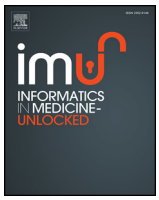
Informatics in Medicine Unlocked 16 (2019) 100191

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu



Artificial intelligence for diabetes case management: The intersection of physical and mental health

Casey C. Bennett[a,b,*]

[a] College of Computing and Digital Media, DePaul University, Chicago, IL, USA
[b] AI & Machine Learning Group, CVS Health, Chicago, IL, USA

ARTICLE INFO

Keywords:
Artificial intelligence
Medical decision making
Machine learning
Diabetes
Case management
Mental health

ABSTRACT

Objective: Diabetes is a major public health problem in the United States, affecting roughly 30 million people. Diabetes complications, along with the mental health comorbidities that often co-occur with them, are major drivers of high healthcare costs, poor outcomes, and reduced treatment adherence in diabetes. Here, we evaluate in a large state-wide population whether we can use artificial intelligence (AI) techniques to identify clusters of patient *trajectories* within the broader diabetes population in order to create cost-effective, narrowly-focused case management intervention strategies to reduce development of complications.

Methods: This approach combined data from: 1) claims, 2) case management notes, and 3) social determinants of health from ~300,000 real patients between 2014 and 2016. We categorized complications as five types: Cardiovascular, Neuropathy, Ophthalmic, Renal, and Other. Modeling was performed combining a variety of machine learning algorithms, including supervised classification, unsupervised clustering, natural language processing of unstructured care notes, and feature engineering.

Results: The results showed that we can predict development of diabetes complications roughly 83.5% of the time using claims data or social determinants of health data. They also showed we can reveal meaningful clusters in the patient population related to complications and mental health that can be used to design a cost-effective screening program, reducing the number of patients to be screened down by 85%.

Conclusion: This study outlines creation of an AI framework to develop protocols to better address mental health comorbidities that lead to complications development in the diabetes population. Future work is described that outlines potential lines of research and the need for better addressing the "people side" of the equation.

1. Introduction

1.1. Problem

Diabetes is a major public health problem in the United States, affecting roughly 30 million people (9.4% of the population) at a total annual cost of $327 billion USD as of 2017 [1]. Diabetes is also associated with the development of a number of serious complications – cardiovascular, renal, neuropathic, ophthalmic – that are major drivers of high costs and poor outcomes [2]. Those complications (e.g. renal failure) additionally lead to increased mortality risk [3,4]. Diabetes has a high comorbidity with mental health issues, such as mood disorders like depression and bipolar. Those mental health comorbidities are known to reduce treatment adherence in diabetes, and increase the risk for development of complications [5,6].

There have a number of efforts to tackle the diabetes problems through analytical solutions using both clinical electronic health record (EHR) data and insurance claims data, primarily focused on identifying patients who are "at-risk" [7–12]. Definitions of risk vary, such as identifying non-diabetic patients via genetic or other information who may develop diabetes in the future, or identifying patients who are pre-diabetic and at-risk for developing full blown diabetes. But for our purposes here (and in most real-world clinical settings for both providers and payors), the *fundamental definition of at-risk* is those diabetic patients who are at-risk for high costs and poor outcomes in the immediate future. In other words, how do we best manage patients who already have diabetes?

The problem for providers and payors is that the patients who are identified as most "at-risk" are, counter-intuitively, not always the ones who represent the most opportunity to change course. Or to put it another way: the real question is which patients have future trajectories (in terms of both costs and outcomes) that are actually *change-able*? In

* Corresponding author. 1613 Florence Ave, Evanston, 1.270.577.0439, IL, USA.
E-mail address: cabennet@indiana.edu.

https://doi.org/10.1016/j.imu.2019.100191
Received 30 March 2019; Received in revised form 8 May 2019; Accepted 10 May 2019
Available online 13 May 2019

real world clinical settings, we want actionable information [13]. A patient may be at-risk, but there may either be nothing we can do about it, or too late to do so. This issue has led to the development of "rising risk" models in both payor and provider analytic departments [14], which seek to identify patients with significant differentials between their current estimated risk and either past or future levels. In other words, patients whose trajectories suggest the individual will become a high risk, high cost patient in the future, even though they are not one currently. The idea is to intervene early before risk escalation occurs. On the flip side, such early intervention often comes with its own sizable costs (e.g. population-wide screening efforts), unless narrowly focused. There are limited resources – we cannot screen every patient for everything. This is a particularly acute problem for healthcare organizations, such as payors, integrated payor-provider systems, and accountable care organizations (ACOs), who are responsible for managing the total cost of care for patients and/or function in value-based care settings [15]. With the growth of value-based care models in the United States, the challenge is only growing.

1.2. Goal

The goal here is to evaluate whether we can cluster *trajectories* of diabetic patients – not necessarily the patients themselves – based on service utilization (detailed in Section 2.2.). The primary concern is what is causing the later development of complications in some diabetes patients, but not others. Then to evaluate the characteristics among those trajectory clusters and look for differences that represent *actionable* intervention opportunities. **In short, are there sub-groups of patient trajectories within the broader diabetes population we can identify where narrowly-focused early intervention strategies could reduce the later development of complications with reasonable costs?**

To this end, the current project was conducted in a statewide population in the southeastern United States comprising roughly ~32,000 individuals with diabetes. We looked at both payor claims data and social determinants of health data. The ultimate output was to create a deployable framework for case management of diabetes patients in real world clinical settings, resulting in a scalable, sustainable system. While we are concerned with the technical development of an artificial intelligence (AI) system for diabetes here, we are also concerned with the "people side" of the equation and how such a system could integrate with existing practices of providers and patients within healthcare systems. The latter is key for successful implementation and user adoption [13,16].

1.3. Previous work

In previous work, we have focused on creating AI systems for finding optimal treatments for a range of chronic illnesses. This includes systems to simulate and augment clinical-decision making in co-occurring physical and mental chronic illness [17,18], data-driven approaches to selecting optimal treatments for mental health [19,20], and robotic applications for patients with dementia and aging-related issues [21,22]. The primary aim across all these was improving treatment of mental health and cognitive issues, as well as understanding the role co-occurring physical illnesses (such as diabetes or cardiovascular issues) play in that.

Along similar lines, a number of researchers have been exploring the possibility of predicting complications development in diabetes care. For instance, the MOSAIC project in Europe has recently been building predictive machine learning models of diabetes complications in Type II diabetes using clinical data, reporting accuracy rates up to 83.8% [23]. Numerous papers in recent years have also constructed computational models of risk assessment for diabetes complications [24,25], while others such as Makino et al. have focused on predicting the development of specific complications like renal disease [26]. Elsewhere, the Centers for Disease Control (CDC) in the United States has begun developing forecasting models to support screening policies for co-morbid mental health issues in diabetic patients [27]. As mentioned in Section 1.1, those mental health comorbidities are known to reduce treatment adherence in diabetes, and increase the risk for development of complications [5,6]. In short, there are a number of ongoing projects looking at using artificial intelligence and machine learning techniques to better manage diabetes care and complications, addressing various facets and coming from varying angles.

1.4. Current work

In this work, we are particularly interested in AI applications focused on the *intersection* of physical and mental health, and how that can be used to produce tools to enhance case management for diabetes care. As described in Section 1.2, this entailed looking for clusters of trajectories of diabetic patients, then identifying differences between the clusters that represent actionable intervention opportunities. For instance, are there particular co-morbidities causing patient trajectories to worsen (i.e. switch from one cluster to another) that are alterable through some case management intervention? And critically, can we not only accurately cluster individual patients, but also *predict* when one may be likely to switch clusters before such a switch occurs?

This prediction piece is important, as our a priori hypothesis is that the trajectory clusters will align with higher or lower incidence of complications in sub-groups of the diabetes population. We base this on the fact that, as noted in Section 1.1, complications are a major cost driver in diabetes, and so should be reflected in service utilization patterns over time.

The final aspect of the work is understanding how this approach fits into current paradigms of case management for diabetes patients within payors and related organizations, such as ACOs. Often there is a lack of rigor in many systems for collecting specific case management interventions performed, where that information is often recorded in unstructured text notes with low fidelity or consistency. We address how this both represents a barrier, and an opportunity, for this kind of work.

2. Methods

2.1. Data

The current project was conducted in late 2017 on a statewide population in the southeastern United States who received health insurance through an Affordable Care Act exchange plan (ACA, known colloquially as "Obamacare") between 2014 and 2016. The total number of unique individuals was approximately ~300,000. Of those, roughly ~32,000 individuals had diabetes, defined as an individual diagnosed with an ICD9 code starting with '250'. 90% of those were individuals with Type II diabetes. In order to be able to analyze comparable windows of patient trajectories (24 months), we limited our analysis to only those individuals whose first appearance of diabetes in the dataset was in 2014 (though as discussed in Section 2.2 below, this is not a cohort analysis). Note this means the first documented insurance claim in the dataset for that individual, not necessarily the first clinical appearance of diabetes. This reduced the final dataset for analysis to 14,941 patients.

Data for each individual patient came from three distinct sources:

1) Insurance Claims Data – including both service claims and medication claims
2) Case Management Notes – which are unstructured text notes of each patient-case manager interaction
3) Social Determinants of Health – information about the individual's lifestyle, habits, and environment within which they live and work from *outside* the clinical context

A brief description of the sources of each data above. For #1 (claims), this was obtained directly from the claims processing system from a large insurance payor in the state, who covered approximately 70% of the population in that state. The total number of claim records available was roughly 29.9 million, alongside nearly 11 million medication prescription claim records. For #2 (case mgmt notes), this was obtained from the backend database for the population health tool used by case managers while managing their patient load. We had nearly 85,000 documented interactions. For #3 (social determinants), this data was obtained from a third party vendor (Acxiom) partnered with the payor. The social determinants data included hundreds of fields comprising things such as socioeconomic status indicators, shopping habits, political affiliation, PersonicX household market segment, off-label vitamin use, occupational data, transportation sources, and more.

Lab values (such as A1C measures) were also available but not included in this analysis, given that such clinical data is not always available outside of clinical settings. However, their inclusion in the future could augment the results seen here.

Data from these disparate sources was consolidated as a unified data warehouse (Postgres 9.5) running in a HIPAA-compliant cloud environment (Healthcare Blocks, https://www.healthcareblocks.com/) built atop Amazon Web Services (AWS). Data was loaded and transformed via ETL processes built in Pentaho, an open-source data integration tool. Data for individual patients from across the three datasets was tied together using a set of unique identifiers provided by the payor and Acxiom. Subsequently, data was loaded into KNIME (Version 3.5.1) [28], an advanced machine learning, modeling, and statistical software package, which also integrates WEKA (Waikato Environment for Knowledge Analysis; Version 3.7) [28]. Some additional modeling was done using Python's SciKit library (http://scikit-learn.org). Appropriate legal consent was obtained between the researchers and the insurance payor prior to accessing the data. We followed all applicable laws in conducting the research, including signed business associate agreements with lawyer approval on both sides.

2.2. Framework overview

The modeling framework described here operates using a number of definitions. First and foremost are the diagnostic definitions. Here, we define diabetes as an individual diagnosed with an ICD9 code starting with '250'. We did not differentiate in most of the analysis between Type I and Type II (90% were Type II), for comparison purposes to other recent research such as Dagliata et al. and Makino et. Al [12,23]. However, there was one curious finding related to Type I individuals from a post-hoc analysis of the clusters (detailed in the Results section 3.2 below). Mental health diagnoses were defined based on the national Agency for Healthcare Research and Quality (AHRQ) classification scheme (see Section 2.4), i.e. ICD9 codes starting with either '29', '30', or '31'. Additional sub-categories were defined specifically for Mood Disorders (e.g. Depression, Bipolar, Cyclothymia … see appendix for include codes). Categorization of diabetes drug stages was also done for the most common medications seen in the dataset, into the following groups: initial medications (e.g. Metformin), Sulfa drugs, GLP-1, DPP-4, SGLT-2, Thiazolidinedione, and Insulin replacement.

Diabetes complications were categorized into five groups: Cardiovascular, Neuropathy, Ophthalmic, Renal, and Other. Cardiovascular Disease was defined using the AHRQ classification codes for ICD9 (see Section 2.4) or codes starting with '2507'. Neuropathy was similarly defined using the AHRQ classification codes for ICD9 or codes starting with '2506'. Ophthalmic was defined as ICD9 codes starting with '3620' or '2505', including retinopathy, edema, and other ophthalmic issues. Renal was defined using the AHRQ classification codes for ICD9 or codes starting with '2504'. The Other category included any codes starting with '2502′, '2508′, or '2509′, including hyperosmolarity and other unspecified complications.

Patient trajectories here are defined based on service utilization, calculated based on the amounts on paid claims (similar to Liu et al. [29]). Like many payor/provider systems in healthcare, consistently collected outcome measurement is uncommon. As such, we use future service utilization as a proxy for whether the patient is getting better or worse over time. The assumption is that worse-off patients generally exhibit higher utilization levels than the rest, and vice versa.

We should also be clear that is *not* a cohort analysis – these patients were at various points of diabetes disease progression (we do not know their medical history prior to 2014), which actually plays a large role in some of our more interesting findings from the trajectory clustering and modeling in the Results section below. This also represents a realistic modeling scenario, at least within the United States and similar healthcare systems, where due to a lack of healthcare data sharing across various payors and providers, when someone signs up for a health insurance plan in a given year, we often know little to nothing about them prior. Real-world models intended to operate given that current reality need to be able to handle such *partial observability*.

2.3. Modeling approach

Different algorithms were used for modeling in different parts of the study, and each of these will be detailed in the relevant sections of the Results below. In short, we used a combination of clustering and classification techniques to solve various questions. Some of the techniques used included Expectation-Maximization (EM) clustering [30], Random Forests [31], J48 Decision Trees (a variant of the classic C4.5 algorithm) [32], Neural Networks [28], Logistic Regression, SVMs [33], and Bayesian Networks [28]. Standard cross-validation techniques from machine learning were used to ensure generalizability [19]. As will be seen in the Results below, there were a few techniques that stood out from the others in terms of performance for this particular problem and dataset. However, this paper is not intended as a comparison of machine learning techniques, as a plethora of those already exist in the literature. Rather here we are focused on the application.

2.4. Feature engineering

A critical part of modeling real-world healthcare datasets is the creation of "meta-data" from the underlying raw data derived from backend databases, which falls under the concept of *feature engineering*. The goal is to create information by intersecting across data sources and fields, constructing "meaning" out of individual data fields that may otherwise lack it [13]. In healthcare, this often takes the form of combining subject matter expertise (SME) with analytic techniques. Another approach is to use blind feature engineering (often on image data) via deep learning, although using such an approach with other clinical data (e.g. electronic health records, EHRs) is fraught with challenges that can produce spurious results, such as issues with data quality, disease heterogeneity, and temporality [34]. Here, we focus on the SME approach using experts from the payor, and highlight a couple of key engineered features which are critical to our results below.

First, we constructed from the underlying raw data a *geographic risk factor*. As we know, where someone lives is often a proxy for things such as socio-economic status, transportation access, lifestyle factors, etc. that correlate with health outcomes (see Fig. 1) [35]. This was done in this study at zip code level given the available data, but could be performed at a more finite level (e.g. census tracts) if available. Geographic risk was calculated by looking at historical utilization patterns in our dataset (scaled per-member) across geographic locales for all patients (not just diabetes) across the state, and then converted into a "risk score" for each zip code on a 0–4 scale. This incorporated two metrics: historical claims cost and the percentage of Top Ten utilizers in that zip code (defined as an individual who fell in the top 10% of members in terms of utilization). This final 0–4 scale was created via equal-widths binning of the raw values based on the mean and standard deviation across all zip codes for each metric separately on a 0–2 scale, then

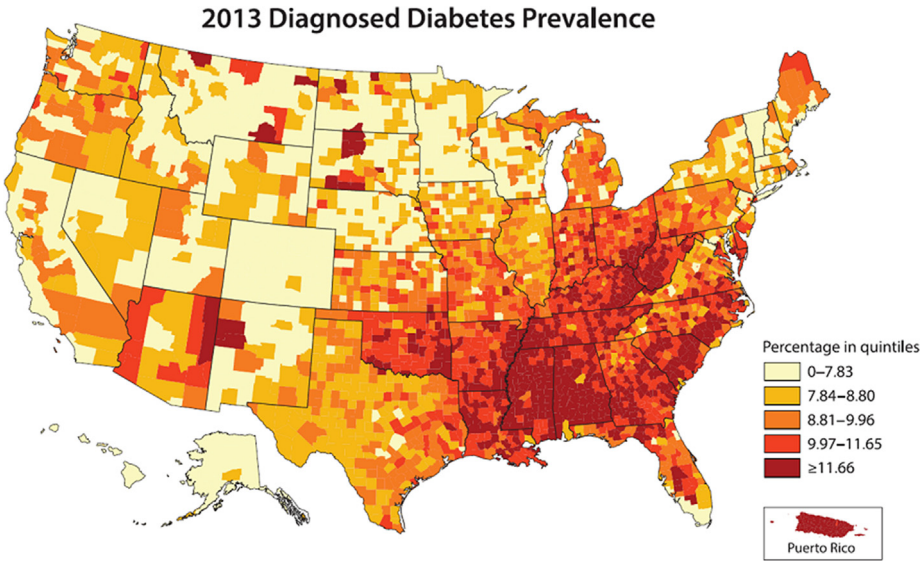


Fig. 1. Diabetes prevalence geographic risk example (taken from Ref. [36]).

adding those together. This was necessary as different metrics were on different scales.

$$Geo\ Risk\ Score = \begin{cases} if\ c > \bar{c} + 0.67\sigma_c & then\ 2 \\ if\ c > \bar{c} - 0.67\sigma_c & then\ 0 \\ else & then\ 1 \end{cases} + \begin{cases} if\ t > \bar{t} + 0.67\sigma_t & then\ 2 \\ if\ t > \bar{t} - 0.67\sigma_t & then\ 0 \\ else & then\ 1 \end{cases}$$

Where c equals historical claim costs for the zip code, and t equals the percentage of Top Ten utilizers for the zip code. We experimented with other approaches for this calculation including different scalings and other metrics like cost ratio (fourteen different approaches total, not shown for brevity), but this particular calculation showed the best performance during analysis.

Another piece of meta-data we termed “Quick-Hitters” (or QH for short). Early on, we had a suspicion that there might be some members who sign up for coverage via the Exchange and use a high number of services, then subsequently do not use many services after that (or in some cases don't even keep their coverage) [37]. This is a potentially significant problem for payors providing Exchange plans under the ACA in the United States, particularly if they drop coverage. We calculated this by looking at people who had high cost ratios, defined as spending over 90% of their total annual claim amounts in the first 60 days of coverage, to create a binary variable. Interestingly, on average, members who were Quick-Hitters tended to actually have *lower claim amounts than average*. However, there are notable exceptions, which led to a further distinction between “low risk” Quick-Hitters and others with higher utilization patterns. Some basic info about Quick Hitters is shown in Table 1. The main takeaway is that Quick-Hitters comprise about 15% of the total exchange member population, and that their utilization patterns are significantly different from the rest of the population.

Finally, we also tagged AHRQ classification codes onto the raw data (https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp). We used that to categorize diagnosis codes (ICD9 and ICD10), CPT procedure codes (including HCPCS), and inpatient revenue codes into smaller categories for analysis. For instance, this allowed us to reduce the roughly 16,000 unique CPT codes in our dataset down to 234 categories. The original AHRQ download files provide instructions for using the classification codes in SAS, but we rewrote that into SQL code procedures to apply the classification codes directly inside a database with claims or clinical data.

3. Results

We are interested in evaluating three questions here: 1) can we discern predictable patterns of complications development in diabetes patients, 2) do those patterns cluster into common trajectories, 3) can we use those clustered trajectories to recommend actionable interventions.

3.1. Predicting diabetes disease progression

Our first step was to explore whether we could predict diabetes disease progression, in particular whether a diabetes patient without any complication-related claims in a calendar year would develop complications in a subsequent 12-month period (binary prediction, yes vs. no). This is a similar approach to what Dagliati et al. did in the recently reported MOSAIC project, though theirs was based on clinical data from a hospital setting [23]. This left a sample size of about 5,000 individuals, with about 30% developing complications in the subsequent time period (the majority of which were cardiovascular-related). Complications were grouped into Cardiovascular, Neuropathy, Ophthalmic, Renal, and Other, as defined in Section 2.2. Given the limited 12-month time frame, we did not attempt to predict specific complications, but complication development overall.

Results can be seen in Table 2, using a Random Forest model (number of trees set to 100, max depth unlimited). The overall accuracy was roughly 83.5%, assessed via standard 10-fold cross validation [19]. We note that is extremely close to Dagliati et al., 2018 who reported 83.8% for predicting complications, theirs from a clinical EHR perspective and ours from a payor claims perspective [23]. We also note

Table 1
Quick-hitter basic info.

Quick Hitter?	Member Count	60 Day Cost	180 Day Cost	365 Day Cost	Cost Ratio
Yes	29083	$2,041	$2,102	$2,972	0.99
No	183138	$506	$2,157	$4,291	0.23

Table 2
Diabetes complication prediction.

Prediction	Non PredPos %	PredPos %	Total Acc	AUC
Diabetes to Complications (2015)	19.9%	87.0%	83.5%	0.9199

Table 3
Claims features for predicting diabetes complications.

Driving Features	Gain Ratio
Prev Year Service Features	0.147
Age	0.095
Prev Year Costs	0.084
Geographic Risk	0.053
Plan Type	0.045
PolicyHolder (Yes/No)	0.036

that similar to Dagliati et al., in order to obtain said performance we had to re-balance the dataset, since most individuals did not develop complications. In our case, we used SMOTE to synthetically oversample the minority class [38]. Note that the PredPos in Table 2 indicates the sensitivity, while "Non PredPos" is equivalent to 1-specificity.

We also tried various other machine learning classification algorithms mentioned in Section 2.3 to predict complications, including SVMs, Neural Networks, Bayesian Networks, J48 Decision Trees, and logistic regression. None of the other algorithms performed as well as Random Forests, with Bayesian Networks and Neural Networks coming closest with accuracy levels around 78–80% (data omitted for brevity).

There were several hundred variables considered in the model, with 37 of those determined to be driving the prediction, based on the gain ratio of each feature [28]. For brevity, we summarize those variables in Table 3. Most of those are self-explanatory, but we do note that "service features" refer to details about claims a member had the previous year, things such as procedure codes, procedure category, diagnoses, diagnosis category, provider specialty, provider type, and place of service. In other words, patients who had more complex medical histories and risker co-morbidities were more likely to develop complications. Plan Type indicated the ACA exchange metallic plan level (Bronze, Silver, Gold), while PolicyHolder indicated whether the member was the primary policy holder or a covered dependent. We also note that the geographic risk score constructed in Section 2.4 turned out to be a critical predictor. On the other hand, we note that the "Quick Hitter" constructed feature did not turn out to be useful in this context, even though it was useful in predicting high utilizers in general. That was not entirely surprising though, given the ongoing chronic nature of diabetes, averse to more episodic diseases.

3.2. Predicting diabetes complications using social determinants of health

We were also interested in using the *social determinants of health* data described in Section 2.1 to predict the development of diabetes complications, using the same approach as above. When we simply added the social determinants data to the existing claims data for the model in Section 3.1 though, performance was roughly the same, indicating the social determinants did not add any *additional* information to the model apart from the patterns the claims data already predicted.

A separate question, however, is whether the social determinants data could be used as a *substitute* for claims and clinical data. To this end, we stripped out all features from the claims, and only included fields from the social determinants data. Using the same Random Forest approach as above, we were able to build models with only social determinants data predicting the development of diabetes complications with approximately 79.1% accuracy, which is comparable to we can achieve with claims data alone. This is valuable, in that it may allow for making patient predictions in the future based on social determinants for new members we know nothing about (i.e. don't have claims or clinical information yet). Or, in other words, allow us to address the *partial observability* problem when making these kinds of healthcare predictions.

There were 109 selected features in the final model (out of 1746 total), which we summarize in Table 4. Of notable interest were

Table 4
Social determinant features for predicting diabetes complications.

Variable	Description
AP models	Vitamin use (A,B,C, Ginger, Echinacea, etc.)
AP models	Weight Loss
AP models	Arthritis indicators
AP models	Whether they received RX online
AP models	Likelihood of making doctor's appt last 12 months
AP006257	Social Setting Risk
AP006258	Smoking Frequency
AP006259	Unhealthy Diet
PersonicX Clusters	Appear to be "lifestyle" clusters
3101–3103	All race/ethnicity (primary risk is if the person was African-American)

variables related to vitamin use, weight loss, and home remedies for arthritis, as well as internet search behavior for medical information and vitamin use. These variables were often, though not always, higher in the group that *did not* develop complications. This seems to indicate that many of these individuals are practicing a fair amount of "self-medicating" outside conventional clinical therapies. Alternatively, it could also be taken as an indicator that the individual is taking a more active self-interest in their own health outcomes. Either way, social determinant variables associated with patient self-care are predictive of higher outcomes and reduced costs. We intend to explore this more in future work.

3.3. Clustering diabetes patient trajectories

While being able to predict whether an individual has a higher likelihood of developing complications is useful, the real question is *what do we do about it*. To this end, we investigated whether the predictive patterns of complication development seen above could translate into common patient trajectory clusters in terms of service utilization over time (see Section 2.2). Critically we were looking for such clusters not created by a human, but *emergent* in the data itself.

The same data as above for all 14,941 diabetes patients in the dataset was processed through an EM clustering algorithm, an unsupervised form of cluster learning (see Section 2.3). Clustering was performed using cross-validation, which found optimal performance with four clusters based on log likelihood. Average per-member service utilization trajectories (cumulative over time, starting in the first quarter of eligible insurance coverage under an Exchange plan) are shown in Fig. 2.

We compared these emergent clusters based on patient characteristics in each one. In summary, there was a high-utilizer Orange group with a lot of complications, and a very high incidence of Renal complications relative to other groups (see Table 5). The low-utilizer Gray group had fewer complications, although they did have a high incidence of minor cardiovascular complications (e.g. hypertension). The Blue group fell somewhere in between those two in terms of utilization and complication incidence rates. The most interesting group is the Yellow group, who appear to be "newer" cases with fewer complications or mental health issues and lower levels of medication prescriptions. They tended to be on the initial stage medications (Metformin) if any, and only prescribed Insulin or Sulfa drugs at half the frequency of the other groups.

In short, there were significant differences among the clusters in terms of complications, mental health comorbidities, medication drug stage (insulin vs. other drugs like DPP-4 and Sulfa), secondary diagnoses severity, and so forth. A couple examples (mental health comorbidity and complications) are shown in Tables 5 and 6.

Interestingly, there was also a significant mental health component (Table 6), with the Blue and Orange groups being twice as likely to have a mental health comorbidity (which on closer analysis were primarily mood disorders such as depression and bipolar disorder), which aligns

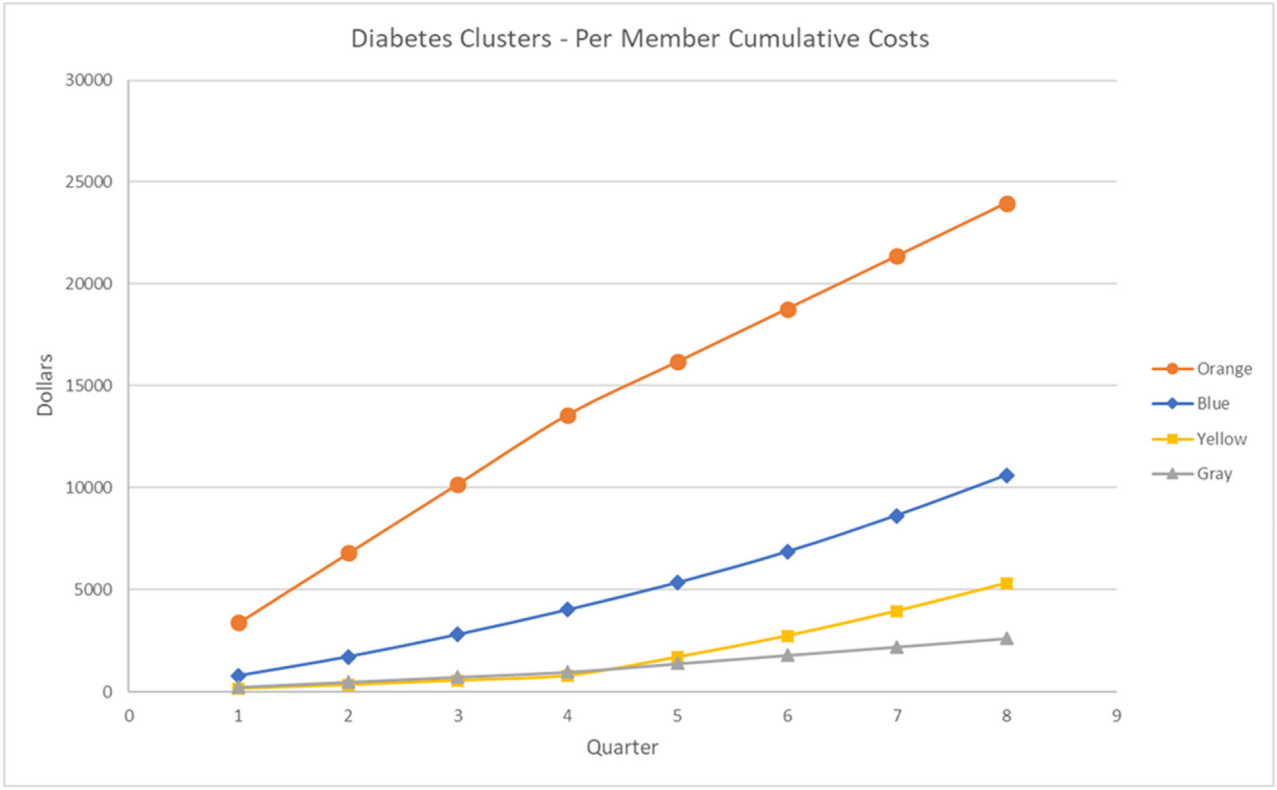


Fig. 2. Diabetes trajectory clusters.

with previous findings [5,39,40]. The TopTen Utilizers (defined as an individual who fell in the top 10% of members in terms of utilization) mirrored that pattern. There appears to be some connection between lower rates of mental health comorbidity, lower rates of more serious complications, and lower service utilization patterns, as evidenced in the Gray and Yellow groups.

As a brief aside, we also note that Type I diabetes appears in normal proportions (~10% give or take a few percent) in all clusters *except* the Orange cluster, which is curious since Type I and Type II had very similar complication rates overall (65.6% vs. 62.6%, respectively). One possible explanation is that since Type I often develops earlier in life as juvenile diabetes, more of those people had bifurcated into that high-cost high-utilization Orange cluster, but more research is needed.

After identifying these clusters, the next question was whether individuals *switch clusters* over time, and at what rate? In other words, if patients get worse or better, that should be reflected in changes to their trajectories. Our analysis showed that for the Gray, Blue, and Orange groups, nearly 70% of individuals remained in the same cluster. However, for the Yellow group, less than half of the individuals (47.3%) remained in the same cluster over time. It appears that, as relatively newer cases, those individuals bifurcate either into more stable, lower utilizer diabetic cases (Gray cluster) or into less stable, higher cost diabetic patients with more complications and needs (Blue and Orange clusters). The latter transition also corresponds to higher rates of mental health comorbidities seen in the Blue and Orange clusters. There is likely a link between the two – development of complications and development of mental health comorbidities – although it is not clear which factor is driving the other. It is also very possible that there are *unaddressed* mental health issues in the Yellow group which are yet to be diagnosed or treated.

All of this suggested that the Yellow cluster would be a likely focus for targeted case management interventions on a subset of diabetes population (~2,000 patients per year), i.e. narrowly-focused early intervention strategies, that could improve outcomes, reduce costs, and be operationalized effectively with limited resources. Given the results above, such targeted interventions would likely need to focus on both early-stage diabetes progression to reduce complication development as well as unaddressed mental health issues. Or in other words, focus on the intersection of physical and mental health. Stymieing one could

Table 6
Mental health comorbidity rates by cluster.

Winner Cluster	Member Cnt	Avg Age	Mental Health comorbid	% TopTen Utilizers
Gray	3864	54.3	21.8%	0.0%
Yellow	2260	49.2	27.1%	0.1%
Blue	6090	50.5	41.4%	13.1%
Orange	2727	50.9	51.7%	81.4%
Total	14941	**51.4**	36.1%	20.2%

Table 5
Complication rates by cluster.

Winner Cluster	Member Cnt	Cardiovascular disease	Neuropathy	Opthalmic	Renal	Other Complications
Gray	3864	94.8%	6.4%	5.0%	3.3%	4.5%
Yellow	2260	4.2%	0.7%	0.6%	0.0%	0.3%
Blue	6090	71.5%	14.7%	8.1%	8.4%	7.2%
Orange	2727	83.9%	22.8%	10.4%	26.6%	18.1%
Total	14941	69.6%	11.9%	6.6%	9.1%	7.4%

Depression Interventions
Administer PHQ9
Assess Symptoms
Assess Med Adherence
Social Needs Assessment
Identify Treatment Barriers
Review Treatment Goals
Arrange PCP visit
Request Physician Med Review
Request Physician Tx Plan Change
Arrange Mental Health visit
Arrange EAP

Fig. 3. Example of depression protocol.

help stymie the other.

3.4. Text mining interventions from case management notes

The final question here for developing an AI system to support diabetes case management is whether currently available population health and care management software typically used by case managers to document care of diabetes patients could effectively track diabetes or mental health interventions that could be leveraged as part of an AI framework. The problem is that in most of those software such data is collected as unstructured text notes, meaning information about interventions needs to be parsed out, and the actual recording of keywords associated with specific interventions or protocols is dependent on case manager training.

In a separate analysis, we looked at a randomly sampled subset of case management notes of encounters with patients having diabetes and depression (roughly 3500) to evaluate whether we could identify interventions from a depression protocol, as shown in Fig. 3.

Code was written in Python (www.Python.org) using the NLTK toolkit (www.nltk.org/) to tag interventions in case management notes via a Bag-of-Words analysis along with Regex to account for spelling variation [41]. The bag of words were based on individual component keywords seen in Fig. 3, but not full phrases. For instance, for "Arrange PCP Visit", we looked for the presence of the term 'pcp' along with any of the terms like 'schedule', 'arrange', 'visit', setup', 'appointment', while using Regex to account for spelling variations. These terminology component keywords were drawn from the training manual that the insurance payor used to train case managers in how to document.

We evaluated this *Intervention Tagger* and found that we could identify at least one intervention in 16.2% of encounters, with about 10% of those encounters having multiple interventions in the same encounter.

That percentage is, as expected, very low. This represents a significant barrier to deploying any sort of AI system for diabetes case management in a real-world setting. Such a barrier would necessitate either 1) enhancements to the software to collect interventions as structured data fields, or 2) changes to care manager training to improve intervention keyword recording in unstructured text fields. Either of these approaches are associated with the broader challenges of deploying technical innovations in real-world clinical practice, i.e. user adoption and implementation science [13,16,42].

4. Discussion

4.1. Main findings

The main takeaway of this study is that there is a critical connection between mental health issues and the development of complications in diabetes, and by stymieing one we can potentially stymie the other. Our findings revealed clusters of *trajectories* in diabetes patients, that show how rates of complications and mental health comorbidities co-occur in unison over time. It also revealed a particular cluster (the "Yellow" cluster) which would be a likely focus for targeted case management interventions as a subset of diabetes population, due to it largely representing "newer" cases which are likely to bifurcate over time into more stable, lower-utilizer diabetic cases or into less stable, higher-cost diabetic patients with more complications. In short, the Yellow cluster represents *actionable* intervention opportunities.

We also found we could predict the development of complications using insurance claims data with 83.5% accuracy using Random Forest techniques, which is very similar to that reported by Dagliati et al., 2018 in the MOSAIC project using clinical data [23]. Interestingly, we were also able to achieve comparable performance using only social determinants of health data.

These findings were converted into the following derived protocol, combining artificial intelligence with changes in clinical practice, for reducing development of complications in diabetes populations:

1) Cluster: Clustering approach to identify patients similar to the Yellow group described above
2) Screen: Screening program for unaddressed mental health issues among the Yellow group
3) Intervene: Targeted case management interventions for mental health for identified patients

This approach provides three key advantages. First, it reduces the population to be screened down by roughly 85% (15,000 to 2,000 patients), providing a more cost-effective, narrow focus. Second, it targets patients who are *likely* to change in the near future, i.e. actionable information. Third, it addresses the mental health comorbidities known to reduce treatment adherence in diabetes, and increase the risk for development of complications [5,6].

4.2. The "people side" of the equation

A major challenge in moving any technical innovation into real-world clinical practice is dealing the "people side" of the equation. Indeed, what is not often acknowledged is that when we deploy artificial intelligence or other analytic solutions, we are in fact engaging in an act of *behavioral re-engineering*. We are trying to change what people do, how they behave, the choices they make. Because in the end, if we don't change any of those things, then we haven't truly "accomplished" anything. The true mark of any technology, clinical or otherwise, is user adoption.

These issues tie back to the field of implementation science [13,43–45], and the work of Kaplan and Green addressing the challenges of creating clinical utility out of technology [16,42]. This is not a trivial problem, which goes beyond the quality of the technology or solution itself. Often there is a misalignment between how the technology "thinks" about the problem, and how clinicians and patients think about the problem. Moreover, there are often issues in other parts of the technology stack or clinical workflow which are limiting factors.

A good example of such a limiting factor is the case note issue described in Section 3.4. The above protocol in Section 4.1 necessitates consistent collection of case management interventions performed at each patient interaction, but analysis of currently available population health and care management software typically used by case managers to document care of diabetes patients showed that such interventions could only identified 16.2% of the time. While enhancements to the software may help address this issue, it is also likely a training issue. Effective measures are needed to improve intervention keyword recording in unstructured text fields. In short, strategies have to simultaneously address both technology and "people side" issues. Successful deployment of any AI solution in real-world practice depends on this holistic understanding. Much of this work remains unaddressed here, and represents an opportunity for future research.

4.3. Limitations

There are a number of limitations to this study, mostly related to the nature of working with claims data and the realities of a real-world insurance payor setting. First and foremost is the problem of *partial observability*. We do not know anything about the patient prior to 2014, nor have their complete medical history. For instance, it is entirely possible, though uncommon, that a patient may have had a diagnosed cardiovascular complication, such as high blood pressure, prior to 2014 and yet went untreated for the entire year of 2014. Even though it is uncommon, there is no way to rule it out. A second major issue is that in this scenario, we only receive data about a patient when a claim occurs, rather than at pre-determined set intervals. In other words, we only know anything about a patient when they receive medical services of some kind. What might be occurring in their day-to-day health outside that scope is unknown. Social determinants data (see Section 3.2) may help address this in the future, as well as in-home sensor technology deployed via smart home devices and social robots [22].

5. Conclusion

We presented here a study using insurance claims and social determinants data to identify clusters of *trajectories* of diabetes patients, in order to support development of an artificial intelligence solution that could reduce development of complications related to diabetes. We showed how we can 1) predict development of complications 83.5% of the time using either claims data or social determinants of health data, 2) identify clusters to reduce the number of patients to be screened down by 85% to create a cost-effective screening program, 3) use a derived protocol with the AI tool to better address mental health comorbidities in the diabetes population. Future work is intended to focus on deployment of this approach on a broader scale, and better address issues related to the "people side" of the equation – e.g. consistent intervention recording in care notes – through participatory design methods [46].

Ethical statement

The authors have no ethical conflicts, financial or personal or otherwise, related to the research presented herein.

Conflicts of interest

The authors have no conflict of interest related to the research presented herein. The information contained herein is not confidential, and has been presented in various public presentations over the past 2 years.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The author would like to acknowledge the support of the following individuals in this work: Dr. Tom Doub, Dr. Karin Chu, Dr. Selma Sabanovic, Dr. Joanna Thomas, Dr. Connie Meeks, Hamza Qadir, and Steve Abell. The work – and the opinions expressed herein – are those of the author, and do not necessarily reflect the views of CVS Health, DePaul University, or their affiliates, nor those of any of the author's previous affiliations or collaborators.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.imu.2019.100191.

Appendix

Table A1Mood Disorder Subcategory ICD9 Codes

ICD9 Code	Description	Subcategory
29600	Bipolar I disorder, single manic episode, unspecified	Bipolar
29601	Bipolar I disorder, single manic episode, mild	Bipolar
29602	Bipolar I disorder, single manic episode, moderate	Bipolar
29603	Bipolar I disorder, single manic episode, severe, without mention of psychotic behavior	Bipolar
29604	Bipolar I disorder, single manic episode, severe, specified as with psychotic behavior	Bipolar
29605	Bipolar I disorder, single manic episode, in partial or unspecified remission	Bipolar
29606	Bipolar I disorder, single manic episode, in full remission	Bipolar
29610	Manic affective disorder, recurrent episode, unspecified	Bipolar
29611	Manic affective disorder, recurrent episode, mild	Bipolar
29612	Manic affective disorder, recurrent episode, moderate	Bipolar
29613	Manic affective disorder, recurrent episode, severe, without mention of psychotic behavior	Bipolar
29614	Manic affective disorder, recurrent episode, severe, specified as with psychotic behavior	Bipolar
29615	Manic affective disorder, recurrent episode, in partial or unspecified remission	Bipolar
29616	Manic affective disorder, recurrent episode, in full remission	Bipolar
29640	Bipolar I disorder, most recent episode (or current) manic, unspecified	Bipolar
29641	Bipolar I disorder, most recent episode (or current) manic, mild	Bipolar
29642	Bipolar I disorder, most recent episode (or current) manic, moderate	Bipolar
29643	Bipolar I disorder, most recent episode (or current) manic, severe, without mention of psychotic behavior	Bipolar
29644	Bipolar I disorder, most recent episode (or current) manic, severe, specified as with psychotic behavior	Bipolar
29645	Bipolar I disorder, most recent episode (or current) manic, in partial or unspecified remission	Bipolar
29646	Bipolar I disorder, most recent episode (or current) manic, in full remission	Bipolar
29650	Bipolar I disorder, most recent episode (or current) depressed, unspecified	Bipolar
29651	Bipolar I disorder, most recent episode (or current) depressed, mild	Bipolar
29652	Bipolar I disorder, most recent episode (or current) depressed, moderate	Bipolar
29653	Bipolar I disorder, most recent episode (or current) depressed, severe, without mention of psychotic behavior	Bipolar
29654	Bipolar I disorder, most recent episode (or current) depressed, severe, specified as with psychotic behavior	Bipolar
29655	Bipolar I disorder, most recent episode (or current) depressed, in partial or unspecified remission	Bipolar
29656	Bipolar I disorder, most recent episode (or current) depressed, in full remission	Bipolar
29660	Bipolar I disorder, most recent episode (or current) mixed, unspecified	Bipolar
29661	Bipolar I disorder, most recent episode (or current) mixed, mild	Bipolar
29662	Bipolar I disorder, most recent episode (or current) mixed, moderate	Bipolar

29663	Bipolar I disorder, most recent episode (or current) mixed, severe, without mention of psychotic behavior	Bipolar
29664	Bipolar I disorder, most recent episode (or current) mixed, severe, specified as with psychotic behavior	Bipolar
29665	Bipolar I disorder, most recent episode (or current) mixed, in partial or unspecified remission	Bipolar
29666	Bipolar I disorder, most recent episode (or current) mixed, in full remission	Bipolar
2967	Bipolar I disorder, most recent episode (or current) unspecified	Bipolar
29680	Bipolar disorder, unspecified	Bipolar
29689	Other bipolar disorders	Bipolar
29620	Major depressive affective disorder, single episode, unspecified	Depression
29621	Major depressive affective disorder, single episode, mild	Depression
29622	Major depressive affective disorder, single episode, moderate	Depression
29623	Major depressive affective disorder, single episode, severe, without mention of psychotic behavior	Depression
29624	Major depressive affective disorder, single episode, severe, specified as with psychotic behavior	Depression
29625	Major depressive affective disorder, single episode, in partial or unspecified remission	Depression
29626	Major depressive affective disorder, single episode, in full remission	Depression
29630	Major depressive affective disorder, recurrent episode, unspecified	Depression
29631	Major depressive affective disorder, recurrent episode, mild	Depression
29632	Major depressive affective disorder, recurrent episode, moderate	Depression
29633	Major depressive affective disorder, recurrent episode, severe, without mention of psychotic behavior	Depression
29634	Major depressive affective disorder, recurrent episode, severe, specified as with psychotic behavior	Depression
29635	Major depressive affective disorder, recurrent episode, in partial or unspecified remission	Depression
29636	Major depressive affective disorder, recurrent episode, in full remission	Depression
29681	Atypical manic disorder	Other Mood Disorders
29682	Atypical depressive disorder	Other Mood Disorders
29690	Unspecified episodic mood disorder	Other Mood Disorders
29699	Other specified episodic mood disorder	Other Mood Disorders
3004	Dysthymic disorder	Other Mood Disorders
30110	Affective personality disorder, unspecified	Other Mood Disorders
30113	Cyclothymic disorder	Other Mood Disorders

References

[1] American Diabetes Association. Economic costs of diabetes in the U.S. In 2017. 2017 accessed Oct.1,2018 http://care.diabetesjournals.org/content/early/2018/03/20/dci18-0007.

[2] American Diabetes Association. Diagnosis and classification of diabetes mellitus. 2014 accessed Oct.1,2018 http://care.diabetesjournals.org/content/37/Supplement_1/S81.

[3] Gregg EW, Li Y, Wang J, Rios Burrows N, Ali MK, Rolka D, et al. Changes in diabetes-related complications in the United States, 1990–2010. N Engl J Med 2014;370(16):1514–23.

[4] Tancredi M, Rosengren A, Svensson AM, Kosiborod M, Pivodic A, Gudbjörnsdottir S, et al. Excess mortality among persons with type 2 diabetes. N Engl J Med 2015;373(18):1720–32.

[5] Ducat L, Philipson LH, Anderson BJ. The mental health comorbidities of diabetes. J Am Med Assoc 2014;312(7):691–2.

[6] Harkness E, Macdonald W, Valderas J, Coventry P, Gask L, Bower P. Identifying psychosocial interventions that improve both physical and mental health in patients with diabetes: a systematic review and meta-analysis. Diabetes Care 2010;33(4):926–30.

[7] Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. Comput Struct Biotechnol J 2017;15:104–16.

[8] Razavian N, Blecker S, Schmidt AM, Smith-McLallen A, Nigam S, Sontag D. Population-level prediction of type 2 diabetes from claims data and analysis of risk factors. Big Data 2015;3(4):277–87.

[9] Buch V, Varughese G, Maruthappu M. Artificial intelligence in diabetes care. Diabet Med 2018;35(4):495–7.

[10] Anderson JP, J.R, Parikh, Shenfeld DK, Ivanov V, Marks C, Church BW, et al. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: an application of machine learning using electronic health records. Journal of Diabetes Science and Technology 2016;10(1):6–18.

[11] Basu S, Raghavan S, Wexler DJ, Berkowitz SA. Characteristics associated with decreased or increased mortality risk from glycemic therapy among patients with type 2 diabetes and high cardiovascular risk: machine learning analysis of the ACCORD trial. Diabetes Care; 2017. p. dc172252.

[12] Makino M, Ono M, Itoko T, Katsuki T, Koseki A, Kudo M. Artificial intelligence predicts progress of diabetic kidney disease-novel prediction model construction with big data machine learning. Diabetes 2018;67(Supplement 1):539.

[13] Bennett CC, Doub TW, Selove R. EHRs connect research and practice: where predictive modeling, artificial intelligence, and clinical decision support intersect. Health Policy and Technology 2012;1(2):105–14.

[14] Mayzell G. Managed care and payer models. Population health: an implementation guide to improve outcomes and lower costs. CRC Press; 2015.

[15] Trotter P. Chronic disease is healthcare's rising-risk. 2016 accessed https://www.healthitoutcomes.com/doc/chronic-disease-is-healthcare-s-rising-risk-0001, Accessed date: 6 October 2018.

[16] Green LW. Making research relevant: if it is an evidence-based practice, where's the practice-based evidence? Fam Pract 2008;25(suppl):i20–4.

[17] Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. Artif Intell Med 2013;57(1):9–19.

[18] Bennett CC, Doub TW. Temporal modeling in clinical artificial intelligence decision-making and cognitive computing: empirical exploration of practical challenges. Proceedings of the 3rd SIAM workshop on data mining for medicine and healthcare. 2014.

[19] Bennett CC, Doub TW. Data mining and electronic health records: selecting optimal clinical treatments in practice. Proceedings of the 6th International Conference on Data Mining. 2010. p. 313–8.

[20] Bennett CC, Doub TW. Expert systems in mental healthcare: AI applications in decision making and consultation. In: Luxton DD, editor. Artificial intelligence in behavioral and mental healthcare. Elsevier Press; 2016. p. 27–51.

[21] Sabanovic S, C Bennett C, Chang WL, Huber L. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR) 2013:1–6.

[22] Bennett CC, Sabanovic S, Piatt JA, Nagata S, Eldridge L, Randall N. A robot a day keeps the blues away. IEEE International Conference on Healthcare Informatics (ICHI) 2017:536–40.

[23] Dagliati A, Marini S, Sacchi L, Cogni G, Teliti M, Tibollo V, et al. Machine learning methods to predict diabetes complications. Journal of Diabetes Science and Technology 2018;12(2):295–302.

[24] Lagani V, Chiarugi F, Thomson S, Fursse J, Lakasing E, Jones RW, et al. Development and validation of risk assessment models for diabetes-related complications based on the DCCT/EDIC data. J Diabetes Complicat 2015;29(4):479–87.

[25] Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. Comput Struct Biotechnol J 2017;15:104–16.

[26] Makino M, Ono M, Itoko T, Katsuki T, Koseki A, Kudo M, et al. Artificial intelligence predicts progress of diabetic kidney disease-novel prediction model construction with big data machine learning. Diabetes 2018;67(Supplement 1):539.

[27] Jin H, Wu S, Di Capua P. Development of a clinical forecasting model to predict comorbid depression among diabetes patients and an application in depression screening policy making. Prev Chronic Dis 2015:12.

[28] Witten IH, Frank E. Data mining: practical machine learning tools and techniques. second ed. San Francisco, CA: Morgan Kaufmann; 2005.

[29] Liu H, Waxman DA, Main R, Mattke S. Utilization of anesthesia services during outpatient endoscopies and colonoscopies and associated spending in 2003-2009. J Am Med Assoc 2012;307(11):1178–84.

[30] Abbas OA. Comparisons between data clustering algorithms. Int Arab J Inf Technol 2008;5(3):320–5.

[31] Breiman L. Random Forests. Mach Learn 2001;45:5–32.

[32] Quinlan JR. C4. 5: programs for machine learning. San Francisco, CA: Morgan Kaufmann; 2003.

[33] Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Their Appl 1998;13(4):18–28.

[34] Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics; 2017.

[35] Wrobel JS, Mayfield JA, Reiber GE. Geographic variation of lower-extremity major amputation in individuals with and without diabetes in the Medicare population. Diabetes Care 2001;24(5):860–4.

[36] Centers for Disease Control. accessed https://www.cdc.gov/diabetes/data/statistics-report/diagnosed.html; 2017, Accessed date: 7 October 2018.

[37] Keehan SP, Cuckler GA, Sisko AM, Madison AJ, Smith SD, Stone DA, et al. National health expenditure projections, 2014–24: spending growth faster than recent trends. Health Aff 2015;34(8):1407–17.
[38] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–57.
[39] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, et al. No health without mental health. The Lancet 2007;370(9590):859–77.
[40] Lin EH, Von Korff M. WHO WMH Survey Consortium. Mental disorders among persons with diabetes – results from the world mental health surveys. J Psychosom Res 2008;65(6):571–80.
[41] Kimia AA, Savova G, Landschaft A, Harper MB. An introduction to natural language processing: how you can get more from those electronic notes you are generating. Pediatr Emerg Care 2015;31(7):536–41.
[42] Kaplan B. Evaluating informatics applications–clinical decision support systems literature review. Int J Med Inform 2001;64(1):15–37.
[43] Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4(1):50.
[44] Karsh BT. Beyond usability: designing effective technology implementation systems to promote patient safety. BMJ Qual Saf 2004;13(5):388–94.
[45] Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010;43(1):159–72.
[46] Bennett CC, Doub TW, Bragg AD, Luellen J, Van Regenmorter C, Lockman J, Reiserer R. Data mining session-based patient reported outcomes (PROs) in a mental health setting: toward data-driven clinical decision support and personalized treatment. IEEE Health Informatics, Imaging, and Systems Biology Conference 2011:229–36.