of using the algorithm was 0.14% (1.5 mmol/mol) for black patients (29,120 visits), 0.09% (1.0 mmol/mol) for white patients (7,444 visits), 0.22% (2.4 mmol/mol) for Hispanic patients (6,732 visits), and 0.11% (1.2 mmol/mol) for all other patients (4,844 visits). The benefit of the algorithm was 0.20% (2.2 mmol/mol) for patients <60 years of age and 0.08% (0.9 mmol/mol) for patients aged ≥ 60 years or older. The benefit was 0.20% (2.2 mmol/mol) for patients with poor glycemic control (i.e., current $HbA_{1c} > 7.0\%$ [53 mmol/mol]) as compared with 0.05% (0.4 mmol/mol) for those with good glycemic control.

Our methodology motivates a provider dashboard that would report information on the demographics, medical history, and response to treatment for patients similar to an index patient. A prototype dashboard visualization for one sample patient visit is shown in Fig. 2. The dashboard would include the patient's demographic and

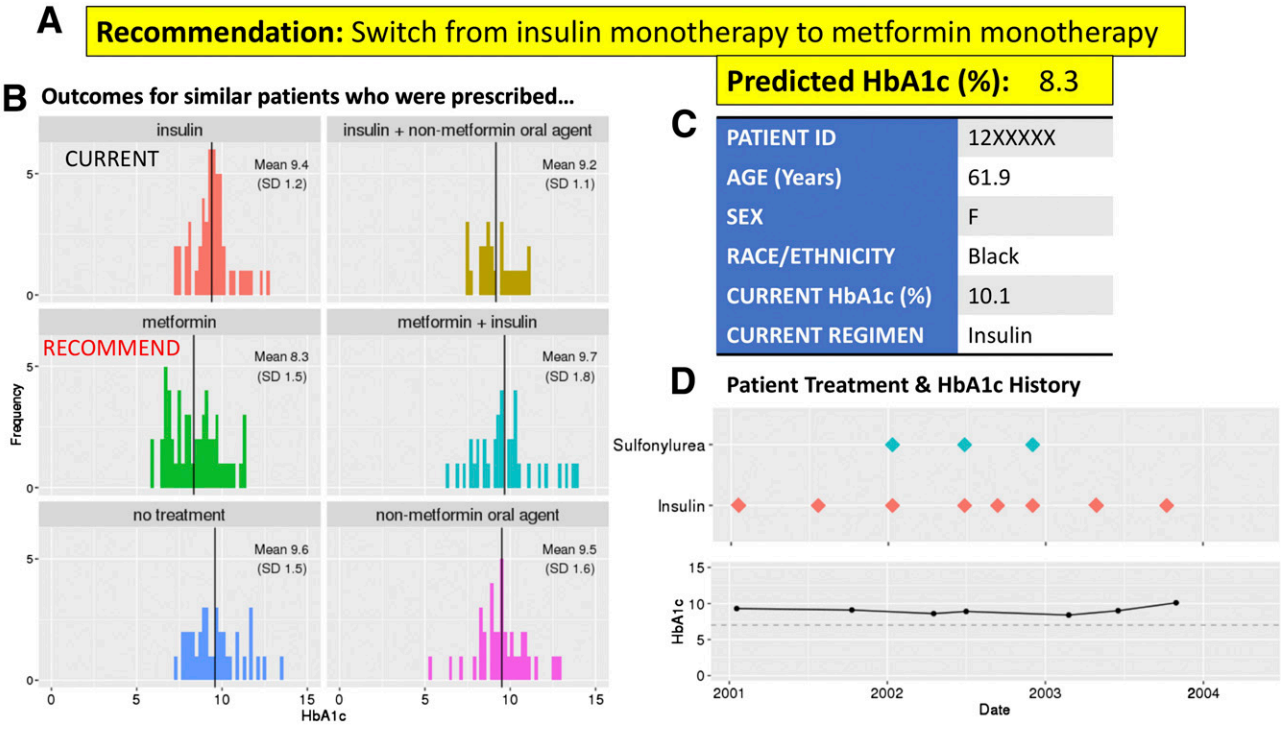

Figure 2—Visualization of prescriptive algorithm: provider dashboard prototype. This figure visualizes how the prescriptive algorithm can be used by providers for a single patient. *A*: Displays the algorithm's treatment recommendation along with the predicted posttreatment HbA_{1c} under that treatment. *B*: Depicts the mean, SD, and full distribution of posttreatment HbA_{1c} outcomes for the k_t^* most similar patient visits in the data set for each of the six regimens on this patient's menu of options. In each subpanel in panel *B*, the posttreatment HbA_{1c} level is on the horizontal axis, and the number of visits is on the vertical axis. The table in panel *C* presents basic information about the patient's demographic and medical history. *D*: Depicts the history of diabetes progression and treatment for the patient, with date along the horizontal axis. The vertical axis of the top subpanel indicates various drug classes; the vertical axis of the bottom subpanel depicts HbA_{1c} percentage.

health information along with visualizations of the patient's treatment history and HbA_{1c} progression. In addition, the dashboard would display the mean, SD, and full distribution of HbA_{1c} outcomes among the k_t^* nearest neighbors who received each treatment in the menu of options. Based on this evidence, the dashboard would display a treatment recommendation. The provider would have the ability to override this recommendation given any special management needs of the patient. For instance, if the patient is elderly and the distribution of HbA_{1c} outcomes indicates that the recommended therapy has an elevated risk of hypoglycemia, the provider may opt for an alternative treatment.

The overall mean HbA_{1c} outcome using the LASSO-based prescriptive algorithm was lower by 0.13 $\pm$ 0.01% (1.4 $\pm$ 0.1 mmol/mol; $P < 0.001$) compared with the mean standard of care outcome. The benefit from using the random forest–based prescriptive algorithm relative to standard of care was 0.07 $\pm$ 0.01% (0.8 $\pm$ 0.1 mmol/mol; $P < 0.001$).

In the sensitivity analyses, under three alternate random splittings of the dataset, the overall mean benefit of using the prescriptive algorithm compared with standard of care ranged from 0.11 to 0.15% (1.2–1.6 mmol/mol; $P < 0.001$ in all instances).

CONCLUSIONS

To our knowledge, we present the first prescriptive method for personalized type 2 diabetes care. Using historical data from a large EMR database, this novel prescriptive method resulted in an average HbA_{1c} benefit of 0.44% (4.8 mmol/mol) at each doctor's visit for which the algorithm's recommendation differed from standard of care.

Our method incorporates patient-specific demographic and medical history data to determine the best course of treatment. Compared with other machine-learning methods considered, the *k*NN prescriptive approach is highly interpretable and flexible in clinical applications. The novelty of our approach is in personalizing the decision-making process by incorporating patient-specific factors. This method can easily accommodate alternative disease-management approaches within specific subpopulations, such as patients with chronic kidney disease and elderly patients. We believe this personalization is the primary driver of benefit relative to standard of care.

In practice, the algorithm can be integrated into existing EMR systems to dynamically suggest personalized treatment paths for each patient based on historical records. The algorithm ingests and analyzes EMR data and generates recommendations. An intuitive, interactive dashboard summarizes the evidence for the recommendation, including the expected distribution of outcomes under alternative treatments (Fig. 2).

Because of the nature of retrospective data from existing EMR, this study has several limitations. Patients were not randomized into treatment groups. Although our matching methodology controls for several confounding factors that could explain differences in treatment effects, we can only estimate counterfactual outcomes. EMR data do not include socioeconomic factors or patient preferences that may be important in treatment decisions. Becaue of a lack of sufficient data, glucagon-like peptide 1 agonists were not considered as a separate drug class. If more data were available, we could further differentiate regimen types beyond the 13 we include in this analysis. In addition, the study population from BMC may not be representative of the U.S. population as a whole.

With EMR medication-order data alone, we cannot be certain whether a prescribed medication was filled or taken and cannot know precisely when the medication was stopped. Although this data quality issue could hamper attempts to make drug efficacy comparisons, our analysis aims to address the question of which drugs to prescribe under real-world scenarios. We optimize for an outcome that takes into account unobserved factors such as nonadherence. For instance, if nonadherence is more prevalent among patients prescribed insulin than other regimens, this perspective may explain why, in our study population, the algorithm recommends insulin less often than it is prescribed in clinical practice.

Our method can be extended to be more flexible and comprehensive. Currently, the prescriptive algorithm does not support individualized glycemic targets; we assume that a lower glycemic level is always preferred. The study currently optimizes only for a single health outcome; a more comprehensive algorithm would consider adverse event outcomes as well.

Despite these limitations, the study establishes strong evidence of the benefit of individualizing diabetes care. The success of this data-driven approach invites further testing using datasets from other hospital and care settings. Testing the prescriptive algorithm in a clinical trial setting would provide even stronger evidence of clinical effectiveness. As large-scale genomic data become more widely available, the algorithm could readily incorporate such data to reach the full potential of personalized medicine in type 2 diabetes.

In this study, we developed a novel data-driven prescriptive algorithm for type 2 diabetes that improves significantly on the standard of care when tested on patient-level EMR data from a large medical center. Our work is a key step toward a fully patient-centered approach to diabetes management.

Acknowledgments. The authors thank Dr. Michael Kane, Massachusetts Institute of Technology, for sharing clinical expertise in the progression and treatment of diabetes and Dr. William Adams, Boston University Clinical and Translational Science Institute, for sharing clinical expertise and assisting with the interpretation of EMR. The authors also thank BMC for use of its i2b2 database and the Associate Editor and the three reviewers for thoughtful comments that improved the paper significantly.

Funding. This research is partially supported by National Science Foundation grant 6926678 ["SHB: Type II (INT): Collaborative Research: Algorithmic Approaches to Personalized Health Care"].

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. D.B. contributed to the definition of the problem, methods development, and reviewing and editing the paper. N.K. contributed to the definition of the problem, methods development, data analysis, and reviewing and editing the paper. A.M.W. contributed to the definition of the problem, methods development, data analysis, and writing, reviewing, and editing the paper. Y.D.Z. contributed to the definition of the problem, methods development, data analysis, and writing, reviewing, and editing the paper. D.B. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 76th Scientific Sessions of the American Diabetes Association, New Orleans, LA, 10–14 June 2016.

References

1. Rodbard HW, Jellinger PS, Davidson JA, et al. Statement by an American Association of Clinical Endocrinologists/American College of Endocrinology consensus panel on type 2 diabetes mellitus: an algorithm for glycemic control. Endocr Pract 2009;15:540–559

2. Inzucchi SE, Bergenstal RM, Buse JB, et al.; American Diabetes Association (ADA); European Association for the Study of Diabetes (EASD). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 2012;35:1364–1379
3. Subramanian S, Hirsch IB. Personalized diabetes management: Moving from algorithmic to individualized therapy. Diabetes Spectr 2014; 27:87–91
4. Zhang C, Zhang R. More effective glycaemic control by metformin in African Americans than in Whites in the prediabetic population. Diabetes Metab 2015;41:173–175
5. Ismail-Beigi F, Moghissi E, Tiktin M, Hirsch IB, Inzucchi SE, Genuth S. Individualizing glycemic targets in type 2 diabetes mellitus: implications of recent clinical trials. Ann Intern Med 2011; 154:554–559
6. Lipska KJ, Bailey CJ, Inzucchi SE. Use of metformin in the setting of mild-to-moderate renal insufficiency. Diabetes Care 2011;34:1431–1437
7. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science 2015;349:255–260
8. Bertsimas D, Kallus N. From predictive to prescriptive analytics [article online], 2015. Available from https://arxiv.org/abs/1402.5481. Accessed 10 October 2016
9. Bertsimas D, O'Hair AK, Pulleyblank WR. *The Analytics Edge*. Charlestown, MA, Dynamic Ideas LLC, 2016
10. Cover T, Hart P. Nearest neighbor pattern classification. IEEE Trans Inf Theory 1967;13: 21–27
11. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55
12. Imbens GW, Rubin DB. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York, Cambridge University Press, 2015
13. Franco RS. Measurement of red cell lifespan and aging. Transfus Med Hemother 2012;39: 302–307
14. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol 1996;58:267–288
15. Breiman L. Random forests. Mach Learn 2001;45:5–32
16. R: a language and environment for statistical computing [Internet], 2016. Available from https://www.r-project.org/. Accessed 10 October 2016