
JMIR Diabetes

Emerging Technologies, Medical Devices, Apps, Sensors, and Informatics to Help People with Diabetes
Volume 1 (2016), Issue 1 ISSN 2371-4379 Editors-in-Chief: Ricardo Correa, MD, EdD; Sheyu Li, MD

Contents

Original Papers

Data Mining of a Remote Behavioral Tracking System for Type 2 Diabetes Patients: A Prospective Cohort Study (e1) Noah Wayne, Nick Cercone, Jiye Li, Ariel Zohar, Joel Katz, Patrick Brown, Paul Ritvo.	2
Information and Communication Technology-Powered Diabetes Self-Management Systems in China: A Study Evaluating the Features and Requirements of Apps and Patents (e2) Ying Li, Jin Tan, Bozhi Shi, Xiaolian Duan, Daidi Zhong, Xiaoling Li, Jianning Qu.	16
Evaluating the Accuracy of Google Translate for Diabetes Education Material (e3) Xuewei Chen, Sandra Acosta, Adam Barry.	29

Original Paper

Data Mining of a Remote Behavioral Tracking System for Type 2 Diabetes Patients: A Prospective Cohort Study

Noah Wayne^{1*}, PhD, RKin; Nick Cercone^{2†*}, PhD; Jiye Li^{2*}, PhD; Ariel Zohar^{1*}, B BSc (Hons), RKin; Joel Katz^{3*}, PhD (CPsych); Patrick Brown^{4,5*}, PhD; Paul Ritvo^{1*}, PhD (CPsych)

¹Health Behaviour Change Lab, School of Kinesiology & Health Science, York University, Toronto, ON, Canada

²Data Mining Lab, Lassonde School of Engineering, York University, Toronto, ON, Canada

³Pain Mechanisms Lab, Department of Psychology, York University, Toronto, ON, Canada

⁴Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

⁵Analytics and Informatics, Prevention and Cancer Control, Cancer Care Ontario, Toronto, ON, Canada

†deceased

*all authors contributed equally

Corresponding Author:

Paul Ritvo, PhD (CPsych)

Health Behaviour Change Lab

School of Kinesiology & Health Science

York University

136 Chemistry Building

4700 Keele Street

Toronto, ON, M3J1P3

Canada

Phone: 1 416 736 2100 ext 22396

Fax: 1 416 736 5774

Email: pritvo@yorku.ca

Abstract

Background: Complications from type 2 diabetes mellitus can be prevented when patients perform health behaviors such as vigorous exercise and glucose-regulated diet. The use of smartphones for tracking such behaviors has demonstrated success in type 2 diabetes management while generating repositories of analyzable digital data, which, when better understood, may help improve care. Data mining methods were used in this study to better understand self-monitoring patterns using smartphone tracking software.

Objective: Associations were evaluated between the smartphone monitoring of health behaviors and HbA1c reductions in a patient subsample with type 2 diabetes who demonstrated clinically significant benefits after participation in a randomized controlled trial.

Methods: A priori association-rule algorithms, implemented in the C language, were applied to app-discretized use data involving three primary health behavior trackers (exercise, diet, and glucose monitoring) from 29 participants who achieved clinically significant HbA1c reductions. Use was evaluated in relation to improved HbA1c outcomes.

Results: Analyses indicated that nearly a third (9/29, 31%) of participants used a single tracker, half (14/29, 48%) used two primary trackers, and the remainder (6/29, 21%) of the participants used three primary trackers. Decreases in HbA1c were observed across all groups (0.97-1.95%), but clinically significant reductions were more likely with use of one or two trackers rather than use of three trackers (OR 0.18, $P=.04$).

Conclusions: Data mining techniques can reveal relevant coherent behavior patterns useful in guiding future intervention structure. It appears that focusing on using one or two trackers, in a symbolic function, was more effective (in this sample) than regular use of all three trackers.

(*JMIR Diabetes* 2016;1(1):e1) doi:[10.2196/diabetes.4506](https://doi.org/10.2196/diabetes.4506)

KEYWORDS

diabetes mellitus, type 2; health coaching; mhealth; telehealth; data mining

Introduction

Diabetes is a cluster of metabolic conditions characterized by dysglycemia from defects in insulin secretion and/or unhealthy behaviors that cause debilitating complications and death [1,2]. In 2013, an estimated 8.3% of the global population lived with diabetes and more than US \$612 billion was spent on treatment [3]. A type 2 diabetes (T2DM) diagnosis, affecting 90-95% of people with diabetes, largely results from genetic predisposition, excess body weight, physical inactivity, and poor diet [1]. Additional studies demonstrate that lower socioeconomic status (SES) populations are at greater risk for developing diabetes [4,5] and often demonstrate poorer disease management, resulting in more frequent (and more expensive) complications and hospitalizations [5].

There is consensus among researchers and clinical professionals that glucose self-monitoring, exercise, and diet-related health behaviors are important in effective T2DM management and well-regulated serum glucose levels [6]. Monitoring health behaviors (ie, exercise, healthy diet, glucose monitoring) may be important in improving glucose control [6]. Mobile tracking technologies can help patients adopt and sustain self-management behaviors and can help health professionals provide better monitoring and support. However, there are challenges in determining the optimal design of self-tracking tools and their optimal use with respect to frequency and duration.

Data mining (DM) refers to analytic approaches useful in detecting coherent patterns in large and complex datasets [7]. The applications of DM methods in analyses of diabetes-related health behaviors are continually being improved, especially in the selection of analytic frameworks that capture key data with sufficient explanatory power [8].

As utilization of electronic health records increases in health care, DM becomes more relevant [9] to chronic disease prevention and management [8]. A recent study utilized descriptive DM algorithms to analyze a dataset with 450 attributes to identify a “short list” of the behavioral correlates of depressive disorder [10]. This study exemplifies use of DM in datasets lacking the uniformity needed for more conventional analyses [11]. DM can help integrate variables of multiple types (eg, diet, exercise, blood pressure, SES, income, geographic location) in better understanding factors affecting diabetes incidence, prevalence, and management [7].

In another diabetes study, DM algorithms were used to construct a model that predicted short-term changes in blood glucose [12]

exemplifying how DM can help identify risk factors for hypoglycemia [13]. Several studies found that information collected on meals, insulin therapy, and physical activity improved the prediction of blood glucose levels [12-14]. Applying DM analyses to self-monitoring data could help identify key indicators in the prevention of life-threatening hypoglycemic events [15]. DM applied to primary medical data could point to useful methods for initiating and maintaining effective T2DM treatment, leading to better resource allocation and treatment personalization [16,17].

Health coaching is a promising clinical role that stimulates and supports health behavior change in patients with varying SES, health problems and diagnosed chronic diseases. When connected with 24 hour/day/7 day/week mobile phone-based counseling, health coaching is associated with benefits for individuals affected by uncontrolled T2DM [18-20] and chronic obstructive pulmonary disease [17].

Objective

Our primary objective was to evaluate associations between the mobile phone monitoring of health behaviors, within a randomized controlled trial (RCT), and clinically significant reductions in glycated hemoglobin (HbA1c).

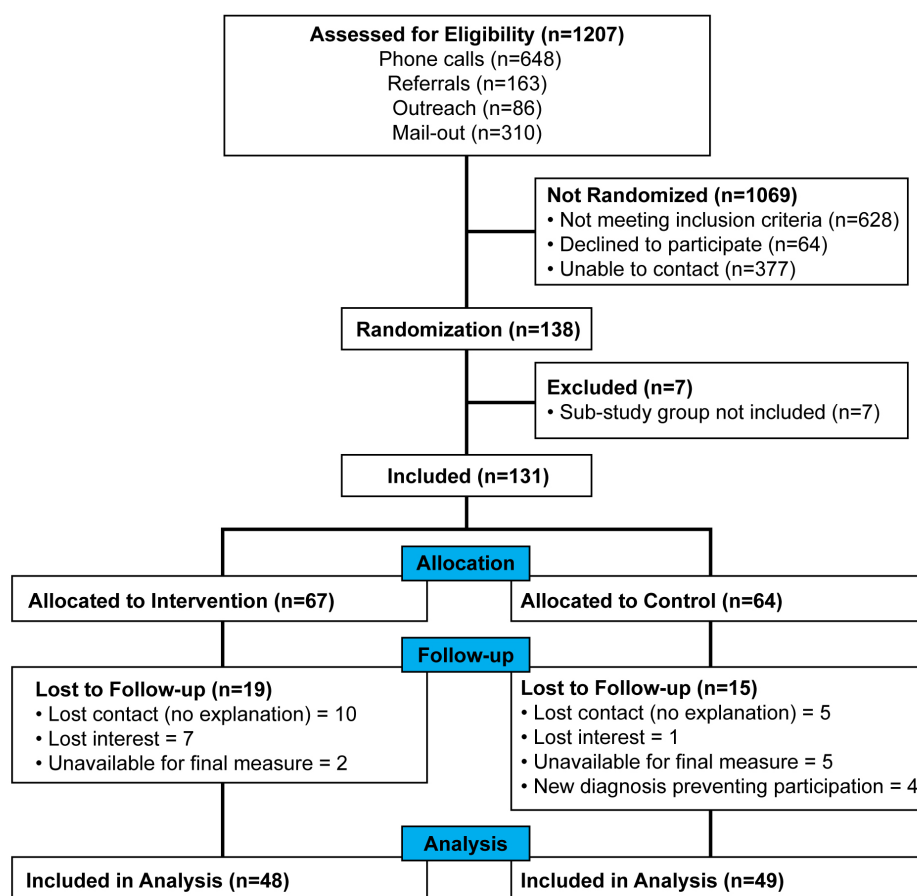
Methods

The RCT protocol was reviewed and approved by the York University research ethics board (Certificate #2012-033), and all patients provided written, informed consent to participate.

The RCT assessed T2DM patients (N=97) assisted by personal health coaches trained in behavior-change theories, practices, and counseling methods (see Figure 1). At baseline, all patients had poor glucose regulation as indicated by glycated hemoglobin (HbA1c) $\geq 7.3\%$. In the mobile phone-assisted intervention arm, participants (n=48) were provided a smartphone (Samsung Galaxy Ace 2) with data service and pre-installed health tracking software (NexJ Systems Inc., Connected Wellness Platform [CWP]) enabling the detailed monitoring of two behaviors (exercise and diet) and a risk-related outcome (blood glucose) throughout a 24-week intervention. Health coaches helped patients use the mobile phone software in ways that best fit their daily routines.

CWP use data were extracted from NexJ Systems servers upon trial completion and compiled into .csv files stored on password-protected portable drives. Study participant IDs were matched with software user IDs, as data were anonymized and prepared for analyses.

Figure 1. Flowchart of enrollment.



Data Analysis

Association Rule Algorithms

To discover useful relationships between self-tracker use and HbA1c outcomes, we employed association-rule algorithms software to find coherent relationships in transactions represented by sets of items, termed “frequent item sets.” For example, when customer A buys bread and cheese, and customer B buys bread, cheese and burgers, bread and cheese appear frequently on both shopping lists. Therefore, bread and cheese are associated and they qualify as a frequent item set [bread, cheese]. Support is a term reflecting the measurement of association frequency, as defined by the percentage of observations to which the item sets belong. In our study, support was defined as the number of times an attribute value (such as *a1c 6 month diff=1.1*) or a set of attribute values (such as *glucose_count=200* and *food_count=150*) appear in participant data, divided by the total number of participants, expressed in a percentage (ie, multiplied by 100).

We used association rule algorithms to identify all common attributes in participants. The support threshold was fixed at a minimum of 5%, such that item sets were generated that occurred in at least 5% of the sample. We used the a priori association rule algorithm implemented by C language for the Unix/Linux environment.

Attribute Selection

Among the 97 T2DM study completers, the present analysis considers patients in the experimental group (n=48). As our objective was to evaluate associations between the use frequency of different trackers in the CWP in relation to HbA1c outcomes, we selected the change in HbA1c over 24 weeks (*a1c6month_diff*) and 4 software uses as the attributes for the association rule algorithm. These attributes were use frequency of the blood glucose tracker (*glucose_count*), use frequency of the food tracker (*food_count*), use frequency of the exercise tracker (*exercise_count*), and the total use frequency of all three trackers (*generic_count*).

Discretization

We proceeded to discretize the attributes to implement the association rule algorithm. For the four tracker attributes (*glucose_count*, *food_count*, *exercise_count*, and *generic_count*), data were discretized into categories relevant to typical use during a standard week, with numerical criteria selected to reflect significant adoption levels by patients (see Table 1). Discretization was compiled differently per tracker to align with expected and actual adoption rates per tracked behavior (or outcome). For example, glucose management in T2DM includes self-monitoring of serum-glucose via finger prick, recommended several times daily for poorly managed patients, and less often if there is better gluco-regulation. Use of the blood glucose tracker was discretized based on frequencies of 1-4.9 uses per

week, 5-7.9 uses per week, 8-13.9 uses per week, and 14-21.9 uses per week. For food tracking, a similar discretization pattern was used. Since use frequency fluctuated, discretization was calculated as up to once per week, 1-3.9 uses per week, 4-6.9 uses per week, 7-13.9 uses per week (1 to <2/day), 14-20.9 uses per week (2 to <3/day), and 21+ per week (3+/day). Since the exercise tracker was the least frequently used, discretization

rules were adjusted to once every other week, 0.5-0.9 uses per week, 1-1.9 uses per week, 2-2.9 uses per week, 3-3.9 uses per week, and 4<6 uses per week. The total use of all trackers was discretized to once per week, 1-3.9 times per week, 4-6.9 times per week (0.5-1/day), 7-13.9 times per week (1-2/day), 14-17.9 times per week (2-4/day), 18-34.9 times per week (4-5/day), and 35-50 times per week (5-7/day).

Table 1. Discretization of tracker use frequency^{a-d}.

Attribute	Group	Range	Meaning
Glucose count	1	0-0	Zero
	2	5-24	Up to once per week
	3	24-120	1-4.9 times/week
	4	120-192	5-7.9 times/week
	5	192-336	8-13.9 times/week
	6	336-520	14-21.9 times/week
Exercise count	1	0-0	Zero
	2	0-5	Minimal (>0-4)
	3	5-12	Once every other week
	4	12-24	0.5-0.9 time/week
	5	24-48	1-1.9 times/week
	6	48-72	2-2.9 times/week
	7	72-96	3-3.9 times/week
	8	96-150	4-6 times/week
Food count	1	0-0	Zero
	2	1-5	>0-5 over 6 months
	3	5-24	Up to 1 time/week
	4	24-96	1-3.9 times/week
	5	96-168	4-6.9 times/week (1/day)
	6	168-336	7-13.9 times/week (1-1.9/day)
	7	336-504	14-20.9 times/week (2-2.9/day)
	8	504-720	21+ times/week (3+/day)
Generic count	1	0-0	Zero
	2	0-24	1 time/week
	3	24-96	1-3.9 times/week
	4	96-168	4-6.9 times/week (0.5-1/day)
	5	168-336	7-13.9 times/week (1-2/day)
	6	336-672	14-17.9 times/ week (2-4/day)
	7	672-840	18-34.9 times/ week (4-5/day)
	8	840-1200	35-50 times/week (5-7/day)

^aAttribute=type of tracker used.

^bGroup=discretization group.

^cRange=frequencies of use for allocation to group.

^dMeaning=frequency of use in terms of use per week/day.

Preprocessing for HbA1c Attribute

For the *alc6month_diff* attribute, conceived of as the most important attribute in investigating the desired associations, we preserved its original value to an optimal degree. Therefore, we did not discretize this attribute but created an extension of attributes based on the *alc6month_diff* attribute. We obtained a set of all possible values for *alc6month_diff* attribute (a), such as $\{-1.4, -0.8, -0.5, -0.3, -0.1, 0, 0.1, 0.2, 0.5, a_n\}$. For patients with an *alc6month_diff* value of a_1 through to a_n , we created the following new attributes: *alc6_at_least_-1.4*, *alc6_at_least_-0.8*, *alc6_at_least_-0.5*, ..., *alc6_at_least_n*. Accordingly, the patient with the lowest reduction value was associated with only one attribute value (of this kind) while the patient with the highest reduction was associated with all the attribute values established.

Postprocessing of Associations

The above pre-processing approaches produced redundancies in the input data for the association rule algorithm. Therefore, we applied three approaches to post-processing after the associations were generated. First, if a frequent item set contained only *alc6_at_least_a_n* attribute values, where a_n is a float number (from a_1 to a_n), we removed such item sets (for example, *alc6_at_least_0.4 alc6_at_least_0.2* and *alc6_at_least_-0.3 alc6_at_least_-0.8 alc6_at_least_-1.4*). The preceding two frequent item sets were thus removed because the extra *alc6_at_least* values were added by the pre-processing approach were not relevant unless generated together with the other attributes.

Second, for a frequent item set, if *alc6_at_least_a_n* the attribute values that appeared multiple times were removed except the *alc6_at_least_a_n* where n is the largest value among the “alc6” values, (for example, *glucose_120 alc6_at_least_0.5 alc6_at_least_0.3*).

In this frequent item set, the *alc6_at_least* attribute occurred twice. However, when a reduction value for “alc6” of at least 0.5 exists, by default, the reduction value of 0.3 must exist. Therefore, *alc6_at_least_0.3* was a redundant attribute value and was therefore removed. The above frequent item set becomes the following after post-processing: *glucose_120 alc6_at_least_0.5*.

Third, after considering the above two situations, redundant item sets existed in the result. For example, *glucose_120 alc6_at_least_0.7 alc6_at_least_0.4* and *glucose_120 alc6_at_least_0.7 alc6_at_least_0.1*. These two frequent item sets both contained redundant “*alc6_at_least*” attribute values. After considering the second situation, these two frequent item

sets became *glucose_120 alc6_at_least_0.7* and *glucose_120 alc6_at_least_0.7*. Since they were identical and their support values were identical, we removed one of them.

Pre- and postprocessing were implemented by Python using the Linux environment. Attribute selection and categorization were used to determine how system use was associated with change in HbA1c per participant. A minimum clinically significant change approach was used to determine what proportion of participants demonstrated 0.5% (5.5 mmol/mol) or greater reductions of HbA1c and used trackers at variable levels of intensity [21]. Therefore, patient data were extracted only with associations containing the attribute-value pair *alc_0.5*, representing changes in HbA1c values that were 0.5% or above, which was true for 29 participants.

Last, a Fisher's exact test for count data was applied to determine the statistical relationship between use of all three trackers, use of one or two trackers, and HbA1c reductions. For this test, the 10 intervention participants who used the software but did not have a clinically significant reduction in HbA1c were used as a comparison group. Results were considered significant at the $P=.05$ level.

Results

Usage Data

In total, 48 intervention patients completed the 24-week trial. Only 39 patients used the CWP software to track health behaviors. Of this software software-user sample, 29 reduced their HbA1c measure by clinically significant levels at trial conclusion, defined as 0.5% (5.5 mmol/mol) or greater (see Figure 2). Demographic and baseline characteristics of the clinically significant HbA1c reduced-users are provided in Table 2 and compared with the full intervention sample. While multiple health trackers were available in the CWP, most participants used one or a combination of two-to-three specific trackers (ie, between (1) glucose monitoring, (2) exercise tracking, and (3) food tracking) (see Figures 3-5). These trackers reflect the behaviors often considered relevant in diabetes management and therefore the trackers often recommended by health coaches for patients use. Of the 29 clinically significant HbA1c reduced-user group (selected as a subpopulation of interest [SOI]) (HbA1c reduction $\geq 0.5\%$), 2 singularly used the food tracker and 7 singularly used the glucose tracker, while no subjects singularly used the exercise tracker. Meanwhile, 11 of these subjects used both glucose and exercise trackers, 3 used both glucose and food trackers, while 0 subjects used both food and exercise trackers. Last, 6 subjects used all three trackers.

Table 2. Subject demographics.

	SOI (n=29)	Full intervention sample (n=48)
Age in years, mean (range) SD	53.4 (26-68) 10.7	53.1 (26-74) 10.9
HbA1c, % (mmol/mol) SD		
Baseline	8.88% (73.6) 1.30	8.69% (71.5) 1.32
6 months	7.52% (58.7) 0.95	7.88% (62.6) 1.17
Reduction (baseline to 6 months)	1.36% (14.9)	0.81% (8.9)
Sex, n (%)		
Male	11 (33)	17 (35)
Female	18 (66)	31 (65)
Education, n (%)		
Less than high school	5 (17)	10 (21)
High school	14 (48)	17 (35)
College diploma or vocational training	6 (21)	11 (23)
University degree	4 (14)	8 (17)
Did not disclose	0 (0)	2 (4)
Car access, n (%)		
Owns a car	10 (34)	19 (40)
Has access to car	8 (28)	9 (19)
No access to car	11 (38)	19 (40)
Not disclosed	0 (0)	1 (2)
Employment status, n (%)		
Unemployed	10 (35)	16 (33)
Student	1 (3)	3 (6)
Part-time	0 (0)	1 (2)
Full-time	8 (28)	13 (27)
Retired	3 (10)	6 (13)
Self-employed	5 (17)	6 (13)
Work in the home (take care of children)	2 (7)	2 (4)
Not disclosed	0 (0)	1 (2)
Income in CAD\$, n (%)		
0-9999	4 (14)	9 (19)
10,000-25,000	8 (28)	10 (21)
25,000-50,000	7 (24)	12 (25)
50,000-75,000	2 (7)	3 (6)
75,000-up	3 (10)	4 (8)
Did not disclose	5 (17)	10 (21)

As seen in [Table 3](#), the HbA1c reductions of subjects who used single trackers (food-tracker only or glucose-tracker only) did not significantly differ from each other. More subjects (11/48, 22.9%) used the glucose and exercise tracker in combination than the glucose and food tracker in combination (3/48, 6.3%), but there were no significant differences in HbA1c levels in these subjects. The 6 of 48 subjects (12.5%) who used all three

trackers (glucose/food/exercise) achieved a mean HbA1c reduction of 1.55%.

The Fisher's exact test for count data indicated subjects who used the software (n=39) were more likely to achieve a clinically significant reduction in HbA1c if they used one or two trackers than if they used all three trackers (OR 0.18, $P=.04$) (see [Table 4](#)).

Table 3. Tracker usage pattern.

Mobile phone tracker	Users, n	Mean reduction of HbA1c, %	SD
Food only	2	1.95	1.20
Glucose only	7	1.74	1.00
Exercise only	0	0	0
Glucose + exercise	11	0.97	0.38
Glucose + food	3	1.07	0.45
Food + exercise	0	0	0
Glucose + food + exercise	6	1.55	0.49

Table 4. Fisher's exact test for count data on three trackers versus less^a.

	Achieved clinical HbA1c reduction (≥0.5%)	Did not achieve clinical HbA1c reduction (<0.5%)
All 3 trackers	6	6
<3 trackers	23	4

^aOR 0.18 and $P=0.04$.

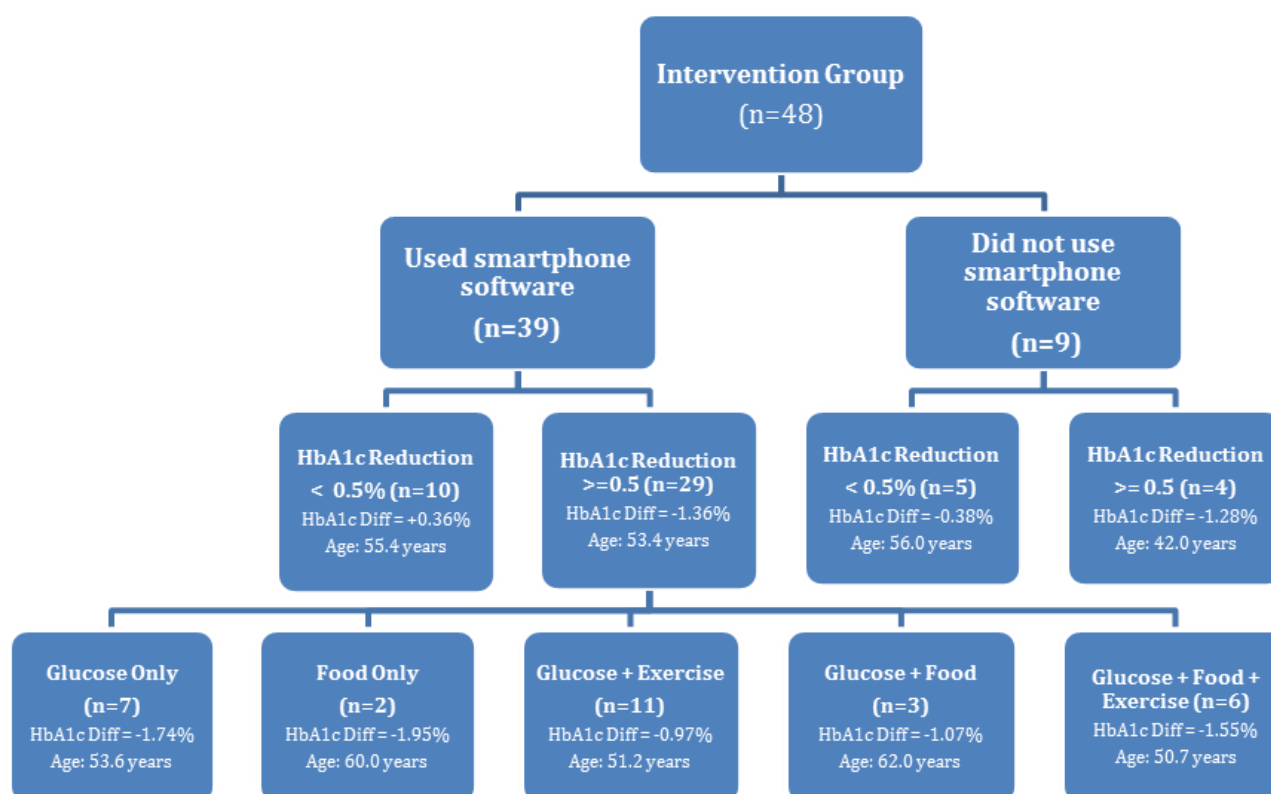
Figure 2. Breakdown of smartphone usage group.

Figure 3. Food tracker.

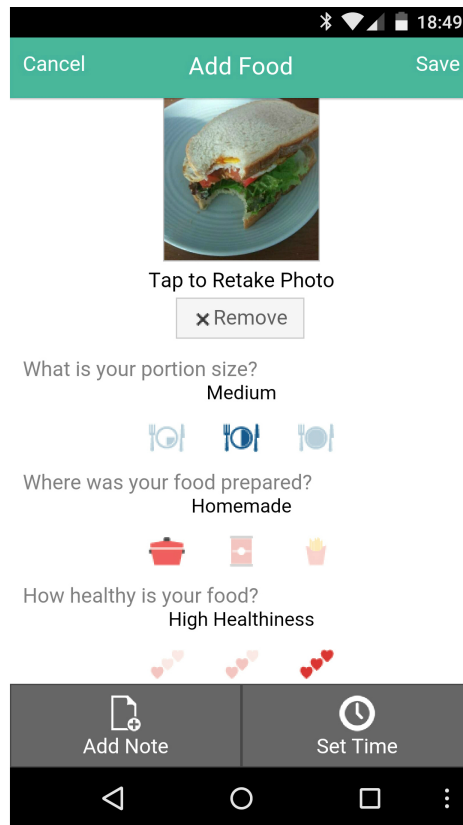


Figure 4. Blood glucose tracker.

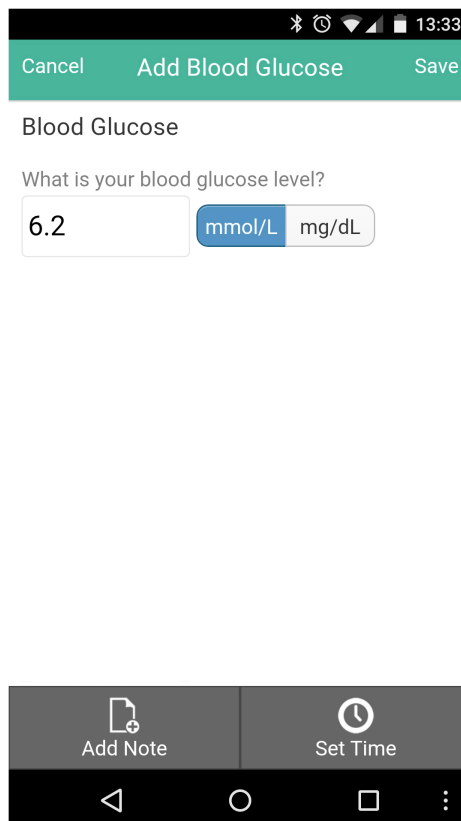


Figure 5. Exercise tracker.

Cancel Add Exercise Save

Exercise

How intense was your exercise?
Moderate

How long did you exercise?
45 minutes

How many calories did you burn?
calories

Add Note Set Time

Data Mining

The glucose tracker was the most frequently used CWP function. Altogether, 22.9% (11/48) of subjects tracked their blood glucose with the software 1-4.9 times per week *and* had a clinically significant reduction in HbA1c (>0.5%) (see [Table 5](#)) while an additional 22.9% with >0.5% reduction in HbA1c demonstrated more frequent glucose tracker use (6/48, 12.5%) and used it 5-7.9 times per week while 10.4% (5/48) used it 8-13.9 times per week.

The food tracking function followed a similar pattern, although with slightly less frequent use. We found 9/48 (18.8%) of

participants who achieved at least a 0.5% reduction in HbA1c used the food tracker a minimal amount (one time or less per week), while 7/48 (14.6%) used the system 1-3.9 times per week, and 4/48 (8.3%) used the system 4-6.9 times per week (see [Table 6](#)).

The least used tracker (and the only tracker that was never singularly used) was the exercise tracker, which was used by 7/48 (14.6%) of the intervention participants between 0.5-0.9 times per week ([Tables 7 and 8](#)). When it was paired with glucose tracking, however, every patient who used the exercise tracker had clinically significant reductions in HbA1c (see [Figure 6](#)).

Table 5. Glucose tracker use.

HbA1c diff.	Glucose 1-4.9 times/week		Glucose 5-7.9 times/week		Glucose 8-13.9 times/week	
	Users, %	n	Users, %	n	Users, %	n
0.1	29.2	14	12.5	6	12.5	6
0.2	27.1	13	12.5	6	12.5	6
0.3	25	12	12.5	6	12.5	6
0.4	22.9	11	12.5	6	12.5	6
0.5	22.9	11	12.5	6	10.4	5
0.6	20.8	10	10.4	5	10.4	5
0.7	18.8	9	10.4	5	10.4	5
0.8	16.7	8	10.4	5	10.4	5
0.9	16.7	8	10.4	5	10.4	5
1.1	14.6	7	10.4	5	8.3	4
1.2	10.4	5	8.3	4		
1.3	10.4	5	8.3	4		
1.5	8.3	4	8.3	4		

Table 6. Food tracker use.

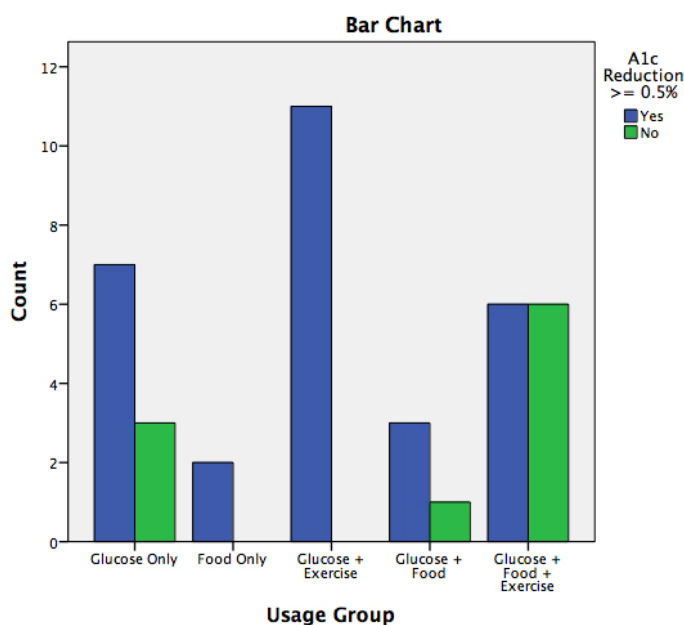
HbA1c diff.	Food up to 1x/wk		Food 1-3.9x/wk		Food 4-6.9x/wk	
	Users, %	n	Users, %	n	Users, %	n
0.1	20.8	10	16.7	8	8.3	4
0.2	18.8	9	16.7	8	8.3	4
0.3	18.8	9	14.6	7	8.3	4
0.4	18.8	9	14.6	7	8.3	4
0.5	18.8	9	14.6	7	8.3	4
0.6	16.7	8	12.5	6	8.3	4
0.7	16.7	8	10.4	5	8.3	4
0.8	16.7	8	10.4	5	8.3	4
0.9	14.6	7	8.3	4	8.3	4
0.11	14.6	7				
0.12	8.3	4				
0.13	8.3	4				

Table 7. Exercise tracker use—Part 1.

HbA1c diff.	Exercise up to 1x/wk		Exercise <5x in 6 months		Exercise 1x every other wk	
	Users, %	n	Users, %	n	Users, %	n
0.1	16.7	8	16.7	8	14.6	7
0.2	16.7	8	14.6	7	14.6	7
0.3	14.6	7	14.6	7	14.6	7
0.4	14.6	7	12.5	6	12.5	6
0.5	14.6	7	12.5	6	12.5	6
0.6	12.5	6	12.5	6	8.3	4
0.7	12.5	6	12.5	6	8.3	4
0.8	12.5	6	12.5	6	8.3	4
0.9	12.5	6	10.4	5	8.3	4
0.11	12.5	6	10.4	5	8.3	4
0.13	8.3	4	10.4	5		
0.15			8.3	4		

Table 8. Exercise tracker use—Part 2.

HbA1c diff.	Exercise 0.5-0.9x/wk		Exercise 1-1.9x/wk		Exercise 2-2.9x/wk	
	Users, %	n	Users, %	n	Users, %	n
0.1	14.6	7	8.3	4	8.3	4
0.2	14.6	7	8.3	4	8.3	4
0.3	14.6	7			8.3	4
0.4	14.6	7			8.3	4
0.5	14.6	7				
0.6	14.6	7				
0.7	12.5	6				
0.8	12.5	6				
0.9	10.4	5				
0.11	8.3	4				

Figure 6. Tracker usage group by clinical HbA1c reduction ($\geq 0.5\%$).

Discussion

Principal Findings

In the emergent field of mobile phone–assisted health coaching, a recurring question is whether the expense of integrating mHealth technology justifies the benefits. This highlights the need to further investigate for whom (which subpopulations) these technologies are useful and which technologies are most useful for them. In addressing these questions, this study pilots a method of investigating RCT participants, focusing on those who derived clinically significant benefit from participation *and* made significant use of the mobile phone software. The study specifies how much use they engaged in, and the associated benefits, using the clinically significant reduction in HbA1c criteria of 0.5% (5.5 mmol/mol) or greater.

By employing attribute categorizations, tracker-use frequencies were determined per subject in an SOI of significant users who also derived clinically significant glucose reductions. This dataset enabled comparison of subjects who used single trackers versus multiple trackers (two or three trackers). Descriptive results indicated that in subjects achieving a 0.5% HbA1c reduction or better, singular use of the glucose and food monitoring was undertaken, while, in contrast, in the same group, no singular use of the exercise tracker was undertaken. In dual-tracker use, glucose and exercise trackers were employed by more subjects than glucose and food tracking. Interestingly, food and exercise (as dual) trackers were not used by any of the subjects who met criteria for 0.5% HbA1c reduction or greater. Last, all three trackers were used by 12.5% (6/48) of subjects and were associated with a substantive HbA1c mean reduction (1.55% or 16.9 mmol/mol).

The Fisher's exact test for count data indicated that for subjects who used the software ($n=39$) and used all three trackers, there was a significantly lower likelihood of achieving a clinically

significant reduction in HbA1c than for those who used a lesser number (ie, one or two trackers) (OR 0.18, $P=.04$). This finding indicates that it might be advisable for patients with type 2 diabetes to focus on one or two trackers, especially if one tracker assesses blood glucose levels. From a behavioral perspective, these data could influence health coaching recommendations for health behavior tracking.

Limitations

Data mining is often used to process large amounts of data. One limitation of the pilot application of data mining in this study was the relatively small user sample size. Nonetheless, the association rule algorithm technique offers a foundation with which to study larger datasets of mHealth tracking technologies as they become available. In terms of diabetes intervention, this was an RCT of typical size (48 intervention participants) and, altogether 10,695 uses of the mobile phone app were analyzed (about 62 uses per month per participant who used the software). Although future DM studies may address larger datasets, this pilot demonstrated application in an RCT dataset within which >10,000 app uses were analyzed (averaging ~1 use per day).

Conclusion

In summary, this study points to a future when the mobile monitoring of health behaviors will increase and provide digital signals representing engagement in discrete behaviors and daily-weekly-monthly outcomes. Whereas previous associations between counseling and outcomes were difficult to obtain and often based on retrospective self-report, mobile phone monitoring offers ongoing records that precisely reflect status improvements, their stability, and fluctuations (eg, relapsing patterns). Altogether, with the increasing collection of wearable data, we may derive a quantifiable perspective on health changes that instructs the patient and health coach in improving chronic disease management.

Acknowledgments

The authors would like to thank NexJ Systems Inc. for their partnership in this trial and for the use of the Connected Wellness Platform as a clinical research tool. Funding was provided by the Public Health Agency of Canada and the Federal Development Agency of Southern Ontario. We offer special thanks to the staff of the Black Creek Community Health Centre and trial participants from the Jane-Finch community of Toronto, Ontario.

Joel Katz is supported by a Canadian Institutes of Health Research Canada Research Chair in Health Psychology.

The authors would like to acknowledge with sadness the untimely passing of study co-author, colleague, and friend, Dr. Nicholas Cercone. Dr. Cercone's expertise and mentorship on data mining theory and technique was invaluable. His inspiring and supportive presence will be deeply missed.

Conflicts of Interest

None declared.

References

1. Seaquist ER. Addressing the burden of diabetes. *JAMA* 2014 Jun 11;311(22):2267-2268. [doi: [10.1001/jama.2014.6451](https://doi.org/10.1001/jama.2014.6451)] [Medline: [24915253](https://pubmed.ncbi.nlm.nih.gov/24915253/)]
2. Riddle MC, Ambrosius WT, Brillon DJ, Buse JB, Byington RP, Cohen RM, Action to Control Cardiovascular Risk in Diabetes Investigators. Epidemiologic relationships between A1C and all-cause mortality during a median 3.4-year follow-up of glycemic treatment in the ACCORD trial. *Diabetes Care* 2010 May;33(5):983-990 [FREE Full text] [doi: [10.2337/dc09-1278](https://doi.org/10.2337/dc09-1278)] [Medline: [20427682](https://pubmed.ncbi.nlm.nih.gov/20427682/)]
3. International Diabetes Foundation. In: Guariguata L, Nolan T, Beagley J, Linnenkamp U, Jacqmain O, editors. *IDF Diabetes Atlas (6th Edition)*. Brussels, Belgium: International Diabetes Federation; 2013.
4. Dinca-Panaitescu S, Dinca-Panaitescu M, Bryant T, Daiski I, Pilkington B, Raphael D. Diabetes prevalence and income: Results of the Canadian Community Health Survey. *Health Policy* 2011 Feb;99(2):116-123. [doi: [10.1016/j.healthpol.2010.07.018](https://doi.org/10.1016/j.healthpol.2010.07.018)] [Medline: [20724018](https://pubmed.ncbi.nlm.nih.gov/20724018/)]
5. Hux J, Booth G, Slaughter P, Laupacis A. *Diabetes in Ontario: An ICES Practice Atlas*. Toronto: Institute for Clinical Evaluative Science; 2003. URL: <http://www.ices.on.ca/~media/Files/Atlases-Reports/2003/Diabetes-in-Ontario/Full%20report.ashx> [accessed 2015-04-06] [WebCite Cache ID 6XabgpTYm]
6. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee. *Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada*. *Can J Diabetes* 2013;37:S1-S212 [FREE Full text]
7. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* 2013 Apr 3;309(13):1351-1352. [doi: [10.1001/jama.2013.393](https://doi.org/10.1001/jama.2013.393)] [Medline: [23549579](https://pubmed.ncbi.nlm.nih.gov/23549579/)]
8. Georga E, Protopappas V, Mougiakakou S, Fotiadis D. Short-term vs. long-term analysis of diabetes data: Application of machine learning and data mining techniques. In: *IEEE. 2013 Presented at: 13th International Conference on Bioinformatics and Bioengineering (BIBE); November 10-13, 2013; Chania p. 1-4.* [doi: [10.1109/BIBE.2013.6701622](https://doi.org/10.1109/BIBE.2013.6701622)]
9. Koh HC, Tan G. Data mining applications in healthcare. *J Healthc Inf Manag* 2005 Jan;19(2):64-72. [Medline: [15869215](https://pubmed.ncbi.nlm.nih.gov/15869215/)]
10. Yoon S, Taha B, Bakken S. Using a data mining approach to discover behavior correlates of chronic disease: a case study of depression. *Stud Health Technol Inform* 2014;201:71-78. [Medline: [24943527](https://pubmed.ncbi.nlm.nih.gov/24943527/)]
11. Yoo I, Alafairet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang J, et al. Data mining in healthcare and biomedicine: a survey of the literature. *J Med Syst* 2012 Aug;36(4):2431-2448. [doi: [10.1007/s10916-011-9710-5](https://doi.org/10.1007/s10916-011-9710-5)] [Medline: [21537851](https://pubmed.ncbi.nlm.nih.gov/21537851/)]
12. Georga EI, Protopappas VC, Ardigo D, Marina M, Zavaroni I, Polyzos D, et al. Multivariate prediction of subcutaneous glucose concentration in type 1 diabetes patients based on support vector regression. *IEEE J Biomed Health Inform* 2013 Jan;17(1):71-81. [doi: [10.1109/TITB.2012.2219876](https://doi.org/10.1109/TITB.2012.2219876)] [Medline: [23008265](https://pubmed.ncbi.nlm.nih.gov/23008265/)]
13. Zecchin C, Facchinetti A, Sparacino G, Cobelli C. Reduction of number and duration of hypoglycemic events by glucose prediction methods: a proof-of-concept in silico study. *Diabetes Technol Ther* 2013 Jan;15(1):66-77. [doi: [10.1089/dia.2012.0208](https://doi.org/10.1089/dia.2012.0208)] [Medline: [23297671](https://pubmed.ncbi.nlm.nih.gov/23297671/)]
14. Pappada SM, Cameron BD, Rosman PM, Bourey RE, Papadimos TJ, Olorunto W, et al. Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes. *Diabetes Technol Ther* 2011 Feb;13(2):135-141. [doi: [10.1089/dia.2010.0104](https://doi.org/10.1089/dia.2010.0104)] [Medline: [21284480](https://pubmed.ncbi.nlm.nih.gov/21284480/)]
15. Barrett MA, Humblet O, Hiatt RA, Adler NE. Big Data and Disease Prevention: From Quantified Self to Quantified Communities. *Big Data* 2013 Sep;1(3):168-175. [doi: [10.1089/big.2013.0027](https://doi.org/10.1089/big.2013.0027)]
16. Marinov M, Mosa ASM, Yoo I, Boren SA. Data-mining technologies for diabetes: a systematic review. *J Diabetes Sci Technol* 2011 Nov;5(6):1549-1556 [FREE Full text] [Medline: [22226277](https://pubmed.ncbi.nlm.nih.gov/22226277/)]
17. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de WL. A pilot study of a tool to stimulate physical activity in patients with COPD or type 2 diabetes in primary care. *J Telemed Telecare* 2014 Jan;20(1):29-34. [doi: [10.1177/1357633X13519057](https://doi.org/10.1177/1357633X13519057)] [Medline: [24414397](https://pubmed.ncbi.nlm.nih.gov/24414397/)]

18. Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. *J Med Internet Res* 2014;16(6):e149 [FREE Full text] [doi: [10.2196/jmir.3180](https://doi.org/10.2196/jmir.3180)] [Medline: [24907918](https://pubmed.ncbi.nlm.nih.gov/24907918/)]
19. Wayne N, Perez DF, Kaplan DM, Ritvo P. Health Coaching Reduces HbA1c in Type 2 Diabetic Patients From a Lower-Socioeconomic Status Community: A Randomized Controlled Trial. *J Med Internet Res* 2015;17(10):e224 [FREE Full text] [doi: [10.2196/jmir.4871](https://doi.org/10.2196/jmir.4871)] [Medline: [26441467](https://pubmed.ncbi.nlm.nih.gov/26441467/)]
20. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care* 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: [10.2337/dc11-0366](https://doi.org/10.2337/dc11-0366)] [Medline: [21788632](https://pubmed.ncbi.nlm.nih.gov/21788632/)]
21. Karyekar CS, Frederich R, Ravichandran S. Clinically relevant reductions in HbA1c without hypoglycaemia: results across four studies of saxagliptin. *Int J Clin Pract* 2013 Aug;67(8):759-767 [FREE Full text] [doi: [10.1111/ijcp.12212](https://doi.org/10.1111/ijcp.12212)] [Medline: [23795975](https://pubmed.ncbi.nlm.nih.gov/23795975/)]

Abbreviations

CWP: Connected Wellness Platform

DM: data mining

HbA1c: glycated hemoglobin

RCT: randomized controlled trial

SOI: subpopulation of interest

T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 06.04.15; peer-reviewed by TW Ho, M Spreeuwenberg, F Lai; comments to author 29.07.15; revised version received 28.10.15; accepted 08.01.16; published 06.04.16.

Please cite as:

Wayne N, Cercone N, Li J, Zohar A, Katz J, Brown P, Ritvo P

Data Mining of a Remote Behavioral Tracking System for Type 2 Diabetes Patients: A Prospective Cohort Study

JMIR Diabetes 2016;1(1):e1

URL: <http://diabetes.jmir.org/2016/1/e1/>

doi: [10.2196/diabetes.4506](https://doi.org/10.2196/diabetes.4506)

PMID: [30291054](https://pubmed.ncbi.nlm.nih.gov/30291054/)

©Noah Wayne, Nick Cercone, Jiye Li, Ariel Zohar, Joel Katz, Patrick Brown, Paul Ritvo. Originally published in *JMIR Diabetes* (<http://diabetes.jmir.org>), 06.04.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Diabetes*, is properly cited. The complete bibliographic information, a link to the original publication on <http://diabetes.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Information and Communication Technology-Powered Diabetes Self-Management Systems in China: A Study Evaluating the Features and Requirements of Apps and Patents

Ying Li¹, ME; Jin Tan¹, ME; Bozhi Shi¹, ME; Xiaolian Duan², MD; Daidi Zhong¹, PhD; Xiaoling Li¹, BE; Jianning Qu³, MD

¹Key Laboratory of Biorcheological Science and Technology, Ministry of Education, Bioengineering College, Chongqing University, Chongqing, China

²Chongqing Academy of Science & Technology, Chongqing, China

³The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

Corresponding Author:

Xiaolian Duan, MD

Chongqing Academy of Science & Technology

2Yangliu Road, Huangshan Avenue

New North Zone

Chongqing, 401123

China

Phone: 86 2367300872

Fax: 86 2377665868

Email: dxl@cast.gov.cn

Abstract

Background: For patients with diabetes, the self-monitoring of blood glucose (SMBG) is a recommended way of controlling the blood glucose level. By leveraging the modern information and communication technology (ICT) and the corresponding infrastructure, engineers nowadays are able to merge the SMBG activities into daily life and to dramatically reduce patient's burden. Such type of ICT-powered SMBG had already been marketed in the United States and the European Union for a decade, but was introduced into the Chinese market only in recent years. Although there is no doubt about the general need for such type of SMBG in the Chinese market, how it could be adapted to the local technical and operational environment is still an open question.

Objective: Our overall goal is to understand the local requirements and the current status of deploying ICT-powered SMBG to the Chinese market. In particular, we aim to analyze existing domestic SMBG mobile apps and relevant domestic patents to identify their various aspects, including the common functionalities, innovative feature, defects, conformance to standards, prospects, etc. In the long run, we hope the outcome of this study could help the decision making on how to properly adapt ICT-powered SMBG to the Chinese market.

Methods: We identified 289 apps. After exclusion of irrelevant apps, 78 apps remained. These were downloaded and analyzed. A total of 8070 patents related to glucose were identified from patent database. Irrelevant materials and duplicates were excluded, following which 39 patents were parsed to extract the important features. These apps and patents were further compared with the corresponding requirements derived from relevant clinical guidelines and data standards.

Results: The most common features of studied apps were blood health data recording, notification, and decision supporting. The most common features of studied patents included mobile terminal, server, and decision supporting. The main difference between patents and apps is that the patents had 2 specific features, namely, interface to the hospital information system and recording personal information, which were not mentioned in the app. The other major finding is that, in general, in terms of the components of the features, although the features identified in both apps and patents conform to the requirements of the relevant clinical guidelines and data standards, upon looking into the details, gaps exist between the features of the identified apps and patents and the relevant clinical guidelines and data standards. In addition, the social media feature that the apps and patents have is not included in the standard requirements list.

Conclusions: The development of Chinese SMBG mobile apps and relevant patents is still in the primitive stage. Although the functionalities of most apps and patents can meet the basic requirements of SMBG, gaps have been identified when comparing the functionalities provided by apps and patents with the requirements necessitated by the standards. One of the most important

gaps is that only a small portion of the studied apps provides the automatic data transmission and exchange feature, which may hamper the overall performance. The clinical guidelines can thus be further developed to leverage new features provided by ICT-powered SMBG apps (eg, the social media feature, which may help to improve the social intervention of patients with diabetes).

(*JMIR Diabetes* 2016;1(1):e2) doi:[10.2196/diabetes.4475](https://doi.org/10.2196/diabetes.4475)

KEYWORDS

apps; diabetes; information and communication technology; self-monitoring of blood glucose; patents

Introduction

Self-Monitoring of Blood Glucose

According to World Health Organization data statistics, in 2014, the global prevalence of diabetes was estimated to be 9% among adults [1]. In most developed countries, diabetes is the fourth highest fatal disease, whereas in developing countries, the number of people living with diabetes is on the rise [2]. Being a developing country, China has the largest population of diabetics in the world [3,4]. With the increasing prevalence of diabetes, various diabetes-related complications have become the main reason for disability or death in these patients [5].

Results of clinical studies have shown that controlling glucose level based on a predefined rule can reduce the incidence and development of chronic complications of diabetes, but whether glucose level can remain stable in an allowed range in daily life depends on diabetes self-management [6,7]. The self-monitoring of blood glucose (SMBG), as a recommended diabetes self-management tool, plays an important role in the self-management of diabetes. It can provide real-time glucose information and timely intervention that can help patients maintain normal blood sugar levels. The demonstrated clinical value of SMBG for patients with diabetes has been widely recognized by stakeholders. Accordingly, the International Diabetes Federation [8], the American Diabetes Association [9], and the National Institute for Clinical Excellence [10] have stressed that SMBG is an integral part of a comprehensive diabetes management, and they recommend that all patients should practice SMBG. However, the data from SMBG can significantly affect physician's decision making because, in their opinion, the influence of the SMBG data for treatment decisions is equal or greater than the glycosylated hemoglobin (HbA1c) levels [11-15]. The guidelines [16] proposed by the Ministry of Health (MoH) also promoted various subtypes of SMBG scheme applicable to patients with diabetes.

Information and Communication Technology

In 2014, there were about 1.5 billion mobile device users worldwide. A year ago, this number was only 1.1 billion. A previous report [17] indicated that P.R. China had the largest amount of mobile phone users in 2013 and the number is still increasing. The information and communication technology (ICT) [18], characterized by multimedialization, popularization, diversity, personalization, and globalization [19], is an attractive platform for health promotion and disease management interventions. The same applies to diabetes self-management as well.

The Chinese government acts positively in deploying ICT in the health care domain, as it has already issued several regulatory policies in recent years, such as the "Twelfth Five (2011-2015)" National Strategic Emerging Industry Development Plan [20], the Biomedical industry development plan [21], and the Opinions on Promoting the Development of Health Services Industry [22]. The top policymakers in the country are pushing stakeholders to develop portable health data-collection devices, with ability to connect to the Internet (eg, via Wi-Fi or mobile Internet), and to increase the deployment of automatic and intelligent health information service.

SMBG System Powered by ICT

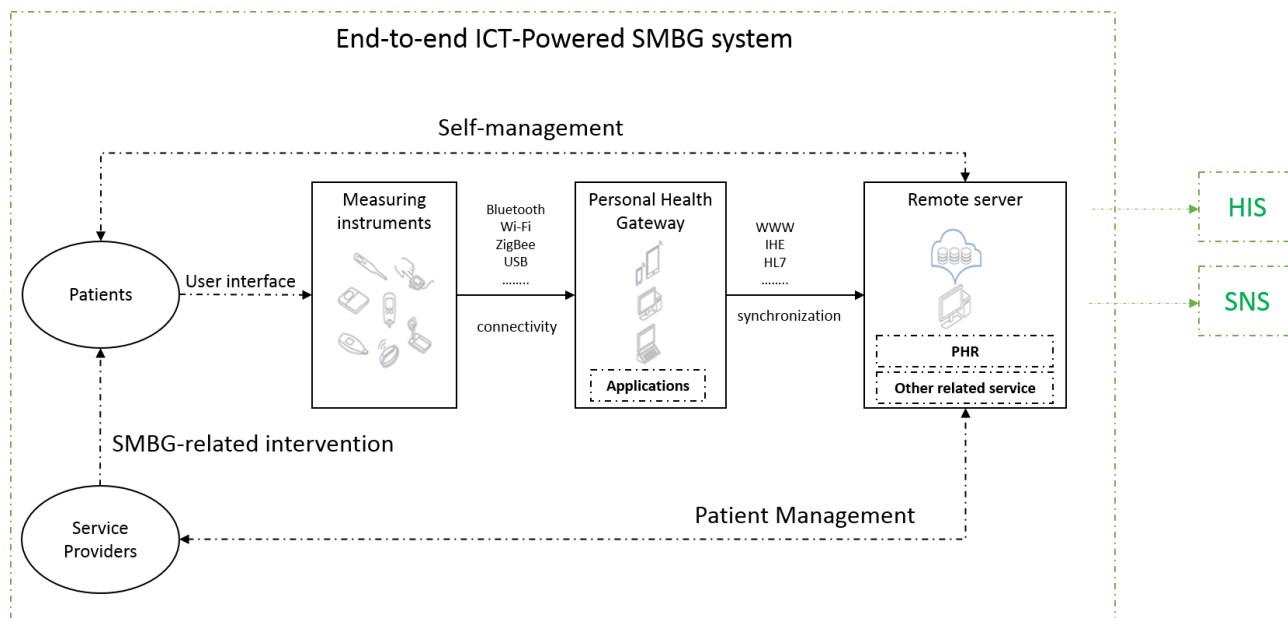
When practicing SMBG, patients normally monitor and manage their glucose level by themselves at home. Such an activity may benefit from an ICT-powered SMBG system, which allows patients to transmit their data to a service provider via an end-to-end data channel. In most cases, such a system is composed of one or more measuring instruments (eg, a glucose meter), a gateway device (eg, a mobile device), and a remote server. Patients can collect their glucose measurements and other related health data with glucose meter and other instruments. The common devices used here are blood glucose monitor (BGM) and continuous glucose monitor (CGM). The remote server contains personal health record (PHR) and other related services. The established connectivity between measuring instruments and personal health gateway, and the connectivity between gateways and remote server, together populate an end-to-end data channel. Through this channel, the collected data can be transmitted from patients to service providers via uplink, and the instructions from service providers can be sent to patients via downlink. This allows patients and service providers to access the health data any time. Service providers can provide further appropriate interventions to patients based on certain data-driven strategy. Moreover, the external partners contain the hospital information system (HIS), the social network site, etc. Figure 1 shows a general architecture of such ICT-powered end-to-end system. It is a simplified version of the architecture described in Continua Design Guidelines [23], which has been widely adopted by vendors in the European Union and the United States. With the rapid development of ICT, there are many mobile apps about disease management in developed countries [24-26], especially about diabetes self-management [27].

The SMBG powered by the aforementioned system can effectively assist patients with lifestyle changes and reduce patients' reliance on doctors. In addition, it allows the service providers to acquire the information about patients in a timely

manner, to promote patients' self-management and to generate revenue for primary care stations. This technology can not only set patients free, but can also reduce the financial burden of hospitalization and medical resource consumption. Furthermore, previous studies have shown that such system may lead to positive clinical effectiveness. Hua [28] and Yan [29] carried out experiments using the ICT-powered SMBG system, and the experimental group outperformed the control group in terms of

the clinical efficiency of diabetes treatment, the stability of controlled blood glucose level, and the efficacy of drug usage. Last, but not least, with the help of the aforementioned system, service providers now have great potential to conduct personalized data mining and analysis based on the acquired health care information, and thus are able to provide various customized treatment solutions [30].

Figure 1. The user flow of the SMBG practice based on an end-to-end system.



Contents and Objectives

The ICT-powered SMBG had already been marketed in the United States and the European Union for a decade, but was introduced into the Chinese market only in recent years. Although there is no doubt about the general need for such a type of SMBG in the Chinese market, how it could be adapted to the local technical and operational environment is still an open question.

Our overall goal is to understand the local requirements and the current status of deploying ICT-powered SMBG to the Chinese market. In particular, we aim to analyze existing domestic SMBG mobile apps and relevant domestic patents to identify their various aspects, including the common functionalities, innovative feature, defects, conformance of standards, prospects, etc. In the long run, we hope that the outcome of this study could help with the decision making on how to properly adapt ICT-powered SMBG to the Chinese market.

Although the end-to-end ICT-powered SMBG system is the target of this study, we employed mobile apps as one of our data resources. From the architecture presented in Figure 1, it can be seen that the mobile app is a key component. Its features may well reflect the features of the entire system. Moreover, it is much easier to acquire metadata of mobile apps, than those of the end-to-end system. This can be achieved by searching the app stores or Internet.

Methods

This section describes how the apps, the patents, and the standards were identified, and what selection strategy had been applied in this study. It also shows how the studied apps and patents were evaluated and analyzed.

Apps and Patents Selection Strategy

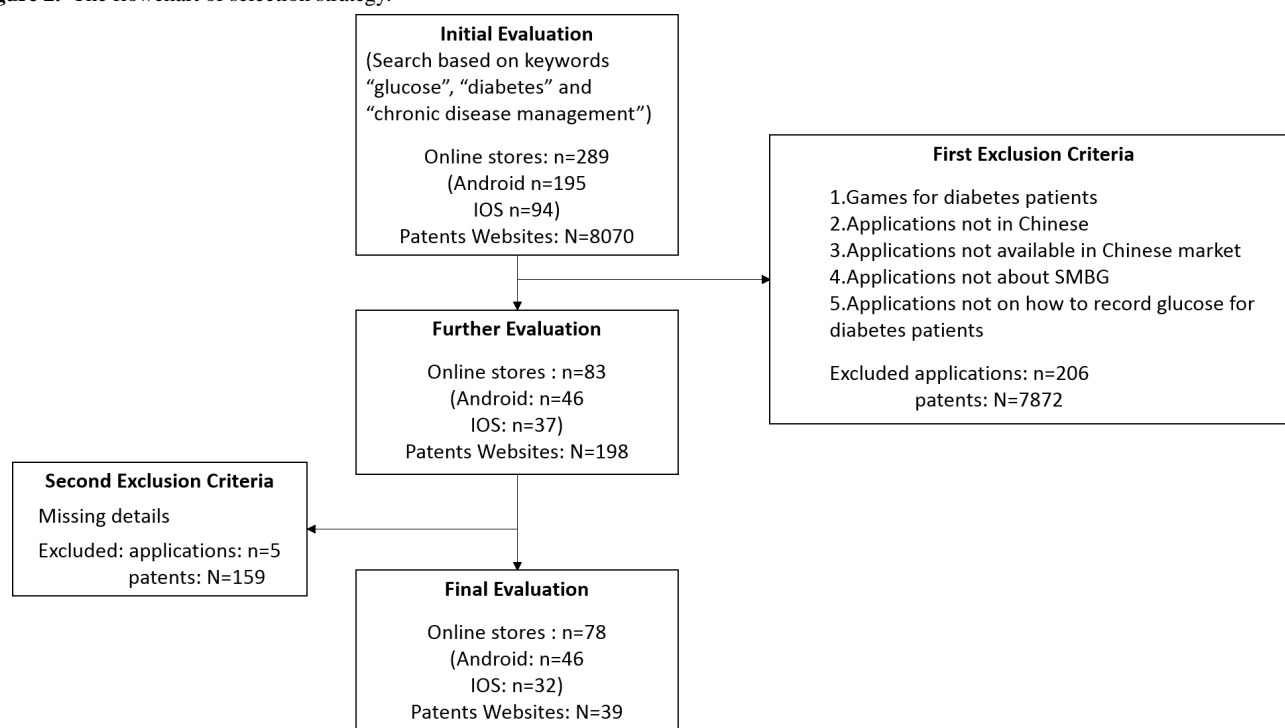
In this study, the search resources are online stores for mobile apps and Chinese national patents database [31]. We searched for diabetes-related apps from globally available app stores including Google Play Store, iTunes, BlackBerry World, Windows Phone, and Nokia's store. The domestic app stores in PR China such as Huawei Store [32] and Xiaomi Store [33] were also searched. We identified 289 apps and 8070 patents from app store and patents database, respectively, using the following keywords: "glucose," "diabetes," and "chronic disease management." Among the search result of apps, in particular, the Google Android market occupies the largest portion (n=195), followed by the Apple iOS (n=94); no results were retrieved from the Blackberry and Windows Phone store.

In the next step, we eliminated general health and lifestyle apps and game apps, because they were not relevant to the SMBG practice (eg, the calorie calculator, glucose calculator). We also excluded apps that are not designed for the Chinese market and are not in Chinese. This is to ensure that all identified apps are available in the Chinese market and that they present the information in Chinese. Finally, we excluded apps that are not related to recording glucose levels for patients with diabetes.

Thus, after the first round of selection, a total of 83 apps and 198 patents remained. The second round of selection aimed at identifying apps on the management of diabetes without much

missing details. A total of 78 apps (Android: n=46 and iOS: n=32) and 39 patents were selected. Figure 2 shows the inclusion and exclusion criteria applied in this study.

Figure 2. The flowchart of selection strategy.



Standards Selection Strategy

The application of various standards plays an important role in regulating the product design, product usage, and clinical practices. Therefore, conformance to the requirements imposed by the respective standard is essential to ensure that the ICT-powered SMBG system implemented is both functionally and clinically suitable. Besides, adherence to necessary standards is also a necessary condition for such implementation to become scalable and widely deployable.

To meet these rationales, 2 types of standards are considered in this study: the clinical guidelines and the data standards. The former ensures that the ICT-powered SMBG implementations are capable of meeting the service provider's operational and clinical needs, and the latter ensures that multiple technical elements (devices and services) work together in an interoperable way. In the context of any end-to-end system having an architecture similar to Figure 1, the so-called data standards actually refer to a family of technical standards, including data exchange protocol, wired or wireless transport, data dictionary, nomenclature code, domain information model, device profile, service configuration, testing procedure, etc. However, providing their technical detail is out of the scope of this article.

To identify relevant data standards, we searched a database of national standard [34], and 3 databases of ministry-level standards, including the MoH [35], Ministry of Industry and Information Technology [36], and China Food and Drug Administration (CFDA) [37] databases. The only reference we found is the *basic dataset of disease management*, which was

published by the MoH in 2012 [38]. However, this standard is purely a data dictionary specifically designed for capturing intrahospital information exchange (eg, HIS) data. The content of the standard alone is not able to support all the operations and data flows generated by an ICT-powered end-to-end system.

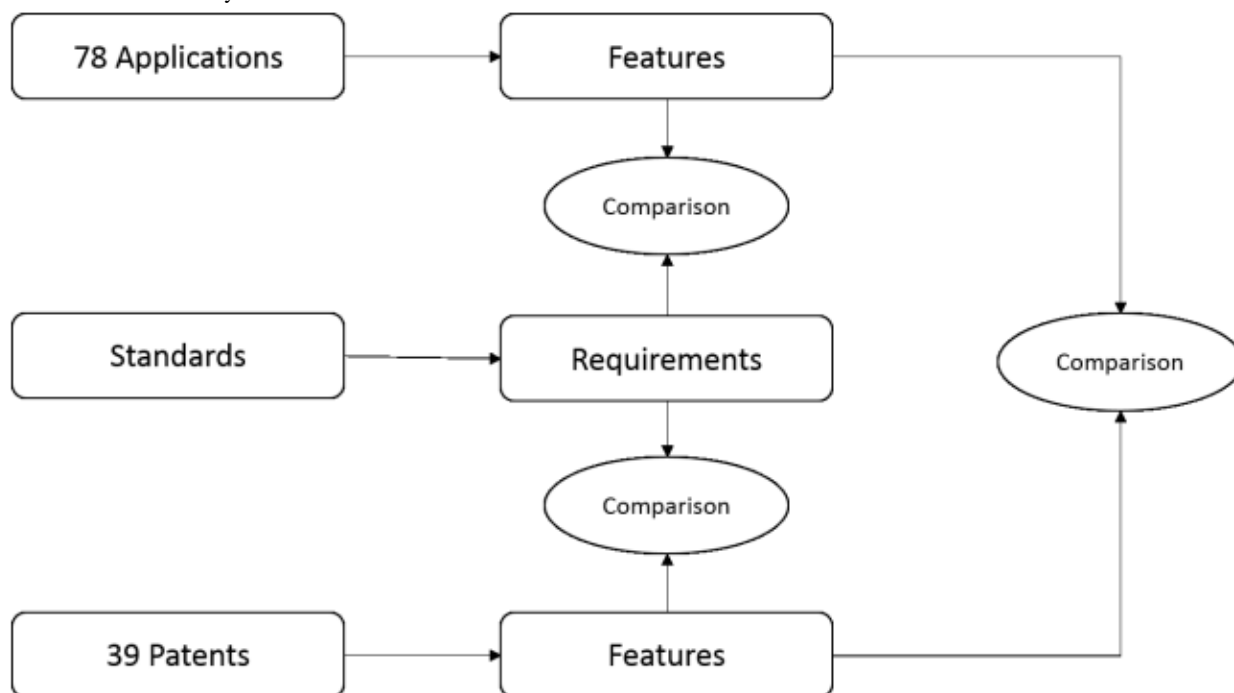
Therefore, as a backup, we employed the following 2 international data standards, whose scopes are similar to what we are looking for, as the baseline for this study: IEEE 11073-10417-2015 (BGM) [39] and IEEE 11073-10425-2014 (CGM) [40], both of which were published by the Institute of Electrical and Electronics Engineers (IEEE) in 2015 and 2014, respectively, and are now being adopted by the International Organization for Standardization.

For the clinical guideline, we identified 2 relevant guidelines [41-42] published by the Chinese Diabetes Society in 2011 and 2012, respectively.

Evaluation and Analysis Method

After obtaining the apps and patents, we summarized the features of the apps and patents, and also extracted relevant requirements from standards. The features of these apps and patents are listed in Multimedia Appendices 1 and . The requirements exacted from standards are listed in the “Data Standards” section, and the following comparisons are presented: (1) comparison between the features of apps and the features of patents; (2) comparison between the features of apps and the requirements of standards; (3) comparison between the features of patents and the requirements of standards. The flowchart of the analysis method we employed is shown in Figure 3.

Figure 3. The flowchart of analysis method.



Results

Data Standards

The data standards we used in this study were IEEE 11073-10417-2015 (BGM) and IEEE 11073-10425-2014 (CGM). Within these 2 data standards, the key measurements, data format, nomenclature codes, data exchange method, device parameters, and device profiles are defined to ensure the interoperability between BGM/CGM devices, gateway devices, and backend servers. Figures 4 and 5 show the domain information model of the 2 standards. Such models may help us to deduce the data-related requirements for SMBG apps, and such requirements come from industry consensus built from the standard development process.

The aforementioned basic dataset of disease management only contains basic glucose level and HbA1c measurement, whereas the 2 IEEE standards contain far more measurements and contextual information, which may better satisfy the needs of SMBG. Such supplementary information can be summarized and categorized as follows:

1. *Direct measurements:* glucose, blood pressure;
2. *Indirect measurement or manual input:* diet, medications, exercise, symptoms, tester, sample locations; and
3. *Instrument information:* interval, author code, alert, calibrations, device status, and solution.

Figure 4. The domain information model in IEEE Std.11073-10417-2015.

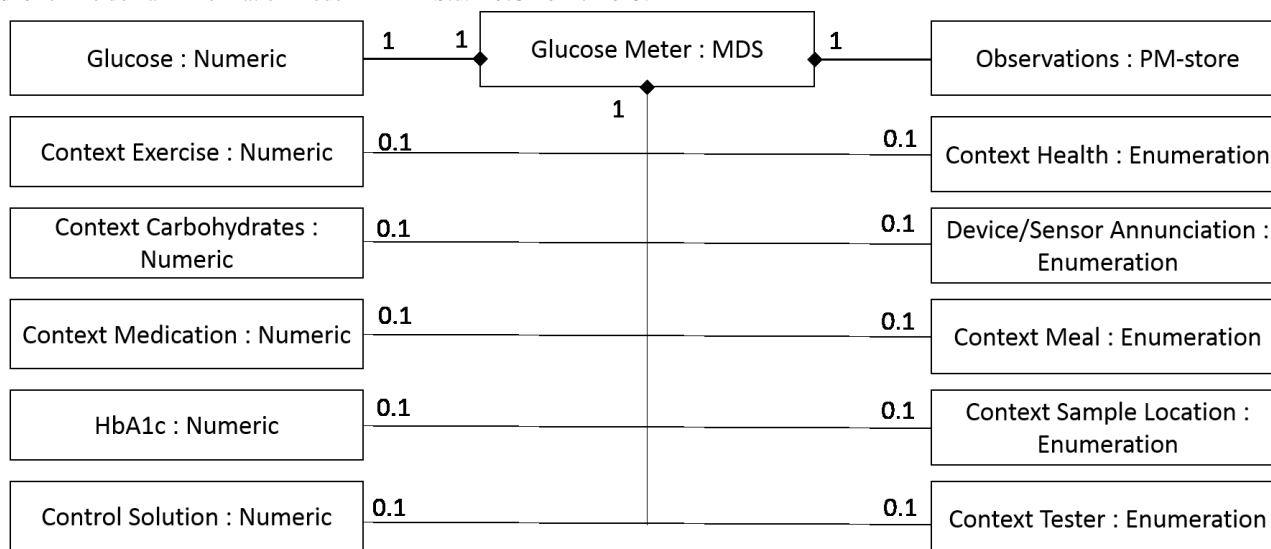
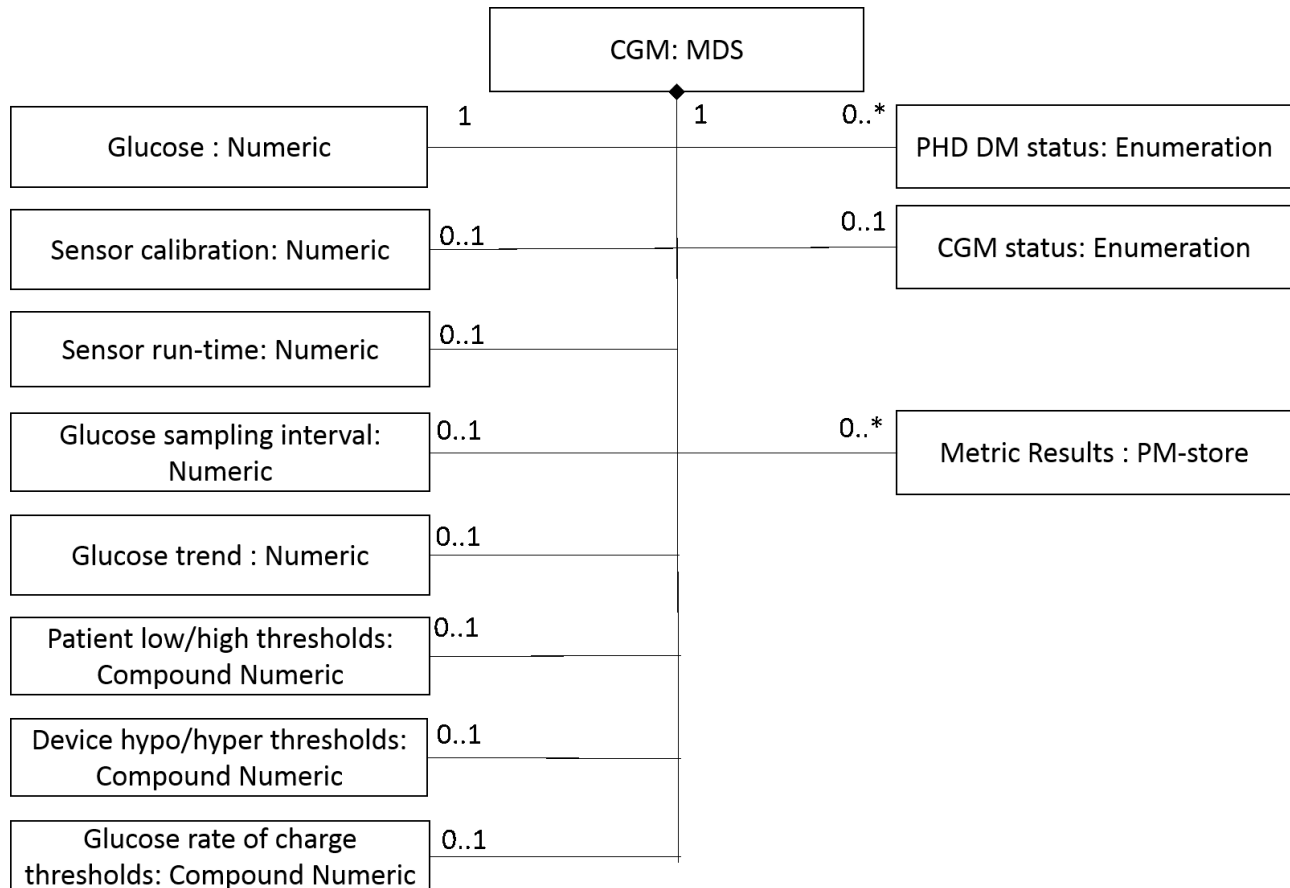


Figure 5. The domain information model in IEEE Std.11073-10425-2014.

Clinical Guidelines

The clinical guidelines used in this study were the Chinese Glucose Monitoring Clinical Guideline and the Chinese Continuous Glucose Monitoring Clinical Guideline. These 2 clinical guidelines define the data that service providers need to know from the patients in SMBG, and provide a standard of glucose monitoring for patients and doctors, including measure counts, exact measuring time, data format, and calculated statistical information of CGM, which is the most useful data for accurate diagnosis. Such standardized presentation of SMBG data can largely help clinicians to understand the conditions of patients with diabetes.

The requirements suggested by such clinical guidelines, either directly or indirectly, can be summarized and categorized as follows:

1. *Accuracy-related information*: accuracy of BGM device, contextual factors of BGM, and accuracy of CGM;
2. *Glucose-related statistics*: frequency and time point of glucose monitoring, mean value, standard deviation, volatility, and absolute difference;
3. *Other necessary context information*: diet, exercise, weight, blood pressure, medications, and symptoms.
4. Data evaluation and analysis method;
5. Decision-supporting information;

6. Diabetes-related educational material;

7. PHR.

Functionality of the End-to-End SMBG System Suggested by Referenced Standards

By jointly considering the data standards and the clinical guidelines, we have identified the following requirements on functionalities:

1. *Self-test*: This includes blood glucose, blood pressure, weight, diet status, record of exercise, record of medications, record of symptoms;
2. Diabetes-related education material;
3. Personalized context-based notification and alert;
4. Decision-supporting information;
5. Data synchronization to PHR; and
6. Data evaluation and analysis method.

The aforementioned functionalities are further elaborated in [Table 1](#). Although these functionalities are mutually exclusive, they have the potential to work as an integrated solution to support SMBG practices. For example, the user can log into the personal portal to view the records of glucose level, blood pressure, weight, diet, exercise, medication, and symptoms to evaluate how these factors are inter-related or how they may affect his/her blood glucose condition.

Table 1. Descriptions of the functionalities implied by standards.

No	Function	Input	Description	Source ^{a,b}
1.1	Blood glucose	Blood glucose values timestamp	User enters values and can view graphs, with low, high, and normal ranges well demarcated.	D/C
1.2	Blood pressure	Blood pressure timestamp	User enters values and can view graphs, with low, high, and normal ranges well demarcated.	D/C
1.3	Weight	Weight timestamp	User enters values and can view graphs.	D/C
1.4	Diet status	Food eaten, carbohydrate timestamp	User enters values and manually selects the type and amount of food.	D/C
1.5	Record of exercise	Exercise type, intensity, duration	User enters values and manually selects the type and intensity of exercise.	D/C
1.6	Record of medications	Medications type, amount, timestamp	User enters values and manually selects the type and amount of drugs.	D/C
1.7	Record of symptoms	Symptoms timestamp	User enters values and manually enters daily physical condition.	C
2	Diabetes-related educational material	Tips, feedback, diabetes health information	Most apps linked to another app for educational material or used Web links. Some had decision support and personalized tips and feedback.	C
3	Personalized context-based notification and alert	Injection reminders, medication reminders, measuring reminder	User enters reminders manually, and also receives reminders for postprandial testing. Some had automatic tailored alerts.	D
4	Decision-supporting information	Emails between doctors and patients	Data can be shared between doctors and patients, and doctors can give remote guidance.	C
5	Data synchronization to personal health record (PHR)	PHR	Synchronization with personal health systems, data can be shared and stored in the cloud platform for easy access.	C
6	Data evaluation and analysis method	The signs data	The data can be presented in a graph to show the effects of data to blood glucose for doctors to determine.	C

^aC: Functionalities from the clinical guidelines.

^bD: Functionalities from the data standards.

Mobile Apps

A total of 78 identified apps have been downloaded, and their functionalities were studied. Most of these apps provide similar functionalities, such as blood health data recording, notification, and offer decision-making support. Most functionalities listed in [Table 1](#) can also be found in the studied apps.

Among them, the blood glucose recording is the most common functionality, which is obvious. A self-management portal allowing patients to store and review various diabetes-related personal health information (such as blood pressure, weight, exercise, diet, medications, and symptoms) appears to be another popular feature in many apps. Furthermore, the personalized notification is included in some apps.

Some of the studied apps provide diabetes-related analytical result and decision-making support for users. Various functionalities such as providing guidelines regarding the treatment and management of diabetes, multimodal educational materials (eg, videos, forums), and PHR synchronization are also included in some apps to examine patients' lifestyle.

A less presented, but interesting, functionality is the social media support, such as the option to email, text message, or any type of communication between stakeholders (eg, patients, service providers, patient's relatives, friends). Few apps also provide data connectivity, which allows for end-to-end data transmission. [Table 2](#) shows the presence of functionalities in the studied apps (N=78).

Table 2. Percentage of studied apps having different features (N=78).

Features	n	%
Blood glucose	78	100
Blood pressure	58	74
Weight	45	58
Notification and alert	42	54
Decision supporting	35	45
Record of exercise	34	44
Diet status	33	42
Social media	30	38
Record of medications	28	36
Diabetes-related education material	19	24
Data evaluation and analysis method	14	18
Data synchronization to personal health record	12	15
Record of symptoms	7	9
Bluetooth	10	13

Patents

A total of 39 identified patents have been downloaded and their functionalities were studied. The described solutions in most of the studied patents are designed as an end-to-end SMBG system. Some of them even go beyond the PHR server and interact with HIS. The personal information (eg, patient's medical records, health insurance information) is also recorded in some of the studied patents.

The studied patents have many functionalities in common. In addition, many functionalities present in the studied apps are also included in the studied patents, such as decision-supporting information, notification and alert, analysis, record of personal health information, and social media. [Table 3](#) shows the most dominating features and the percentage of the studied patents (N=39) that have these tools.

Table 3. Percentage of the studied patents having different features (N=39).

Features	n	%
Mobile terminal	33	85
Server	27	69
Bluetooth	9	23
Decision supporting	10	26
Notification and alert	7	18
Data evaluation and analysis method	6	15
Social media	5	13
Record of personal information	5	13
Record of personal health information	4	10
Interface to the hospital information system	7	18
Diabetes-related educational material	6	15

Discussion

Comparison of Functionalities Between Apps and Standards

All the functionalities related to data summary are shown in [Tables 1 and 2](#), which indicate good conformance of the studied apps to the aforementioned referenced standards. However, for

some other types of functionalities, the situation is different. The details are discussed in the following sections.

Connectivity

The connectivity is a key component in an end-to-end SMBG system. According to the results presented in [Table 2](#), it can be seen that only 13% (10/78) of the studied apps have connectivity; however, all of these use Bluetooth [43]. The deployment of traditional manual input may have negative

impacts, such as poor usability, wrong data entry, and barriers of integration. Most apps have no connectivity to the measuring instruments; a possible reason for this may be the obsolete mind-set of domestic app developers and instrument manufacturers. When building their own products, their design logic is purely app centric or device centric, rather than having a holistic vision established over an interconnected infrastructure. By contrast, as leading tech companies, Google recently released Google Fit [44] and Apple released Health Kit [45]; both of them leverage connectivity technologies to simplify or automate the health data-collection process. It is quite obvious that market players are keen on adopting such functionality in this domain. We believe a similar trend will appear in the future in the Chinese market, but the questions are just “when” and “how.”

End-to-End Data Synchronization Function

As shown in Table 2, only 15% (12/78) of the apps have the end-to-end data synchronization function that allows the data generated by devices to be transmitted and stored in the PHR database, possibly via certain gateway device. Among these 15% apps (12/78), only 1 provides synchronization with a third-party platform (the Microsoft Health Vault [46]); all others connect to their own proprietary health portal.

Despite the benefits of PHR synchronization reported in some studies [47,48], the adoption rate of such functionality is less than what we initially expected, partially because of end-to-end usability issues. Our interpretation of the possible reason is that gateway device vendors are quite dominant in the current Chinese ICT market (including the ecosystem of ICT-powered SMBG), and the app vendors mainly focus on the data synchronization with the gateway device, rather than on the cloud platform. Such strategy helps them to quickly build a relationship with gateway device vendors, and start to share the revenue in the near future; however, this may not be an optimal strategy in the long run, because the service providers are likely to be situated behind the cloud. Without the service providers, the ICT-powered SMBG implementations will become much less attractive to the users. To improve the aforementioned usability issue, joint effort from all the stakeholders (operators, integrators, medical device manufacturers, app developers, health care providers, etc) must be in place to establish a data synchronization channel.

Standardized Data Structure

Among all the portals we have studied, the core diabetes-related data structures are proprietary. None of them is consistent with the data structure suggested by the aforementioned referenced data standards. In this situation, such portals could not be integrated into arbitrary HISs in a massive scale, and thus, these data cannot be exchanged. As a result, most of the portals have only minimal dataset, which is insufficient to support the decision making of clinicians. Luckily, the MoH has noticed this gap, and has begun to tackle it. We expect to see the first wave of standards released within 3-4 years.

Social Media

Findings from this study suggest that only 38% (30/78; Table 2) of apps contain social media functionality. Most apps that

include social media support only provide data-sharing interface to the well-known Chinese social networking services such as WeChat [49] and Weibo [50]. It is hard to see any other interaction between the studied SMBG apps and the linked social networking services. We have also noticed that it is not easy to share graphs and data via the studied SMBG apps with friends or relatives in social networks.

However, we still consider the availability of this functionality as a positive sign, because the psychological motivation coming from others is a good stimulator for patient’s self-management. For example, Chen [51] reported the importance of social aspects and experience sharing among people with diabetes. It is interesting to find that the social media functionality available in these studied apps is not listed in Table 1, that is, it is not included in the current clinical guidelines. We are eager to see how and when such innovative functionality coming from the ICT world could someday result in a change in the traditional clinical world.

Analysis of Patents

According to our study, glucose measuring module, mobile terminal, and remote server contained in most patents constitute the general technical architecture of the end-to-end and closed-loop SMBG systems.

Because inventors often focus sharply on the innovative contents, most patents in our study did not necessarily contain all the SMBG functions defined by referenced standards. The claims of most patents just introduce some common functionalities of the SMBG system. The social media functionality is also found in the contents of studied patents, which is not included in Table 1.

Comparison Between the Studied Patents and Apps

Although many of the functionalities identified in this study were found to be common in app and patents, functionalities such as “interface to the HIS” and “recording personal information” have been found to be missing in many apps. Figure 6 shows the distribution of functionalities in the studied apps and patents.

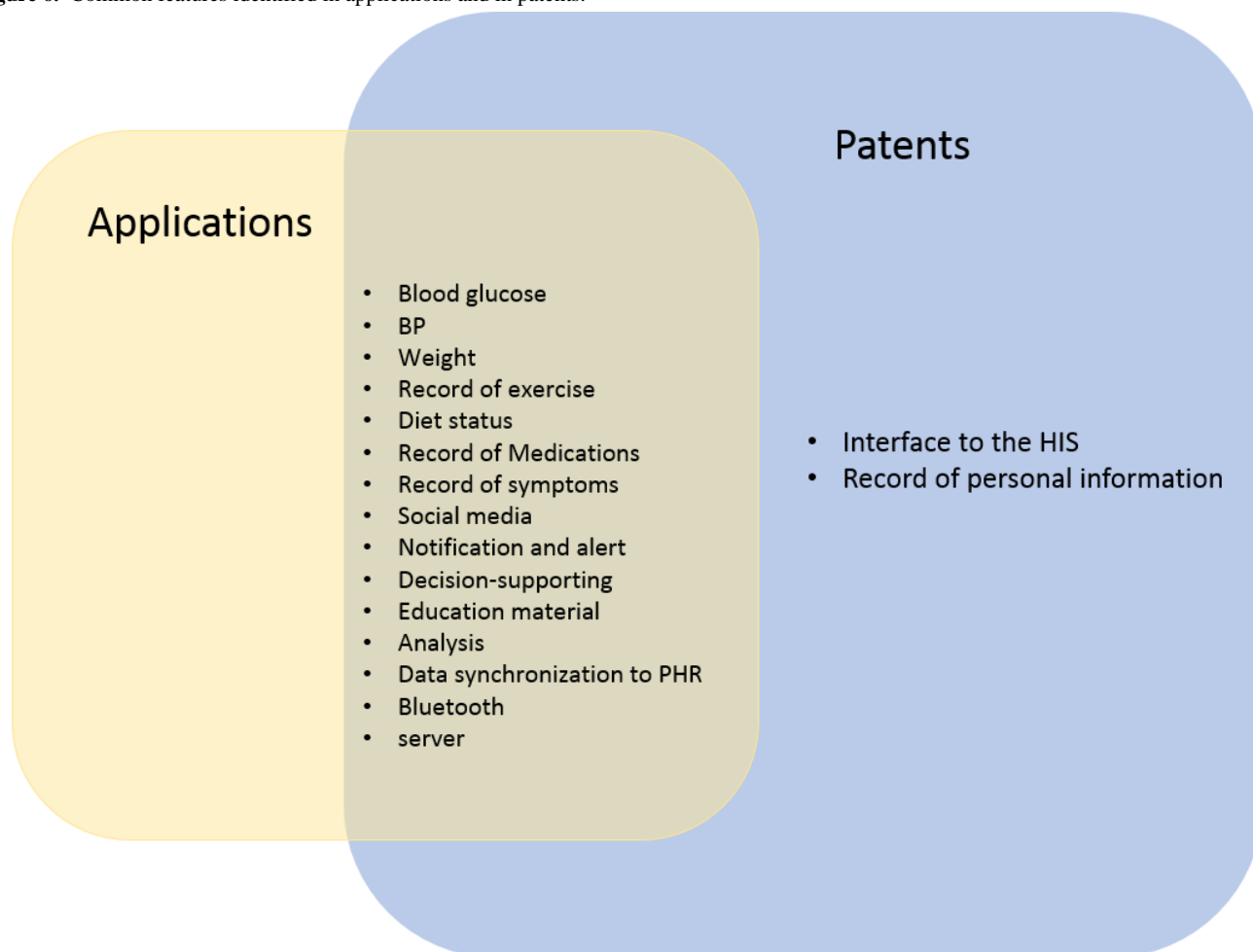
This finding suggests that when writing the patents, the big picture of end-to-end SMBG system is considered; however, when it comes to implementation, only practically feasible requirements are considered due to many practical difficulties (eg, the connectivity interface between the hospital and apps). At present, regulators in both MoH and CFDA have not yet realized the potential operational and security risks caused by ICT-powered SMBG systems, and thus, no regulatory policy includes any holistic requirement taking the end-to-end scenario into consideration. However, by judging the recent policy developments in the European Union and the United States, we believe the situation will be changed soon in China.

Compared with the patents (N=33), less portion of apps (n=10) contains the data transmission function. This is because writing a patent does not require any actual implementation of functionality. By contrast, app developers have to prioritize their tasks according to reality. They tend to implement any functionality that is feasible to be implemented at that moment.

In particular, for implementing the data transmission function, the developers may confront difficulties in engaging its supplier,

interfacing with peer device, and the corresponding development cycle might be lengthy and costly.

Figure 6. Common features identified in applications and in patents.



Limitations

Many of the apps found in the online stores were not available for installation. As a result, some of the functionalities were recorded only from the verbal description or from published articles. Often there are discrepancies between the text description and the actual features, and some functionalities are not apparent until the app is installed and tested.

Another limitation may be that only patents in Chinese were included in this study. In fact, many scholars and manufactures in China file their patents internationally. These international patents have not been searched in our study.

This study only focused on the functionality of apps and patents in Chinese. The clinical efficiency and cost efficiency, however, are not evaluated. We will address these limitations in our future study.

Conclusions

In this article, we reviewed the ICT-powered SMBG systems in PR China. We reviewed 289 apps and 8070 patents. Eventually, 78 apps and 39 patents were analyzed and compared in multiple aspects, including innovative functionalities, defects, prospects, conformance to data standards, conformance to clinical guidelines, etc.

The main finding is that the apps still have some missing links, including the connectivity, the end-to-end data synchronization function, and the standardized data structure. The other major finding of this study is that the glucose measuring module, mobile terminal, and remote server contained in most patents constitute the general technical architecture of the end-to-end and closed-loop SMBG systems. The other difference between the patents and apps is that a couple of features—“interface to the HIS” and “record of personal information”—mentioned in the patents do not appear in the apps. It is also interesting to find that the social media feature that the apps and patents have is not included in the standard requirements list. In the future, the clinical guidelines can be further developed to leverage new features provided by ICT-powered SMBG apps.

The findings provide insights into the design and development of the ICT-powered diabetes self-management system to assist patients suffering from diabetes. Hence, to be able to come up with a good design of a diabetes support tool, all the aforementioned features have to be incorporated into a single app for good monitoring, better follow-up by health care professionals, and better lifestyle management for patients with diabetes.

Acknowledgments

Partial funding was provided by Chongqing University (Project No 2011BAI14B04 and CSTC2013JCSF10034). As the sponsor of this study, Chongqing University did not involve in the review and approval of the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Diabetes self-management applications list.

[[XLSX File \(Microsoft Excel File\), 14KB - diabetes_v1i1e2_app1.xlsx](#)]

Multimedia Appendix 2

Diabetes self-management patents list.

[[XLSX File \(Microsoft Excel File\), 11KB - diabetes_v1i1e2_app2.xlsx](#)]

References

1. Jing W, Wei G. The new progress in diagnosis and treatment of diabetes: Interpreting the ADA diabetes treatment guidelines (2010 edition) [in Chinese]. *Chin J Gen Pract* 2010;9(10):668-671. [doi: [10.3760/cma.j.issn.1671-7368.2010.10.003](https://doi.org/10.3760/cma.j.issn.1671-7368.2010.10.003)]
2. Pan XR, Yang WY, Li GW, Liu J. Prevalence of diabetes and its risk factors in China, 1994. National Diabetes Prevention and Control Cooperative Group. *Diabetes Care* 1997 Nov;20(11):1664-1669. [Medline: [9353605](#)]
3. Li L, Rao K, Kong L, Yao C, Xiang H, Zhai F, Technical Working Group of China National Nutrition Health Survey. A description on the Chinese national nutrition and health survey in 2002. *Zhonghua Liu Xing Bing Xue Za Zhi* 2005 Jul;26(7):478-484. [Medline: [16334996](#)]
4. Kean W, Tianlin L, Hongding X. Study on the epidemiological characteristics of diabetes mellitus and IGT in China. *Chin J Epidemiol* 1998;19(5):282-285. [doi: [10.3760/j.issn:0254-6450.1998.05.007](https://doi.org/10.3760/j.issn:0254-6450.1998.05.007)]
5. News Medical. China, Now the Largest Diabetic Population in the World. URL: <http://www.news-medical.net/news/20100325/China-now-the-largest-diabetic-population-in-the-world.aspx> [accessed 2015-03-24] [WebCite Cache ID [6XGgJZEG3](#)]
6. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. *Patient Educ Couns* 2002;48(2):177-187. [Medline: [12401421](#)]
7. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [Medline: [12867348](#)]
8. International Diabetes Federation. Guideline on Self-Monitoring of Blood Glucose in Non-Insulin Treated Type 2 Diabetes. URL: http://www.idf.org/webdata/docs/SMBG_EN2.pdf [accessed 2015-03-24] [WebCite Cache ID [6XGfq11pt](#)]
9. ADA. Executive summary: Standards of medical care in diabetes--2010. *Diabetes Care* 2010 Jan;33 Suppl 1:S4-10 [FREE Full text] [doi: [10.2337/dc10-S004](https://doi.org/10.2337/dc10-S004)] [Medline: [20042774](#)]
10. National Institute for Clinical Excellence. Clinical Guidelines for Type 2 Diabetes Mellitus: Management of Blood Glucose 2010. URL: <http://www.nice.org.uk/> [accessed 2016-02-26] [WebCite Cache ID [6XGgEppz8V](#)]
11. Halimi S. Benefits of blood glucose self-monitoring in the management of insulin-dependent (IDDM) and non-insulin-dependent diabetes (NIDDM). Analysis of the literature: mixed results. *Diabetes Metab* 1998 Nov;24 Suppl 3:35-41 [FREE Full text] [Medline: [9881230](#)]
12. Banerji MA. The foundation of diabetes self-management: glucose monitoring. *Diabetes Educ* 2007 Apr;33 Suppl 4:87S-90S. [doi: [10.1177/0145721707302059](https://doi.org/10.1177/0145721707302059)] [Medline: [17485416](#)]
13. EDPG. Guidelines for diabetes care. A desktop guide to type 1 (insulin-dependent) diabetes mellitus. European Diabetes Policy Group 1998. *Diabetes Nutr Metab* 1999 Oct;12(5):345-378. [Medline: [10741350](#)]
14. EDPG. A desktop guide to Type 2 diabetes mellitus. European Diabetes Policy Group 1999. *Diabet Med* 1999 Sep;16(9):716-730. [Medline: [10510947](#)]
15. ADA. Consensus statement on self-monitoring of blood glucose. *Diabetes Care* 1987;10(1):95-99. [Medline: [3552518](#)]
16. Chinese Diabetes Society. China glucose monitoring clinical guidelines (2011 edition). *Natl Med J China* 2011;91(10):656-664. [doi: [10.3760/cma.j.issn.1674-5809.2011.01.004](https://doi.org/10.3760/cma.j.issn.1674-5809.2011.01.004)]
17. Xiaosheng B. The Worldwide Largest Number of Smartphone Users Ranking in 2013 [in Chinese]. URL: <http://www.chinapp.com/shidapinpai/62147> [accessed 2015-03-24] [WebCite Cache ID [6XGhRiEFq](#)]

18. Baike B. ICT. URL: http://baike.baidu.com/link?url=r1TmokrN8iAnNs_WoGm-a3s93z4jyylrCf4XNwB1vPUIXscsQ7CN5Vzp4R6cGh0uEGN7jK71zuyfpcx1tRbdk3StxZibcPY0ZpGftw [accessed 2015-03-25] [WebCite Cache ID 6XHVSZ7k8]
19. Zhenzhou L. The status and prospect of ICT in China (in Chinese). *Contemp Commun* 2005;19:76-77.
20. The State Council. "Twelfth Five (2011-2015)" National Strategic Emerging Industry Development Plan. URL: http://www.gov.cn/xxgk/pub/govpublic/mrlm/201207/t20120720_65368.html [WebCite Cache ID 6XHVVftYI]
21. The State Council. The Issuance of Biomedical Development Plan. URL: http://www.gov.cn/zwgk/2013-01/06/content_2305639.htm [accessed 2015-03-25] [WebCite Cache ID 6XHWvjPuq]
22. The State Council. Opinions on Promoting the Development of Health Services Industry. URL: http://www.gov.cn/zwgk/2013-10/14/content_2506399.htm [accessed 2015-03-25] [WebCite Cache ID 6XHVC6C0r]
23. The Continua Alliance. Continua Design Guidelines-Version 2015. URL: <http://www.continuaalliance.org/products/design-guidelines> [accessed 2015-08-28] [WebCite Cache ID 6b7K41NZM]
24. Martínez-Pérez B, de la Torre-Díez Isabel, López-Coronado M, Sainz-De-Abajo B. Comparison of mobile apps for the leading causes of death among different income zones: a review of the literature and app stores. *JMIR Mhealth Uhealth* 2014;2(1):e1 [FREE Full text] [doi: [10.2196/mhealth.2779](https://doi.org/10.2196/mhealth.2779)] [Medline: [25099695](https://pubmed.ncbi.nlm.nih.gov/25099695/)]
25. Bender JL, Yue Rossini Ying Kwan, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. *J Med Internet Res* 2013;15(12):e287 [FREE Full text] [doi: [10.2196/jmir.2661](https://doi.org/10.2196/jmir.2661)] [Medline: [24366061](https://pubmed.ncbi.nlm.nih.gov/24366061/)]
26. West JH, Hall PC, Hanson CL, Barnes MD, Giraud-Carrier C, Barrett J. There's an app for that: content analysis of paid health and fitness apps. *J Med Internet Res* 2012;14(3):e72 [FREE Full text] [doi: [10.2196/jmir.1977](https://doi.org/10.2196/jmir.1977)] [Medline: [22584372](https://pubmed.ncbi.nlm.nih.gov/22584372/)]
27. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. *J Med Internet Res* 2011;13(3):e65 [FREE Full text] [doi: [10.2196/jmir.1874](https://doi.org/10.2196/jmir.1874)] [Medline: [21979293](https://pubmed.ncbi.nlm.nih.gov/21979293/)]
28. Hua F. The effectiveness analysis of type 2 diabetes with community mobile medical (in Chinese). *J Med Theory Pract* 2013;26(10):1387-1388. [doi: [10.3969/j.issn.1001-7585.2013.10.088](https://doi.org/10.3969/j.issn.1001-7585.2013.10.088)]
29. Yan C. Phone alarm or SMS enhance the role of diabetes medication on time (in Chinese). *J Nurs Sci* 2011;26(9):61. [doi: [10.3870/hlxzz.2011.17.061](https://doi.org/10.3870/hlxzz.2011.17.061)]
30. Weimin W, Gang C. Application and prospect of remote monitoring technology on chronic disease management. *Chin Hosp* 2012;6:73-74. [doi: [10.3969/j.issn.1671-0592.2012.06.024](https://doi.org/10.3969/j.issn.1671-0592.2012.06.024)]
31. The State Intellectual Property Office of the P.R.China. The State Intellectual Property Office. URL: <http://www.sipo.gov.cn/> [accessed 2015-03-25] [WebCite Cache ID 6XHdlt0aj]
32. Huawei Store. URL: <http://www.vmall.com/> [accessed 2015-03-25] [WebCite Cache ID 6XHeSE9J1]
33. Xiaomi Store. URL: <http://m.mi.com/v2.html> [accessed 2015-03-25] [WebCite Cache ID 6XHePMRLx]
34. The Database of National Standard. URL: <http://www.sac.gov.cn/> [accessed 2015-08-31] [WebCite Cache ID 6bBf2rbOE]
35. The Standard Databases for Ministry of Health. URL: <http://wsbzw.wsjdx.gov.cn/wsbzw/> [accessed 2015-08-31] [WebCite Cache ID 6bBesBedU]
36. The Databases for Ministry of Industry and Information Technology. URL: <http://www.cape.com.cn/> [accessed 2015-08-30] [WebCite Cache ID 6bAexiOUy]
37. The Standard Databases for China Food and Drug Administration. URL: <http://wsbzw.wsjdx.gov.cn/wsbzw/> [accessed 2015-08-30] [WebCite Cache ID 6bAfIPcON]
38. Ministry of Health. The Basic Dataset of Disease Management. URL: <http://www.bzsoso.com/standard/Soft/WSXYBZ/2012/04/18/182978.html> [accessed 2016-02-26] [WebCite Cache ID 6bAeqES9d]
39. IEEE PHD Standard Group. Health Informatics—Personal Health Device Communication-Part 10417: Device Specialization—Glucose Meter. URL: <http://standards.ieee.org/findstds/standard/11073-10417-2015.html> [accessed 2015-08-30] [WebCite Cache ID 6bAeoNAOA]
40. IEEE PHD Standard Group. Health Informatics—Personal Health Device Communication-Part 10417: Device Specialization—Continuous Glucose Meter. URL: <http://standards.ieee.org/findstds/standard/11073-10425-2014.html> [accessed 2015-03-25] [WebCite Cache ID 6XHWCXIUU]
41. Weiping J. Chinese glucose monitoring clinical guideline (2011 edition) (in Chinese). *Chin J Front Med Sci* 2011;04-72. [doi: [10.3760/cma.j.issn.1674-5809.2011.01.004](https://doi.org/10.3760/cma.j.issn.1674-5809.2011.01.004)]
42. Weiping J. Chinese continuous glucose monitoring clinical guideline (2012 edition) (in Chinese). *Chin J Diabetes* 2012;4(10):582-590. [doi: [10.3760/cma.j.issn.1674-5809.2012.10.003](https://doi.org/10.3760/cma.j.issn.1674-5809.2012.10.003)]
43. Bluetooth. URL: <http://www.bluetooth.com/Pages/Bluetooth-Home.aspx> [accessed 2015-03-25] [WebCite Cache ID 6XHZHKLnx]
44. Google. Google Fit. URL: <http://www.androidcentral.com/app/google-fit> [accessed 2015-08-30] [WebCite Cache ID 6bAej66OC]
45. Apple. Health Kit. URL: <https://developer.apple.com/healthkit/> [accessed 2015-08-30] [WebCite Cache ID 6bAejw1dH]
46. Microsoft. Microsoft Health Vault. URL: <https://www.healthvault.com/cn/zh> [accessed 2015-08-30] [WebCite Cache ID 6bAemVIW]

47. Wiljer D, Urowitz S, Apatu E, DeLenardo C, Eysenbach G, Harth T, Canadian Committee for Patient Accessible Health Records. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. *J Med Internet Res* 2008;10(4):e34 [FREE Full text] [doi: [10.2196/jmir.1061](https://doi.org/10.2196/jmir.1061)] [Medline: [18974036](https://pubmed.ncbi.nlm.nih.gov/18974036/)]
48. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *J Am Med Inform Assoc* 2006;13(2):121-126 [FREE Full text] [doi: [10.1197/jamia.M2025](https://doi.org/10.1197/jamia.M2025)] [Medline: [16357345](https://pubmed.ncbi.nlm.nih.gov/16357345/)]
49. Tecent. URL: <http://weixin.qq.com/> [accessed 2015-08-28] [WebCite Cache ID 6b7Jw2NAAt]
50. Sina. URL: <http://weibo.com/.Archiving> [WebCite Cache ID 6b7JzwTXJ]
51. Chen Y. Take it personally: Accounting for individual difference in designing diabetes management systems. New York: ACM; 2010 Presented at: Proceedings of the 8th ACM Conference on Designing Interactive Systems; 2010; New York p. 252-261 URL: <http://dl.acm.org/citation.cfm?id=1858218>

Abbreviations

BGM: blood glucose monitor
CFDA: China Food and Drug Administration
CGM: continuous glucose monitor
HbA1c: glycosylated hemoglobin
HIS: hospital information system
ICT: information and communication technology
IEEE: Institute of Electrical and Electronics Engineers
MoH: Ministry of Health
PHR: Personal health record
SMBG: self-monitoring of blood glucose

Edited by G Eysenbach; submitted 28.03.15; peer-reviewed by J West, O El-Gayar, Y Wang; comments to author 29.07.15; revised version received 01.09.15; accepted 22.01.16; published 06.04.16.

Please cite as:

Li Y, Tan J, Shi B, Duan X, Zhong D, Li X, Qu J

Information and Communication Technology-Powered Diabetes Self-Management Systems in China: A Study Evaluating the Features and Requirements of Apps and Patents

JMIR Diabetes 2016;1(1):e2

URL: <http://diabetes.jmir.org/2016/1/e2/>

doi: [10.2196/diabetes.4475](https://doi.org/10.2196/diabetes.4475)

PMID: [30291083](https://pubmed.ncbi.nlm.nih.gov/30291083/)

©Ying Li, Jin Tan, Bozhi Shi, Xiaolian Duan, Daidi Zhong, Xiaoling Li, Jianning Qu. Originally published in *JMIR Diabetes* (<http://diabetes.jmir.org>), 06.04.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Diabetes*, is properly cited. The complete bibliographic information, a link to the original publication on <http://diabetes.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Evaluating the Accuracy of Google Translate for Diabetes Education Material

Xuewei Chen¹, MEd; Sandra Acosta², PhD; Adam Etheridge Barry³, PhD

¹Transdisciplinary Center for Health Equity Research, Department of Health and Kinesiology, Texas A&M University, College Station, TX, United States

²Texas A&M University, Department of Educational Psychology, Texas A&M University, College Station, TX, United States

³Texas A&M University, Department of Health and Kinesiology, Texas A&M University, College Station, TX, United States

Corresponding Author:

Xuewei Chen, MEd

Transdisciplinary Center for Health Equity Research

Department of Health and Kinesiology

Texas A&M University

Blocker 311B

4243 TAMU

College Station, TX, 77843

United States

Phone: 1 979 676 0758

Fax: 1 (979) 847 8987

Email: xuweichen@tamu.edu

Abstract

Background: Approximately 21% of the US population speaks a language other than English at home; many of these individuals cannot effectively communicate in English. Hispanic and Chinese Americans, in particular, are the two largest minority groups having low health literacy in the United States. Fortunately, machine-generated translations represent a novel tool that non-English speakers can use to receive and relay health education information when human interpreters are not available.

Objective: The purpose of this study was to evaluate the accuracy of the Google Translate website when translating health information from English to Spanish and English to Chinese.

Methods: The pamphlet, "You are the heart of your family...take care of it," is a health education sheet for diabetes patients that outlines six tips for behavior change. Two professional translators translated the original English sentences into Spanish and Chinese. We recruited 6 certified translators (3 Spanish and 3 Chinese) to conduct blinded evaluations of the following versions: (1) sentences translated by Google Translate, and (2) sentences translated by a professional human translator. Evaluators rated the sentences on four scales: fluency, adequacy, meaning, and severity. We performed descriptive analysis to examine differences between these two versions.

Results: Cronbach's alpha values exhibited high degrees of agreement on the rating outcome of both evaluator groups: .919 for the Spanish evaluators and .972 for the Chinese evaluators. The readability of the sentences in this study ranged from 2.8 to 9.0 (mean 5.4, SD 2.7). The correlation coefficients between the grade level and translation accuracy for all sentences translated by Google were negative (eg, $r_{\text{Meaning}} = -.660$), which indicates that Google provided accurate translation for simple sentences. However, the likelihood of incorrect translation increased when the original English sentences required higher grade levels to comprehend. The Chinese human translator provided more accurate translation compared to Google. The Spanish human translator, on the other hand, did not provide a significantly better translation compared to Google.

Conclusion: Google produced a more accurate translation from English to Spanish than English to Chinese. Some sentences translated by Google from English to Chinese exhibit the potential to result in delayed patient care. We recommend continuous training and credential practice standards for professional medical translators to enhance patient safety as well as providing health education information in multiple languages.

(*JMIR Diabetes* 2016;1(1):e3) doi:[10.2196/diabetes.5848](https://doi.org/10.2196/diabetes.5848)

KEYWORDS

health literacy; health education; health communication; language translation; diabetes; machine translation; human interpreter; translator

Introduction

Health promotion and education material from health organizations, as well as mass media, are primarily written and delivered in English. While public health professionals are working to produce more health content and material in other languages, current availability remains limited [1,2]. For patients and caregivers with limited English proficiency (LEP), this lack of health information in their native language can be especially burdensome and represents an important public health issue.

LEP individuals, defined as any person age 5 and older who speaks English “less than very well” [3], represent a vulnerable population that experiences significant health disparities in the United States [4]. Compared to the native English-speaking population, LEP individuals are less likely to receive and understand health information or correctly interpret health education messages [5].

As a result of their lack of comprehension and/or misinterpretation, LEP individuals (1) spend extra time and money seeking and using health care services, (2) have unsatisfactory experiences with health care providers, (3) make inappropriate health decisions, (4) have limited access and use of preventive health care services, (5) are more challenging to recruit into health education programs, (6) take incorrect dosages of medication, and (7) have worse health status [6-11]. These issues become increasingly important to address as the LEP population in the United States continues to steadily grow. According to the US Census Bureau, approximately 21% of the US population (60.6 million) speaks a language other than English at home [12]. Moreover, among foreign-born US adults, nearly three out of four speak limited English or do not speak English at all [13].

Machine-generated translations represent a novel tool that non-English speakers can use to receive and relay health education information when human interpreters are not available. With the proliferation of online technology, 87% of US adults had access to the Internet in 2014, compared to 43% in 2000 [14]. Moreover, the Internet is becoming increasingly prevalent among minority populations [15]. Perry and Mittelmark [16] contend that digital translation tools “offer substantial time and cost saving...can thus be used not only to immediately collect information when the content is not translated, but also to immediately deliver information generated in one language to speakers of other languages” (p. 199). However, miscommunication through translation is possible given that words often have different meanings depending on the context in which they are used [16].

Khanna et al [17] compared English-to-Spanish translation accuracy between Google and human translators for patient education texts, reporting that Google Translate made more errors than human translators and people preferred the human translation for complex sentences. Similarly, Sharif and Tse [18] reported an overall 50% error rate for medicine labels

translated from English to Spanish by computer programs. Google Translate has also exhibited a high rate of translation errors when translating content on state and national public health websites from English to Chinese [19]. However, to date, we are unaware of any studies evaluating the outputs of a machine translation tool when translating from English to multiple languages drawn from health education material on diabetes. Therefore, it is critical to identify and evaluate available translation tools for helping LEP speakers of different languages understand English health education material.

The purpose of this pilot study was to evaluate the feasibility and accuracy of the Google Translate website as a tool to help LEP persons understand chronic condition management and prevention strategies. Specifically, Google Translate was used for translating a diabetes patient education pamphlet, distributed by the National Diabetes Education program, from English to Spanish and English to Chinese (Mandarin). We chose to focus on Spanish and Chinese for several reasons. First, Spanish and Chinese speakers are more likely to have limited English proficiency. In the United States, approximately 47% of the foreign-born population from Mexico speaks English “not well” or “not at all,” and 32% of the foreign-born population from China speaks English “not well” or “not at all” [13]. Second, among the LEP population, Chinese (68%) respondents exhibit low health literacy, followed by Latinos (45%) [20]. Third, the prevalence of diabetes is rapidly increasing among Hispanic and Chinese Americans [21]. The following research questions guided this investigation:

1. What is the accuracy of Google translations of written sentences from English to Spanish, when compared to professional human translators?
2. What is the accuracy of Google translations of written sentences from English to Chinese, when compared to professional human translators?
3. Can Google Translate be a safe and accurate alternative to human interpreters for providing translation services on health education materials to LEP patients?

Methods**Materials to be Translated**

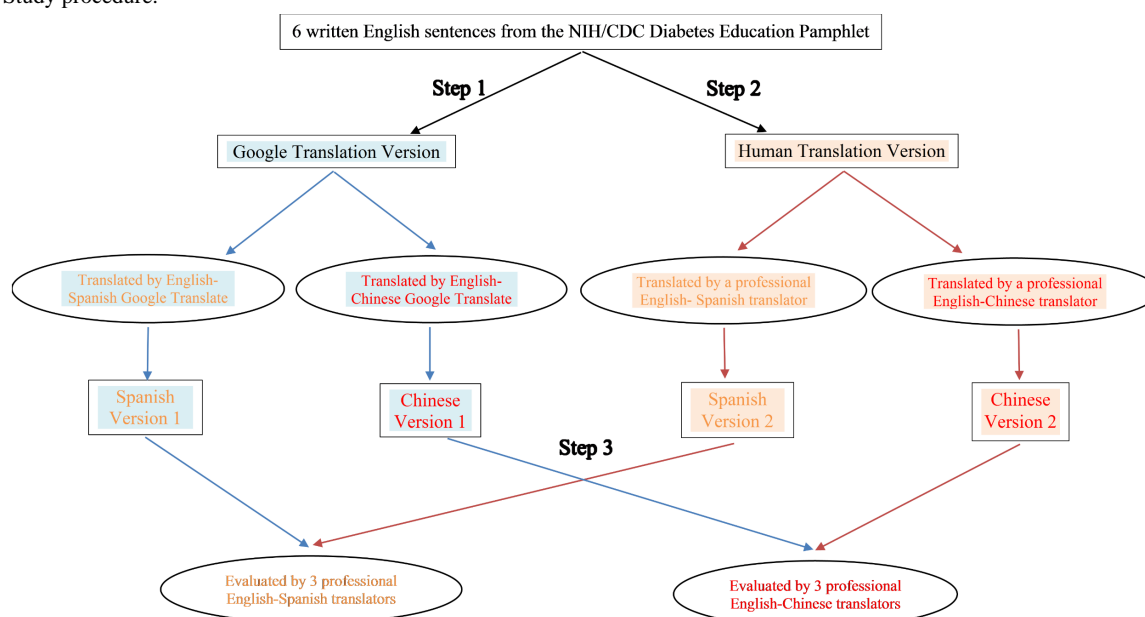
We chose a freely accessible diabetes patient education pamphlet as a heuristic example for evaluating the accuracy of machine translation devices. The pamphlet, “You are the heart of your family...take care of it,” is published by the National Institutes of Health and the Centers for Disease Control and Prevention and distributed by the National Diabetes Education Program. This pamphlet includes six written sentences as behavior change suggestions for managing diabetes and three recommended questions for patients to ask their clinicians. This paper examines the accuracy of Google Translate when translating the six written diabetes prevention and management strategies to determine the differences between machine and human translators, which

could be used to direct further research. This study was approved by the Texas A&M University Institutional Review Board.

Procedures

Following are the overall procedures (see [Figure 1](#)) used throughout this investigation.

Figure 1. Study procedure.



Step 1. Google Translate

We used Google Translate, a free language translation website that instantly translates text and Web pages, to translate the six sentences from English into both Spanish and Chinese.

Step 2. Human Translate

Two professional medical translators translated the original English pamphlet into Spanish and Chinese, respectively. Both were American Translators Association (ATA) certified translators (one certified in English to Spanish and the other in English to Chinese). The ATA website lists all the certified translators' contact information. We approached both translators as regular customers seeking translation services. We did not inform them that their translation product would be evaluated. We sent the original English materials to them by email; they returned the translated sentences in Microsoft Word to us by email. All human translation services were paid for based on quotes provided by the individual translators.

Step 3. Evaluation

After having the materials translated, we separately recruited 6 ATA-certified translators to evaluate the two translation versions (one by Google Translate and the other one by professional human translators). The two translators who provided the human translation versions did not serve as evaluators, nor were they aware we would have evaluators evaluate their translations. Evaluators were also approached via email. We randomly sent invitation emails to 12 English-Spanish translators and 12 English-Chinese translators. We emailed the survey package to the first 6 translators (3 Spanish and 3 Chinese respectively) who accepted our study invitation. They became the evaluators for this study. Each evaluator received US \$15 after submitting the evaluation survey package via email.

Survey Package

To minimize bias, we did not inform the evaluators which version was created by a human or a machine; instead, we marked the products as version 1 (sentences translated by Google) and version 2 (sentences translated by a human). The survey package contained three separate Microsoft Word documents: (1) an evaluation rubric, (2) translation version 1, and (3) translation version 2. Both versions consisted of six written sentences with the original English sentences listed first, followed by the translated sentences (Spanish or Chinese). We asked the evaluators to score each of the translated sentences based on the included evaluation rubric.

Evaluation Rubric

Our evaluation rubric, which was adapted from Khana et al [17], asked evaluators to rate the translation sentences based on Fluency, Adequacy, Meaning, and Severity on a 5-point scale (1 indicating low accuracy and 5 indicating high accuracy). The Fluency and Adequacy evaluations are standard domains for assessing machine translation accuracy [22]. The Fluency domain evaluated readability, grammar, and understandability. The Adequacy domain evaluated how much of the original information had been preserved. The Meaning domain assessed whether the translation product had the same meaning as the original sentence. If a translation product added extra information, it could still receive a high Adequacy score as long as it included all the original information. The Meaning score, however, could identify misleading added information [23]. The evaluators also rated the Severity domain, which provided insight into the degree of negative impact on the patient's health outcome. The detailed evaluation rubric (see [Table 1](#)) defined the different categories for each domain.

Table 1. Evaluation rubric.

Fluency	Adequacy	Meaning	Severity
1 No fluency; no appreciable grammar, not understandable	0% of information conveyed from the original	Totally different meaning from the original	Dangerous to patient
2 Marginal fluency; several grammatical errors	25% of information conveyed from the original	Misleading information added/omitted compared to the original	Impairs care in some way
3 Good fluency; several grammatical errors, understandable	50% of information conveyed from the original	Partially the same meaning as the original	Delays necessary care
4 Excellent fluency; few grammatical errors	75% of information conveyed from the original	Almost the same meaning as the original	Unclear effect on patient care
5 Perfect fluency; like reading a newspaper	100% of information conveyed from the original	Same meaning as the original	No effect on patient care

Data Analysis

We used Cronbach's alpha to assess the degree of agreement among the evaluators. We calculated two sets of means to represent the scores in each of the four domains (ie, Fluency, Adequacy, Meaning, and Severity) from the Chinese and Spanish evaluator groups. We performed descriptive analysis to capture the trend of change from sentence to sentence. Pearson correlation coefficients were also reported to examine the relationship between translation accuracy and the readability of the original English sentences. The readability statistics were generated using Microsoft Word's Flesch-Kincaid Grade Level, which assesses the degree of difficulty for readers to understand a certain sentence or paragraph [24].

To examine the correlational patterns in the data, we considered using multivariate analysis of variance (MANOVA) for investigating whether there was a statistically significant difference between Google and the professional translators with regard to the translation accuracy. MANOVA allows for the comparison of two groups on these four translation accuracy domains simultaneously [25]. However, *P* values are closely dependent on sample size [26]. Thus, such significance testing is not appropriate in this study due to our small sample size

Table 2. Flesch-Kincaid grade level.

Original sentences	Flesch-Kincaid grade level
S1. Eat more fruits, vegetables, beans, and whole grains.	3.7
S2. Eat tasty foods that have less salt, saturated fat, and trans fat.	4.8
S3. Get at least 30 minutes of physical activity on most days or every day. Physical activity helps you keep a healthy weight.	8.5
S4. Stop smoking.	2.8
S5. Take medicines the way your doctor tells you.	3.7
S6. Ask your doctor about taking medicine to protect your heart, such as aspirin or a statin.	9.0
Mean (SD)	5.4 (2.7)

The higher grade level indicates that the text is more difficult for readers to understand. As shown in Table 3, the correlation coefficients between the grade level and translation accuracy for all sentences translated by Google (both Spanish and Chinese) were negative. None of the correlation coefficients

(*N*=6) and the violation of MANOVA assumptions (eg, normality and homogeneity of variance). Therefore, instead of conducting MANOVA, we presented two sets of graphs to visually compare the translation accuracy between Google and human.

Results

Inter-rater Reliability

Cronbach's alpha was used to assess the rating reliability across each evaluator. Cronbach's alpha values exhibited high degrees of agreement on the rating outcome of both evaluator groups: .919 for the Spanish evaluators and .972 for the Chinese evaluators.

Grade Level and Correlations With Accuracy Scores

Table 2 shows the Flesch-Kincaid Grade Level for all six original English sentences. The Flesch-Kincaid readability test rates text on a US school grade level [24]. The readability of the sentences in this study ranged from 2.8 to 9.0 (mean 5.4, SD 2.7). Shorter sentences with simpler vocabulary received lower scores (eg, grade level=2.9 for S4), and longer sentences containing more medical terms received higher scores (eg, grade level=9.0 for S6).

was statistically significant at alpha <.05 level due to the small sample size in our study (*N*=6). However, these negative associations were relatively strong, especially among the Chinese Google group (eg, $r_{\text{Meaning}}=-.660$). For the sentences translated by the professional human translators, there was only

one negative correlation between grade level and translation accuracy scores ($r_{\text{Fluency}}=-.447$). The correlation coefficients between the grade level and translation accuracy scores show that Google provides more accurate translation for easier

sentences but produces more translation errors for more complex sentences. However, the accuracy scores of translated sentences provided by human translators had no strong negative associations with the readability level of the sentences.

Table 3. Correlations between grade level and translation accuracy.

Domains	Flesch-Kincaid grade level			
	Spanish		Chinese	
	Google	Human	Google	Human
Fluency	-.374	-.447	-.373	.679
Adequacy	-.162	.120	-.371	.481
Meaning	-.259	.207	-.660	.481
Severity	-.097	.341	-.469	^a

^aCorrelation coefficient cannot be computed because all sentences translated by the Chinese human translator had a constant severity score (Severity=5).

Spanish Translation: Google Versus Human

As shown in Table 4, in the Fluency domain, all sentences translated by Google had at least good fluency (Fluency \geq 3). All sentences translated by the Spanish human translator had excellent or perfect fluency.

In the Adequacy domain, most sentences from both versions conveyed more than 75% of the original information. One sentence translated by the Spanish human translator (S5) conveyed 50% of the original information (Adequacy=3).

In the Meaning domain, similarly, all sentences from both versions had almost the same meaning as the original information. However, S5 translated by the Spanish human translators had partially the same meaning as the original sentence (Meaning=3).

In the Severity domain, all evaluators agreed that S5 translated by Google had an unclear effect on patient care (Severity=4). That same sentence translated by the Spanish human translator delayed necessary patient care (Severity=3).

Table 4. Spanish Google versus human.

Original sentences	Google				Human			
	Fluency	Adequacy	Meaning	Severity	Fluency	Adequacy	Meaning	Severity
S1. Eat more fruits, vegetables, beans, and whole grains.	4.67	5	5	5	4.33	5	4.67	5
S2. Eat tasty foods that have less salt, saturated fat, and trans fat.	3	4.67	4.33	5	4.67	4.67	4.67	4.67
S3. Get at least 30 minutes of physical activity on most days or every day. Physical activity helps you keep a healthy weight.	3	4.33	4	4.67	4.67	5	4.67	5
S4. Stop smoking.	5	5	5	5	5	5	5	5
S5. Take medicines the way your doctor tells you.	4.33	4.33	4.33	4	4.67	3	3	3
S6. Ask your doctor about taking medicine to protect your heart, such as aspirin or a statin.	4.67	5	5	4.67	4.33	4.33	4.67	5

Chinese Translation: Google Versus Human

As shown in Table 5, in the Fluency domain, S2, S3, and S5 translated by Google had marginal or no fluency (Fluency \leq 2). Every evaluator agreed that S5 was not understandable. All sentences translated by the Chinese human translator had excellent or perfect fluency.

In the Adequacy domain, S5 translated by Google conveyed less than 50% of the original information (Adequacy<3). All

sentences translated by the Chinese human translator conveyed almost 100% of the original information.

In the Meaning domain, S3 and S5 translated by Google had less than partially the same meaning as the original information (Meaning<3). All sentences translated by the Chinese human translator had the same or almost the same meaning as the original ones.

In the Severity domain, S5 and S6 translated by Google delayed necessary care for patients (Severity<3). All sentences translated by the Chinese human translator had no effect on patient care (Severity = 5).

Table 5. Chinese Google versus human.

Original sentences	Google				Human			
	Fluency	Adequacy	Meaning	Severity	Fluency	Adequacy	Meaning	Severity
S1. Eat more fruits, vegetables, beans, and whole grains.	4.67	5	4.67	5	5	5	5	5
S2. Eat tasty foods that have less salt, saturated fat, and trans fat.	2	4.33	3.67	4.67	4.67	5	5	5
S3. Get at least 30 minutes of physical activity on most days or every day. Physical activity helps you keep a healthy weight.	1.67	3.67	2.67	4	5	5	5	5
S4. Stop smoking.	5	5	5	5	4.67	4.67	4.67	5
S5. Take medicines the way your doctor tells you.	1	2.67	2.67	2.33	4.67	5	5	5
S6. Ask your doctor about taking medicine to protect your heart, such as aspirin or a statin.	3	3.67	3	2.33	5	5	5	5

Visually Comparing Google and Human Versions

As shown in [Figures 2 and 3](#), to better compare and capture the trends among sentences with regard to the accuracy scores on four domains, we ranked the sentences according to their grade levels—presenting the easiest sentence (S4) first and the most difficult sentence (S6) last. As shown in [Figure 2](#), when sentences were translated from English to Spanish, S2 and S3 (more difficult sentences) had a considerable difference between Google and human in the Fluency domain, where the human translator did much better than Google. For the relatively easy sentences (S4 and S1), there was not much difference between Google and human in any of the four domains. Interestingly, there was not much difference for the most difficult sentence (S6) either. We also noticed some obvious gaps for S5 (medium

difficult sentence) in the Adequacy, Meaning, and Severity domains, where Google received a higher translation accuracy (English to Spanish) than the human translator did. As shown in [Figure 3](#), when sentences were translated from English to Chinese, S5, S2, S3, and S6 (more difficult sentences) had a considerable difference between Google and human in all four domains, where the human did much better than Google (except S2 in the Severity domain). Similar to what we found in the Spanish set, there was not much difference between Google and human in all domains for the easier sentences (S4 and S1). When comparing between [Figures 2 and 3](#), results showed that the general distance between Google and human for Chinese is larger than Spanish, indicating that Google provided higher accuracy translation service in Spanish than in Chinese.

Figure 2. Spanish Google versus human.

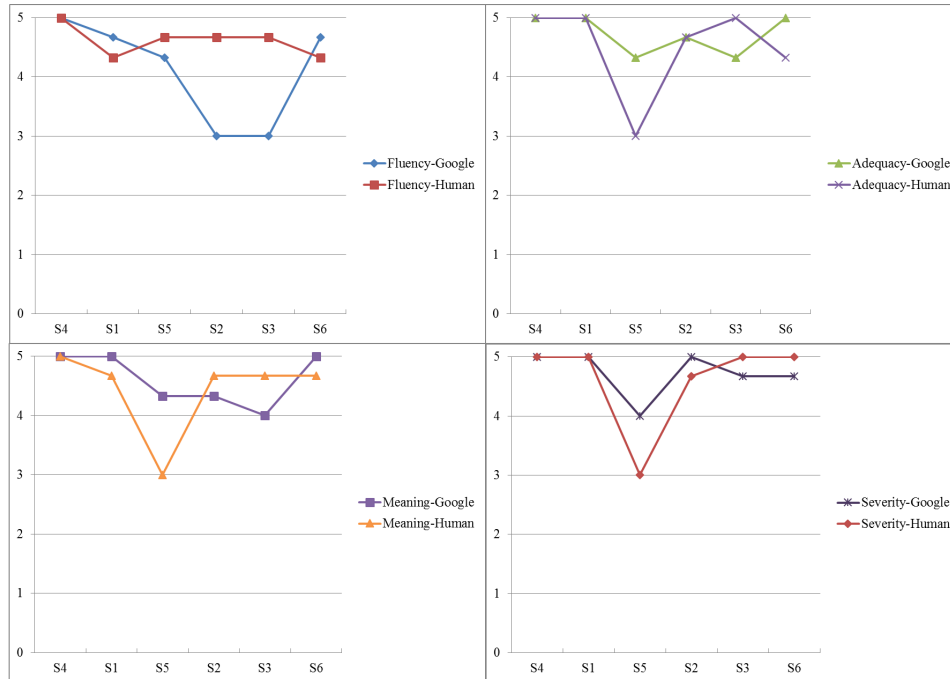
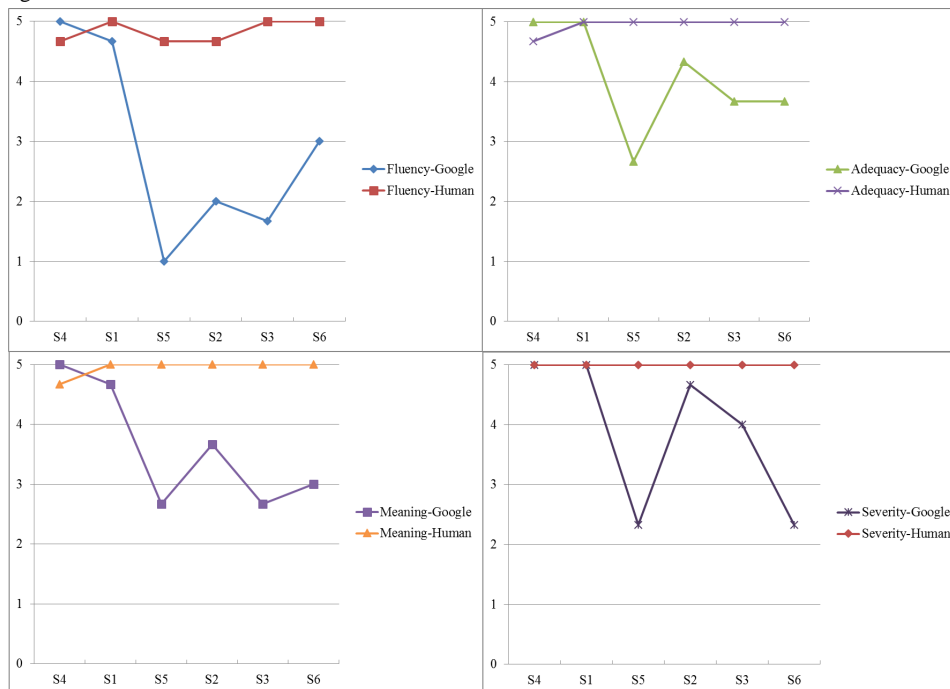


Figure 3. Chinese Google versus human.



Discussion

Principal Findings

This pilot study evaluated the accuracy of Google Translate when translating diabetes patient education materials from English to Spanish and English to Chinese. We found that Google provided accurate translation for simple sentences, but the likelihood of incorrect translation increased when the original English sentences required higher grade levels to comprehend. For example, the most simple sentence in our study (“Stop smoking”) translated by Google received full scores on every domain for both languages, while Google received

lower scores on more difficult sentences (eg, S5 and S6) for both languages. The Chinese human translator provided much more accurate translation than Google did. The Spanish human translator, on the other hand, did not provide a significantly better translation compared to Google. Additionally, we identified some sentences translated by Google from English to Chinese that might lead to delayed patient care. Similarly, one sentence translated by the professional human translator from English to Spanish could also have a negative impact on patients. The results demonstrate that Google is capable of producing a more accurate translation from English to Spanish than English to Chinese.

Google provided more accurate translations for sentences with lower readability levels but made more translation errors on sentences with higher readability levels, especially when translating sentences from English to Chinese. Although we did not find any statistically significant correlation between readability and translation accuracy due to our small sample size, our findings seem to be consistent with previous investigations that document a significant negative correlation between sentence length and translation accuracy ($r=-.4393$, $P<.05$), indicating that the machine was less likely to provide correct translation for longer sentences [27].

Google yielded high error rates when translating English sentences to Chinese. We identified several problematic sentences. S2 and S3 translated by Google from English to Chinese had marginal fluency with several grammatical errors, but the evaluators were able to make enough sense of them to get a meaning close to the original sentences. Thus, these two sentences did not have much negative impact on patient care. However, S5 (“Take medicines the way your doctor tells you”) translated by Google from English to Chinese had no fluency and was not understandable. After translation, this sentence in Chinese did not make sense to the evaluators. Therefore, this sentence might cause delayed patient care. Likewise, S6 (“Ask your doctor about taking medicine to protect your heart, such as aspirin or a statin”) translated by Google from English to Chinese added misleading information into the original sentence by translating it as “Ask your doctor about taking Chinese medicine to protect your heart, such as aspirin or a statin.”

Compared to Chinese, Google provided noticeably higher accuracy when translating sentences from English to Spanish. All the Spanish sentences conveyed more than 75% of the original information and had almost the same meaning as the original sentences. Moreover, none of them had a severe impact on patient care (Severity \geq 4.67). Consistent with our findings, Zeng-Treitler et al [27] also found that Spanish machine translation had higher accuracy than other languages: Spanish had 33.8% correctness compared to the correctness of Chinese, Russian, and Korean, which ranged from 7.98% to 11.74%.

Textbox 1. Translation standard with 10 key components.

1. Develop the English text and/or test the translation with members of the target LOTE (a language other than English)-speaking community.
2. Undertake a cultural and linguistic assessment of the English text in preparation for its translation.
3. Undertake a subject matter expert assessment of the English text as appropriate.
4. Organize for the English text to be translated by a professional translator.
5. Undertake a cultural and linguistic assessment of the translation.
6. Organize for the translation to be proofread by a professional translator.
7. Include the title of the text in English on the translation.
8. Include the name of the target language in English, on both the English text and the translation.
9. Distribute the translation in bilingual format—English and LOTE.
10. Date, monitor, evaluate, and update the English text and the translation as part of an ongoing review program.

In addition to ensuring human translation accuracy, improvements to machine translation tools are also necessary prior to use by patients and health care providers. Health educators should make efforts to achieve higher translation

Zeng-Treitler et al [27] contend that “one possible explanation for this may well lie in the fact that English and Spanish are more similar (eg, word order, inflections) than English and Chinese, Korean or Russian” (p. 76).

The Chinese human translator provided much more accurate translation than Google; however, the Spanish human translator did not provide a significantly better translation than Google. In contrast to our findings, Khanna et al [17] reported that Google made more errors than human translators when translating patient education materials from English to Spanish. Zeng-Treitler et al [27] concluded that Babelfish was not a good machine translation tool because of its high percentage of inaccuracy.

We identified one problematic sentence (S5 “Take medicines the way your doctor tells you”); the translation by the Spanish human translator might cause delayed patient care. This sentence was also problematic when translated by Google from English to Chinese. It conveyed half of the original information and partially the same meaning as the original sentence. The Spanish human translator twisted the meaning of the original English sentence by creating a Spanish sentence saying “Tome las medicinas recetadas por su médico,” meaning “Take the medicine prescribed by your doctor.” Such incorrect translation provided by the Spanish human translator might lead to delayed necessary patient care.

We also wish to highlight that in some cases professional human translators might also make severe errors that negatively impact patients’ health compared to machine translation tools. Flores et al [28] contend that the most common types of mistake by human interpreters, which could potentially cause medical accidents, include omission, false fluency, substitution, editorialization, and addition. For this reason, we recommend continuous training and credential practice standards for professional medical translators to enhance patient safety. For example, Michael et al [29] developed a translation standard to guide the language-translating process for health education information (see [Textbox 1](#)) with 10 key components (p. 550).

accuracy for machine tools and ultimately make sure health education information is not misinterpreted and necessary care not delayed. Mismatches between the vocabulary bank in machine translation systems and the terminologies used in the

original language texts are common sources of machine translation errors [30]. Developing a universal code system for machine translation can improve language translation accuracy [31]. Therefore, we call for collaborations between computer science engineers and public health/health education professionals to work on this language translation technique, which could assist LEP populations better understand health information.

Furthermore, health education information should be written in multiple languages other than English and Spanish. In one study, Becker [1] examined 125 websites that provided health information in the United States and reported that only 10% of the state sites provided Spanish versions. Moreover, these Spanish webpages contained many English texts such as Web link buttons labeled in English. Most health institutions do not provide information in multiple languages besides English on their websites, but Internet users prefer searching for health information using local languages instead of English. Immigrants in particular prefer seeking and reading health information in their native languages rather than the languages of the adopted country [32].

Limitations

Our study has three limitations that should be noted. First, we recruited ATA-certified translators as evaluators who, because of their professional training, had more credibility for scientifically evaluating translation accuracy than non-professional bilinguals such as graduate students. Translators also have different translation styles and knowledge of second language audiences. The selection of certified translators might cause measurement bias because these professional translators are different from general LEP patients.

For instance, compared to LEP patients, certified translators are bilingual, well-educated, and have higher literacy levels. Thus, sentences that are understandable to them might not be understandable to LEP patients. Future research might recruit LEP participants to evaluate these translation products, and researchers might conduct cognitive interviews while participants read these sentences. Second, our study mainly focused on describing the translated products from a technical perspective instead of assessing message consumers' experience from a user perspective. Testing LEP diabetes patients' knowledge and behavior change after using Google Translate to process health education messages is another direction for future study. Finally, our study sample size was small. We evaluated six original English sentences and recruited 6 evaluators, which had less power for generalizability. Researchers should include a large sample of original sentences and evaluators for future study.

Conclusions

Notwithstanding these limitations, this investigation provides important contributions to the ever-growing literature base examining the effectiveness of machine translation tools. In particular, our findings highlight that as sentences become more complex in health information and require higher levels of reading ability, the likelihood of machine translation tools making errors increases. As shown in the paper, these errors have the potential to negatively impact patient health behaviors. Given that medical or health advice is not always delivered in short, easy-to-understand sentences, such as those at a 2.8 grade reading level (eg, "Stop smoking"), it is imperative that future investigations continue to examine the real-world application of machine translation tools and their associated impact on patient and population health.

Acknowledgments

This study was supported by the College of Education and Human Development (CEHD) at Texas A&M University under the CEHD Graduate Research Grant Award. We want to give sincere thanks to Dr. Bruce Thompson for his assistance in the data analysis process. We also thank the ATA translators who participated in this study.

Conflicts of Interest

None declared.

References

1. Becker S. A Study of Web Usability for Older Adults Seeking Online Health Resources. *ACM Transactions on Computer-Human Interaction (TOCHI)*. 2004. (4) p. 387-406 URL: <http://www2.psych.ubc.ca/~pgraf/Psy583Readings/Becker%202004.pdf> [WebCite Cache ID 6iaPzT5uH]
2. Spatz M. *The Medical Library Association guide to providing consumer and patient health information*. Lanham: Rowman & Littlefield; 2014.
3. Pandya C, McHugh M, Batalova J. Limited English Proficient Individuals in the United States: Number, Share, Growth, and Linguistic Diversity. LEP Data Brief. Migration Policy Institute. 2011. URL: <http://www.migrationpolicy.org/research/limited-english-proficient-individuals-united-states-number-share-growth-and-linguistic> [WebCite Cache ID 6iaPzT5uH]
4. Kutner M. *The health literacy of America's adults: results from the 2003 National Assessment of Adult Literacy*. Washington, D. C. Washington, DC: US Dept. of Education, National Center for Education Statistics; 2006.
5. Lopez-Quintero C, Berry EM, Neumark Y. Limited English proficiency is a barrier to receipt of advice about physical activity and diet among Hispanics with chronic diseases in the United States. *J Am Diet Assoc* 2009 Oct;109(10):1769-1774. [doi: [10.1016/j.jada.2009.07.003](https://doi.org/10.1016/j.jada.2009.07.003)] [Medline: [19782177](https://pubmed.ncbi.nlm.nih.gov/19782177/)]

6. Jacobs EA, Karavolos K, Rathouz PJ, Ferris TG, Powell LH. Limited English proficiency and breast and cervical cancer screening in a multiethnic population. *Am J Public Health* 2005 Aug;95(8):1410-1416. [doi: [10.2105/AJPH.2004.041418](https://doi.org/10.2105/AJPH.2004.041418)] [Medline: [16043670](https://pubmed.ncbi.nlm.nih.gov/16043670/)]
7. Nair EL, Cienkowski KM. The impact of health literacy on patient understanding of counseling and education materials. *Int J Audiol* 2010 Feb;49(2):71-75. [doi: [10.3109/14992020903280161](https://doi.org/10.3109/14992020903280161)] [Medline: [20151879](https://pubmed.ncbi.nlm.nih.gov/20151879/)]
8. Ponce NA, Hays RD, Cunningham WE. Linguistic disparities in health care access and health status among older adults. *J Gen Intern Med* 2006 Jul;21(7):786-791 [FREE Full text] [doi: [10.1111/j.1525-1497.2006.00491.x](https://doi.org/10.1111/j.1525-1497.2006.00491.x)] [Medline: [16808783](https://pubmed.ncbi.nlm.nih.gov/16808783/)]
9. Thomson MD, Hoffman-Goetz L. Challenges of recruiting ESL immigrants into cancer education studies: reflections from practice notes. *J Cancer Educ* 2011 Mar;26(1):170-174. [doi: [10.1007/s13187-010-0132-8](https://doi.org/10.1007/s13187-010-0132-8)] [Medline: [20526867](https://pubmed.ncbi.nlm.nih.gov/20526867/)]
10. Toci E, Burazeri G, Jerliu N, Sørensen K, Ramadani N, Hysa B, et al. Health literacy, self-perceived health and self-reported chronic morbidity among older people in Kosovo. *Health Promot Int* 2015 Sep;30(3):667-674. [doi: [10.1093/heapro/dau009](https://doi.org/10.1093/heapro/dau009)] [Medline: [24604915](https://pubmed.ncbi.nlm.nih.gov/24604915/)]
11. Wilson E, Chen AHM, Grumbach K, Wang F, Fernandez A. Effects of limited English proficiency and physician language on health care comprehension. *J Gen Intern Med* 2005 Sep;20(9):800-806 [FREE Full text] [doi: [10.1111/j.1525-1497.2005.0174.x](https://doi.org/10.1111/j.1525-1497.2005.0174.x)] [Medline: [16117746](https://pubmed.ncbi.nlm.nih.gov/16117746/)]
12. Ryan C. US Census Bureau. 2013 Aug. Language use in the United States: 2011 URL: <https://www.census.gov/prod/2013pubs/acs-22.pdf> [accessed 2016-06-23] [WebCite Cache ID 6iULP88nl]
13. Gambino C, Acosta Y, Grieco E. US Census Bureau. 2014 Jun. English-speaking ability of the foreign-born population in the United States:2012 URL: <http://www.census.gov/content/dam/Census/library/publications/2014/acs/acs-26.pdf> [accessed 2016-06-23] [WebCite Cache ID 6iULMkK00]
14. Fox S, Rainie L. Pew Research Center. Washington, DC; 2014. The Web at 25 in the US The overall verdict: The internet has been a plus for society and an especially good thing for individual users URL: http://www.pewinternet.org/files/2014/02/PIP_25th-anniversary-of-the-Web_0227141.pdf [accessed 2016-06-23] [WebCite Cache ID 6iULJxocJ]
15. Lopez M, Gonzalez-Barrera A, Patten E. Closing the digital divide: Latinos and technology adoption. Washington, DC: Pew Research Center, Pew Hispanic Center; 2013. URL: http://www.pewhispanic.org/files/2013/03/Latinos_Social_Media_and_Mobile_Tech_03-2013_final.pdf [accessed 2016-06-23] [WebCite Cache ID 6iULV02wd]
16. Perry MW, Mittelmark MB. The use of emerging technology to build health promotion capacity in regions with diversity in language and culture. *Promot Educ* 2006;13(3):197-202. [Medline: [17294710](https://pubmed.ncbi.nlm.nih.gov/17294710/)]
17. Khanna RR, Karliner LS, Eck M, Vittinghoff E, Koenig CJ, Fang MC. Performance of an online translation tool when applied to patient educational material. *J Hosp Med* 2011 Nov;6(9):519-525. [doi: [10.1002/jhm.898](https://doi.org/10.1002/jhm.898)] [Medline: [22034196](https://pubmed.ncbi.nlm.nih.gov/22034196/)]
18. Sharif I, Tse J. Accuracy of computer-generated, spanish-language medicine labels. *Pediatrics* 2010 May;125(5):960-965 [FREE Full text] [doi: [10.1542/peds.2009-2530](https://doi.org/10.1542/peds.2009-2530)] [Medline: [20368321](https://pubmed.ncbi.nlm.nih.gov/20368321/)]
19. Turner AM, Dew KN, Desai L, Martin N, Kirchoff K. Machine Translation of Public Health Materials From English to Chinese: A Feasibility Study. *JMIR Public Health Surveill* 2015;1(2):e17 [FREE Full text] [doi: [10.2196/publichealth.4779](https://doi.org/10.2196/publichealth.4779)] [Medline: [27227135](https://pubmed.ncbi.nlm.nih.gov/27227135/)]
20. Sentell T, Braun KL. Low health literacy, limited English proficiency, and health status in Asians, Latinos, and other racial/ethnic groups in California. *J Health Commun* 2012;17 Suppl 3:82-99 [FREE Full text] [doi: [10.1080/10810730.2012.712621](https://doi.org/10.1080/10810730.2012.712621)] [Medline: [23030563](https://pubmed.ncbi.nlm.nih.gov/23030563/)]
21. Kandula NR, Diez-Roux AV, Chan C, Daviglius ML, Jackson SA, Ni H, et al. Association of acculturation levels and prevalence of diabetes in the multi-ethnic study of atherosclerosis (MESA). *Diabetes Care* 2008 Aug;31(8):1621-1628 [FREE Full text] [doi: [10.2337/dc07-2182](https://doi.org/10.2337/dc07-2182)] [Medline: [18458142](https://pubmed.ncbi.nlm.nih.gov/18458142/)]
22. White J, O'Connell T, O'Mara F. The ARPA MT evaluation methodologies: evolution, lessons, and future approaches. 1994. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.137.1288&rep=rep1&type=pdf> [WebCite Cache ID 6iaQ1cdZI]
23. Eck M, Hori C. Overview of the IWSLT 2005 evaluation campaign. 2005. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.331.7046&rep=rep1&type=pdf> [WebCite Cache ID 6iaQD7RoR]
24. Stockmeyer N. Using Microsoft Word's readability program. *Michigan Bar Journal* 2009;88:46.
25. Stevens J. *Applied multivariate statistics for the social sciences*. New York, NY: Routledge; 2009.
26. Thompson B. *Foundations of Behavioral Statistics: An Insight-Based Approach*. New York: The Guilford Press; 2006.
27. Zeng-Treitler Q, Kim H, Rosemblat G, Keselman A. Can multilingual machine translation help make medical record content more comprehensible to patients? *Stud Health Technol Inform* 2010;160(Pt 1):73-77. [Medline: [20841653](https://pubmed.ncbi.nlm.nih.gov/20841653/)]
28. Flores G, Laws MB, Mayo SJ, Zuckerman B, Abreu M, Medina L, et al. Errors in medical interpretation and their potential clinical consequences in pediatric encounters. *Pediatrics* 2003 Jan;111(1):6-14. [Medline: [12509547](https://pubmed.ncbi.nlm.nih.gov/12509547/)]
29. Michael J, Ayles T, Ogrin R. Development of a Translation Standard to support the improvement of health literacy and provide consistent high-quality information. *Aust Health Rev* 2013 Sep;37(4):547-551. [doi: [10.1071/AH13082](https://doi.org/10.1071/AH13082)] [Medline: [24018059](https://pubmed.ncbi.nlm.nih.gov/24018059/)]
30. Dillinger M, Lommel A. *LISA Best Practice Guide. Implementing Machine Translation*. Switzerland: Localization Industry Standards Association 2004.

31. Vreeman DJ, Chiaravalloti MT, Hook J, McDonald CJ. Enabling international adoption of LOINC through translation. *J Biomed Inform* 2012 Aug;45(4):667-673 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2012.01.005](https://doi.org/10.1016/j.jbi.2012.01.005)] [Medline: [22285984](#)]
32. Singh PM, Wight CA, Sercinoglu O, Wilson DC, Boytsov A, Raizada MN. Language preferences on websites and in Google searches for human health and food information. *J Med Internet Res* 2007;9(2):e18 [[FREE Full text](#)] [doi: [10.2196/jmir.9.2.e18](https://doi.org/10.2196/jmir.9.2.e18)] [Medline: [17613488](#)]

Abbreviations

ATA: American Translators Association
LEP: limited English proficiency
LOTE: language other than English
MANOVA: multivariate analysis of variance

Edited by G Eysenbach; submitted 03.05.16; peer-reviewed by S Alexander, A Khalid; comments to author 15.06.16; revised version received 17.06.16; accepted 17.06.16; published 28.06.16.

Please cite as:

Chen X, Acosta S, Barry AE

Evaluating the Accuracy of Google Translate for Diabetes Education Material

JMIR Diabetes 2016;1(1):e3

URL: <http://diabetes.jmir.org/2016/1/e3/>

doi: [10.2196/diabetes.5848](https://doi.org/10.2196/diabetes.5848)

PMID: [30291056](https://pubmed.ncbi.nlm.nih.gov/30291056/)

©Xuewei Chen, Sandra Acosta, Adam Etheridge Barry. Originally published in *JMIR Diabetes* (<http://diabetes.jmir.org>), 28.06.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Diabetes*, is properly cited. The complete bibliographic information, a link to the original publication on <http://diabetes.jmir.org/>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>