Contents

Original Papers

Analysis of Diabetes Apps to Assess Privacy-Related Permissions: Systematic Search of Apps (e16146)
José Flors-Sidro, Mowafa Househ, Alaa Abd-Alrazaq, Josep Vidal-Alaball, Luis Fernandez-Luque, Carlos Sanchez-Bocanegra. ................................. 2

Using Virtual Reality to Improve Health Care Providers’ Cultural Self-Efficacy and Diabetes Attitudes: Pilot Questionnaire Study (e23708)
Elizabeth Beverly, Carrie Love, Matthew Love, Eric Williams, John Bowditch. ........................................................................................................ 15

Exchanges in a Virtual Environment for Diabetes Self-Management Education and Support: Social Network Analysis (e21611)
Carlos Pérez-Aldana, Allison Lewinski, Constance Johnson, Allison Vorderstrasse, Sahiti Myneni. ................................................................. 26

Ability of Current Machine Learning Algorithms to Predict and Detect Hypoglycemia in Patients With Diabetes Mellitus: Meta-analysis (e22458)
Satoru Kodama, Kazuya Fujihara, Haruka Shiozaki, Chika Horikawa, Mayuko Yamada, Takaaki Sato, Yuta Yaguchi, Masahiko Yamamoto, Masaru Kitazawa, Midori Iwanaga, Yasuhiro Matsuyabashi, Hirohito Sone. ............................................................... 37

Public Perspectives on Anti-Diabetic Drugs: Exploratory Analysis of Twitter Posts (e24681)
Su Golder, Millie Bach, Karen O’Connor, Robert Gross, Sean Hennessy, Graciela Gonzalez Hernandez. .............................................................. 51

Diabetes Distress and Glycemic Control in Type 2 Diabetes: Mediator and Moderator Analysis of a Peer Support Intervention (e21400)
Kara Mizokami-Stout, Hwajung Choi, Caroline Richardson, Gretchen Platt, Michele Heisler. ................................................................. 64
**Analysis of Diabetes Apps to Assess Privacy-Related Permissions: Systematic Search of Apps**

José Javier Flors-Sidro¹, MSc; Mowafa Househ², PhD; Alaa Abd-Alrazaq², PhD; Josep Vidal-Alaball³,4, MD, MPH, PhD; Luis Fernandez-Luque⁵,6, PhD; Carlos Luis Sanchez-Bocanegra⁷, PhD

¹Information Systems Department, Consorci Hospitalari Provincial de Castelló, Castelló de la Plana, Spain
²Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar
³Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Fruitós de Bages, Spain
⁴Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la recerca a l’Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain
⁵Salumedia Labs, Sevillla, Spain
⁶Adhera Health Inc, Palo Alto, CA, United States
⁷Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

**Corresponding Author:**
Alaa Abd-Alrazaq, PhD
Division of Information and Computing Technology
College of Science and Engineering
Hamad Bin Khalifa University
Education City
Doha,
Qatar
Phone: 974 55708549
Email: aabdralrazaq@hbku.edu.qa

**Abstract**

**Background:** Mobile health has become a major vehicle of support for people living with diabetes. Accordingly, the availability of mobile apps for diabetes has been steadily increasing. Most of the previous reviews of diabetes apps have focused on the apps’ features and their alignment with clinical guidelines. However, there is a lack of knowledge on the actual compliance of diabetes apps with privacy and data security guidelines.

**Objective:** The aim of this study was to assess the levels of privacy of mobile apps for diabetes to contribute to the raising of awareness of privacy issues for app users, developers, and governmental data protection regulators.

**Methods:** We developed a semiautomatic app search module capable of retrieving Android apps’ privacy-related information, particularly the dangerous permissions required by apps, with the aim of analyzing privacy aspects related to diabetes apps. Following the research selection criteria, the original 882 apps were narrowed down to 497 apps that were included in the analysis.

**Results:** Approximately 60% of the analyzed diabetes apps requested potentially dangerous permissions, which pose a significant risk to users’ data privacy. In addition, 28.4% (141/497) of the apps did not provide a website for their privacy policy. Moreover, it was found that 40.0% (199/497) of the apps contained advertising, and some apps that claimed not to contain advertisements actually did. Ninety-five percent of the apps were free, and those belonging to the “medical” and “health and fitness” categories were the most popular. However, app users do not always realize that the free apps’ business model is largely based on advertising and, consequently, on sharing or selling their private data, either directly or indirectly, to unknown third parties.

**Conclusions:** The aforementioned findings confirm the necessity of educating patients and health care providers and raising their awareness regarding the privacy aspects of diabetes apps. Therefore, this research recommends properly and comprehensively training users, ensuring that governments and regulatory bodies enforce strict data protection laws, devising much tougher security policies and protocols in Android and in the Google Play Store, and implicating and supervising all stakeholders in the apps’ development process.

*(JMIR Diabetes 2021;6(1):e16146)*  doi:10.2196/16146
Introduction

Background

Diabetes mellitus (DM) is one of the most common chronic conditions around the globe. The number of people with DM has risen globally from 108 million in 1980 to 422 million in 2014 [1]. Its prevalence has been increasing everywhere, especially in middle-income countries, from 4.7% in 1980 to 8.5% in 2014. DM increases the risk of serious health problems such as myocardial infarction, renal failure, stroke, and lower limb amputation [2]. Diabetic retinopathy is one of the most important causes of blindness worldwide, especially in developed countries [3]. DM has also been linked to an increased risk of other conditions such as dementia, depression, and some types of cancer [4]. In order to reduce the risk of complications, intensive patient education and support are needed, which can be enhanced by the use of mobile technology.

Along with the exponential increase in the number of health apps [5,6], in particular the number of diabetes apps has increased significantly in the last several years [7]. Mobile health (mHealth) has become a major vehicle of support for people living with diabetes, and the availability of mobile apps for diabetes has been steadily increasing. Most of the previous reviews of diabetes apps have focused on their features and their alignment with clinical guidelines [8,9]. However, there is a lack of knowledge on the actual compliance of diabetes apps with privacy and data security guidelines.

Therefore, there is a growing concern to review diabetes apps because in many cases they do not possess the quality and content that they should according to their own declared purposes [10,11]. In addition, some studies that have investigated the effectiveness of mobile apps clearly demonstrate data privacy problems [12], as well as a lack of transparency with the provided information [13].

Studies on mHealth and privacy have raised some serious concerns in recent years. Because very sensitive information is increasingly accessed and shared using mobile apps, there is an obvious need for clinicians, software developers, users, and patients to be aware of and trained on information privacy aspects. Personal data may be collected through different means, such as being entered directly by the user or being recorded by the phone’s camera, microphone, or paired wireless device (eg, Bluetooth gluometer apps). It is crucial to note that the treatment of these critical data demands a special approach regarding security and privacy. However, some apps do not even provide information regarding their privacy policies. In some instances, these privacy terms are difficult to understand by nontechnical users, and some privacy policies may even be regarded as abusive. To make matters worse, the ecosystem of mobile apps is so complex that even app developers and users may not know with whom the data is being shared and for what purpose [14-16].

An additional challenge is that very often stakeholders are not involved in the app development process and consequently cannot provide feedback on privacy preferences [10].

To deal with these issues, some researchers such as Stoyanov et al [17] have attempted to develop a suitable framework—the Mobile App Rating Scale—that allows for the evaluation of the quality of apps. Alternatively, other investigations have focused specifically on privacy or legal issues [18]. In the case of mHealth for diabetes, recent reviews looked into aspects linked to the efficacy of interventions [19,20] but did not address aspects related to privacy. Other research has investigated privacy aspects in generic mHealth apps [12,21]. However, to the best of our knowledge, this study is the first to focus on investigating privacy issues and dangerous permissions in diabetes mobile apps. Studies looking at diabetes apps have not conducted in-depth analyses of dangerous permissions on the Android platform [22].

Objectives

The aim of this study was to evaluate the privacy-related permissions of Android diabetes apps in Google’s Play Store using a semiautomatic approach that relies on the extraction of privacy-related features (eg, permissions, terms of usage). This approach was designed to assist in identifying strategies to raise the awareness of app users, patients, and clinicians. To illustrate our approach, we provide two case studies of diabetes apps that were comprehensively analyzed (Multimedia Appendix 1).

Methods

Study Design

The first step in this study was the extraction of metadata from mobile apps’ metadata using a web-based application programming interface (API) [23]. We used the platform 42Matters, which offers a web-based commercial tool that facilitates access to the Android Google Play Store and to other mobile platforms’ apps’ metadata through a proprietary API [24]. Searches were conducted with the developed script module 42Matters’ index of Android apps. Since the 42Matters platform did not allow the extraction of privacy-related permissions from Apple’s App Store, the research centered on Android apps from Google’s Play Store. Data extraction was focused on potentially dangerous permissions [25] that allow the requesting app access to private user data or control over the mobile device, both of which can negatively impact the user. Because this type of permission introduces potential risk, the system does not automatically grant it to the requesting app. Our methodology was based on similar studies of health apps that used the 42Matters platform, but focusing on privacy-related information [26,27].

In order to complement the quantitative results already presented, we described and investigated two very popular and well-rated diabetes apps (presented in Multimedia Appendix 1) from a qualitative perspective.
For the extraction of the diabetes apps’ metadata, we first devised the architecture [28] and subsequently developed the corresponding software module for the automatic extraction of mobile app metadata using the web-based API of 42Matters. The output of this module is a data set stored locally in a comma-separated values (CSV) file. The source code for the module was released under the GNU AGPLv3 license and can be found on the GitHub link [29]. This module is capable of querying the API of the 42Matters platform to retrieve metadata related to diabetes apps, including the Android permissions required by the apps. The module was designed to extract apps with the following search parameters: (1) language (we searched for English-language apps), (2) keyword search (we searched for apps whose titles included the root words “diabet” and “mellitus”), and (3) app categories (we selected the categories medical, health and fitness, lifestyle, and education).

The resulting apps were manually reviewed (see Multimedia Appendix 1) to assess whether they were related to diabetes. All apps were related to diabetes, but we did not address the quality of their content. As explained in the “Limitations” section, choosing a method where search fields matched the description—and not only the title—would have resulted in more apps, many of which would not have been related to diabetes.

Once the most suitable app categories were identified, it was then possible to move on to design the entire app selection process, which consisted of the following steps (see Figure 1):
Figure 1. App selection process flowchart.

- Step 1: “Identification” phase—all of the diabetes apps that contained the root words “diabet” or “mellitus” in an app’s title field were selected, resulting in 882 apps; by matching diabet or mellitus, it was possible to ensure that any relevant potential variations of the words that contained these root words (ie, diabetes, diabetic, diabetics, mellitus, etc) were included in the search.
- Step 2: “Category filtering” phase—in order to guarantee that only relevant diabetes apps were included in the study, all the retrieved apps that did not belong to the medical, health and fitness, education, or lifestyle categories [30] were automatically filtered out by the 42Matters script module and excluded from the study; this filtering resulted in 732 apps.
- Step 3: “Screening” phase—in this phase, we manually filtered apps and excluded 5 diabetes apps related to pets, 1 discontinued app, and 55 duplicated apps; this screening resulted in 671 apps.
- Step 4: “Eligibility” phase—we excluded apps that did not have a minimum of 50 downloads, and therefore discarded 174 apps.
• Step 5: “Inclusion” phase—the resulting 497 apps were analyzed, which were the objects of analysis of this research.

Data Extraction: Retrieved Metadata Fields

After the final set of apps was selected in June 2019, a process was initiated to extract all the relevant metadata and information, which were stored in a CSV file. All the retrieved fields are described in the table below.

Table 1. Description of apps’ retrieved metadata as provided by 42Matters.

<table>
<thead>
<tr>
<th>App’s metadata field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Main name of the app</td>
</tr>
<tr>
<td>Price</td>
<td>Price and currency (0 if it was free)</td>
</tr>
<tr>
<td>Permission</td>
<td>Required Android permissions of the app</td>
</tr>
<tr>
<td>Rating</td>
<td>App’s average rating from 0 to 5 (0=worst, 5=best)</td>
</tr>
<tr>
<td>Number of downloads</td>
<td>Number of times the app was downloaded</td>
</tr>
<tr>
<td>Number of ratings</td>
<td>Number of times the app was rated</td>
</tr>
<tr>
<td>Contains advertising</td>
<td>True if the app contained advertising and false if it did not</td>
</tr>
<tr>
<td>Category</td>
<td>Category to which the app belonged (medical, health and fitness, education, or lifestyle)</td>
</tr>
<tr>
<td>Short description</td>
<td>Short description of the app’s declared purpose</td>
</tr>
<tr>
<td>Website</td>
<td>Website of the app</td>
</tr>
<tr>
<td>Privacy policy</td>
<td>Website showing the app’s privacy policy</td>
</tr>
</tbody>
</table>

Extraction of Android Privacy-Related Permissions

Starting with Android 6.0 (API 23 level), users grant permissions to apps while using them, not when an app is installed. On the one hand, this approach simplifies the process of installing the app because the user does not need to grant permissions when installing or updating the app. In addition, it provides the user with more control over the app’s functionalities because users can revoke the granted permissions from the app’s configuration screen at any time. On the other hand, this new approach complicates the app’s usability because dangerous permissions have to be granted while using the app, which poses an additional challenge for untrained users. Android distinguishes between 4 categories of permissions: normal, signature, dangerous, and special [31].

Signature and special permissions will not be explained here because they are rarely used and were not found in any of the apps included in our research. The most frequently requested permissions are normal and dangerous permissions. If an app declares a normal permission in its manifest, the system grants permission in its manifest, the system grants permission to it automatically without the user’s intervention. On the other hand, Android considers dangerous permissions as critical because they allow apps to access users’ critical data.

More concretely, an Android dangerous permission [25,32] allows the requesting app access to private user data or control over the mobile device. Because this type of permission allows developers to access users’ data, photos, and videos stored on the device, it introduces potential risk, and the system does not automatically grant it to the requesting app [33,34].

In brief, normal permissions do not put the user’s privacy at risk directly. Consequently, if an app declares a normal permission in its metadata, the system grants permission to it automatically without the user’s intervention. On the other hand, a dangerous permission allows an app to access the user’s critical data, and consequently the user should explicitly authorize this permission [35]. The 10 most required dangerous permissions found in this research are shown in Multimedia Appendix 2.

Results

App Functions

The process described in the “Methods” section retrieved a total of 497 apps (Multimedia Appendix 3). The breakdown of privacy-related permissions is summarized in Table 2. Most of the apps required at least one dangerous permission.

Table 2. Summary of the privacy-related main features of retrieved diabetes apps.

<table>
<thead>
<tr>
<th>Assessed parameter</th>
<th>Diabetes apps (N=497), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not require any permissions (either normal or dangerous)</td>
<td>89 (17.9)</td>
</tr>
<tr>
<td>Only requires normal permissions</td>
<td>111 (22.3)</td>
</tr>
<tr>
<td>Requires at least one dangerous permission</td>
<td>297 (59.8)</td>
</tr>
<tr>
<td>Does not provide a website link to its privacy policy</td>
<td>141 (28.4)</td>
</tr>
<tr>
<td>Contains advertising</td>
<td>199 (40.0)</td>
</tr>
</tbody>
</table>
The reason for apps not requesting any permissions is that they serve very basic functions (eg, calculators, logs, diaries, etc) that only need access to very basic and noncritical Android resources. Only 22.3% (111/497) of the apps required normal (noncritical) permissions alone. On the other hand, 59.8% (297/497) of the apps required at least one dangerous permission. This might be partially justified by these apps’ more advanced functionalities (eg, doctor-patient interaction, connecting to a glucometer, calorie-burning calculation, scanning the barcode of diabetic food, etc).

Regarding privacy, it was worrying to discover that 28.4% (141/497) of the apps did not return the privacy policy metadata field, consequently posing additional difficulty for users to adequately understand how these apps would treat very sensitive personal information.

Finally, 40.0% (199/497) of the apps contained advertising, which can imply the sharing of critical personal data (eg, a user’s precise location) with unknown third parties for geolocated advertisement. Consequently, because the advertising business model in the mobile ecosystem is usually linked to the sharing or selling of critical personal data [36], the aforementioned findings unquestionably confirm the necessity to educate users and raise awareness regarding user privacy in diabetes apps.

**Dangerous Permissions**

As explained below, dangerous permissions refer to permissions that might lead to data breaches of private information [37]. From the 497 diabetes apps included in our final analysis, a substantial number of them—297 (59.8%)—required dangerous permissions. Table 3 shows, in decreasing order, which dangerous permissions were most frequently requested by the apps.

<table>
<thead>
<tr>
<th>Dangerous permission</th>
<th>Diabetes apps that requested it (N=497), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write external storage</td>
<td>272 (54.7)</td>
</tr>
<tr>
<td>Read external storage</td>
<td>169 (34.0)</td>
</tr>
<tr>
<td>Access coarse location</td>
<td>103 (20.7)</td>
</tr>
<tr>
<td>Access fine location</td>
<td>95 (19.1)</td>
</tr>
<tr>
<td>Camera</td>
<td>89 (17.9)</td>
</tr>
<tr>
<td>Get accounts</td>
<td>82 (16.5)</td>
</tr>
<tr>
<td>Read phone state</td>
<td>81 (16.3)</td>
</tr>
<tr>
<td>Record audio</td>
<td>39 (7.8)</td>
</tr>
<tr>
<td>Call phone</td>
<td>23 (4.6)</td>
</tr>
<tr>
<td>Read contacts</td>
<td>22 (4.4)</td>
</tr>
<tr>
<td>Others (the sum of the remaining dangerous permissions)</td>
<td>28 (5.6)</td>
</tr>
</tbody>
</table>

In addition, Figure 2 illustrates the number of apps that required each of the top 14 dangerous permissions, arranged by category. The four quadrants represent each of the four categories to which the apps belonged: education, health and fitness, medical, and lifestyle. In addition, the “Advertising” tag indicates whether an app contained advertising: the ones in blue contained advertising, while the ones in red did not. The x-axis shows the number of apps, while the y-axis lists the 14 most requested dangerous permissions.
Figure 2. The top 14 dangerous permissions by app category (lifestyle, medical, education, and health and fitness) and type of privacy-related permission requested, as well as whether they included advertising (“True”) or not (“False”).

Discussion

Principal Results and Comparison With Previous Work

Although we identified the apps requesting access to the camera (89/497, 17.9%), we need to study the actual usage of apps in order to fully understand the context before we consider that access to be a potential risk. For instance, in the case of diabetes, it is very common to use the camera for food logging. On the other hand, except for advertising or fitness tracking (eg, calorie counting), the need for the user’s geolocation data seems difficult to justify. In this sense, what might be acceptable in one app might not be reasonable in others. Similar studies found that 77 of 186 (41.4%) permissions requested by 58 popular German mHealth apps were not related in any way to the apps’ functionalities [38]. Moreover, 15 of 42 (35.7%) Android health and well-being apps accredited by the UK’s NHS Health Apps Library requested critical permissions for unjustifiable reasons [12]. Similarly, other research concluded that several popular mental health apps and mHealth apps requested permissions that were not aligned with the apps’ stated purposes [14,21]. One of the consequences of requesting unnecessary dangerous permissions is a decrease in users’ trust, acceptance, and use of these apps.

Another finding of this study was that 95.4% of the apps were free of charge. The business model of free apps is, in most cases, based on advertising (through services such as Google AdMob), resulting in the disclosure of users’ critical data, either directly (through the app itself) or indirectly (through Google’s commercial advertising platforms).

The reliance on advertising of some of the studied apps might be linked to the high number of apps requesting geolocation, since location can increase advertisement revenue. A study on
NHS-accredited apps found some evidence that patients’ data were information for advertisers [12]. Other studies also found that users’ information was shared in 19 of 24 popular medication-related apps in the United Kingdom, the United States, Canada, and Australia [39]. Research of privacy in the top 36 mental health and smoking cessation apps also found a lack of compliance with disclosing or sending data to third-party providers [40]. Although app developers usually claim that they do not collect or share personally identifiable data, users can be easily identified by correlating advertising services using data analytics [39].

In addition, 28.4% of the studied apps did not provide a privacy policy website, which corroborates results from other research that demonstrated that 48% of 17,991 free Android apps did not have a privacy policy [18]. Building on this finding, 81% of 154 Android apps related to hypertension and diabetes did not refer to a privacy policy [33]. In addition, a privacy policy was missing in 417 of 600 (69.5%) prominent mHealth apps [41]. Most likely, had we not discarded less reliable apps in our research, the percentage of apps that did not provide a link to a website with their privacy policy would have been higher [34]. The lack of a privacy policy is a critical fault, as it prevents users from properly understanding how apps treat their very sensitive personal information. Further, the discrepancy between apps’ privacy policies and their actual features has been reported in several studies [12,18]. This issue might be partially attributed to the fact that app developers have insufficient knowledge about privacy best practices [42].

In our study, 59.8% of apps required at least one dangerous permission, the two most requested being write external storage (54.7%) and read external storage (34.0%). This finding confirms the results from previous research. For instance, the most common dangerous permissions requested by the most popular freeware mHealth apps were write external storage (90%) and read external storage (50%) [34]. For prominent mental health apps in the Google Play Store, the most frequently requested permissions were also write (73%) and read (73%) external storage. In addition, these two permissions were the most requested (79%) in medicine-related apps in the Google Play Store in the United Kingdom, the United States, Canada, and Australia [38]. These permissions may indeed jeopardize users’ privacy because they allow developers to access users’ data, photos, and videos stored on the device [33,34]. Another relevant finding was that health and fitness apps usually requested more dangerous permissions than apps belonging to other categories [21].

Apps’ ever-changing functionality and privacy policies, as well as their complexity, do not facilitate matters, either. Moreover, having to manually accept dangerous permissions when using an app poses an additional challenge that can have detrimental consequences, particularly for less knowledgeable users. For instance, individuals with low literacy rates or the elderly would require adequate training to truly understand what they are consenting to before using diabetes apps. Existing tools to evaluate eHealth literacy skills [43] do include security awareness as one of their dimensions. However, the complexity of potential security issues is increasing, and it might be necessary to develop new tools and training methods for both patients and health care providers.

**Practical Implications**

These findings have very important practical implications for users, physicians, developers, and policy makers [44,45]. To select an appropriate mobile app for diabetes, end users should be aware of what type of personal data is collected, used, and shared by a certain app by carefully reading the app’s description, terms of use, and privacy policy.

In addition, it is imperative to emphasize the need for training so that users are able to understand complex privacy policies and terms of service and are fully aware of the privacy risks derived from the sharing of their data with third parties. Users should also be knowledgeable about the different types of dangerous permissions so that they can discern how each particular permission may jeopardize their data. The ultimate goal is to empower users so that they can autonomously and proficiently deny access to any unjustifiable dangerous permission.

To minimize the privacy risks derived from using diabetes apps, savvy users should use AdBlock or encryption apps [33]. Moreover, health care providers should ensure that the apps they recommend to patients adhere to a strict privacy code, and they should assist users in selecting suitable apps by explaining both the apps’ benefits and their risks.

App developers should enforce their apps’ full compliance with internationally recommended standards and practices [46-49]. Specifically, developers must ensure that their apps’ privacy policies are always readily available, very simple to read, and able to be understood by any user. Further, their apps should never request dangerous permissions not directly related to the apps’ declared purpose. Developers should not—without the users’ explicit consent—collect, use, or share user data for any purpose outside of the predefined scope of the app, and all data sharing practices should be transparently disclosed to users. Last but not least, developers should be aware of diverse privacy laws and data protection legislation, which differ greatly depending on the country or region of use.

In terms of privacy laws, apps tend to adhere to the data protection legislation in the developers’ country of origin but not in the apps’ country of use. Therefore, regulators around the world should collaborate to establish a specific international accreditation program for diabetes apps. Such a program should be based on unified privacy best practices in which user privacy is the main priority. Because app developers reserve the right to change their privacy policies at any given time and modify their apps’ declared purpose and functionalities, regulators should regularly monitor developers’ adherence to the recommended privacy practices. As well, regulators should emphasize developers’ responsibility and accountability for protecting user data. In addition, app stores should mandate stringent principles and standards that actually compel developers to provide simple and intelligible privacy policies in their apps, especially taking into consideration untrained or illiterate users.
Limitations

We opted to use the free version of the commercial platform 42Matters instead of the Google Play Store because the Google Play Store had a limit of 250 apps per query.

Another limitation was that the developed module exclusively searched for all diabetes apps that contained the root words diabet or mellitus in the title field. There are some diabetes apps in which the aforementioned root words appear in the app’s description but not in the app’s name. Therefore, some diabetes-related apps may have been excluded from the study. However, this criterion was selected for two principal reasons: (1) to ensure that only truly diabetes-related apps were retrieved, and (2) to make the best use of limited resources (there was neither enough time nor enough labor to thoroughly screen 4700+ apps, many of which bore no relation whatsoever to diabetes). In this sense, our research was not intended to be exhaustive. Rather, we wanted to quantify and evaluate the overall privacy characteristics of the most representative sample of diabetes-related apps. A broader search (ie, to query for all apps that contained the root words diabet or mellitus in the apps’ descriptions) would certainly have yielded many false positives of apps unrelated to diabetes and hence required a very resource-intensive manual screening of the apps, which would have been an unnecessary complication of the overall analysis process.

The study did not comprehensively address either the fact that the number of permissions an app requests does not necessarily reflect how risky the app may be. For instance, an app requesting, unnecessarily, a single dangerous permission, could seriously endanger users’ personal data by collecting and illegitimately sharing them. On the other hand, an app requesting multiple dangerous permissions, but for valid technical or functional needs, could be considered safe. Therefore, the amount of personal information that users are putting at risk depends on many factors, such as the app’s functionality, the permissions it requests, and the context in which these permissions are being used [50]. To perform a more complete assessment of apps’ privacy risks, additional technical, human, and contextual research (eg, analysis of the skills of patients using diabetes apps) should be conducted. For example, when dealing with privacy issues in health apps, an important factor to be considered would be the legitimacy of the request, as highlighted in a recent publication on mHealth apps for cancer in which the authors evaluated a new scale to assess the privacy policies of mHealth apps [51]. Tracking users’ location might be fair in the case of reporting a medical emergency (eg, hypoglycemic crisis).

Although the methodology employed in this research was robust and Google is continuously improving Android and the Play Store’s security policy, this study found evidence that it is extremely difficult to prove whether diabetes apps actually comply with their privacy policies. In fact, even Google cannot control the many malicious apps that are frequently uploaded by hackers in its Play Store and is consequently forced to periodically remove massive numbers of these fraudulent apps [52-54]. Further, a recently published two-year study discovered 2040 potential counterfeit apps that contained malware in the Google Play Store [55].

This study did not cover all of the elements related to the privacy and security of diabetes apps. Privacy protection cannot be guaranteed solely by controlling permissions; for instance, unsecure internet connections can also jeopardize the privacy of mobile app users. Finally, our study only evaluated the apps on one app store; the privacy policies and the requested dangerous permissions in other app stores, such as Apple’s App Store or Samsung’s Galaxy Store, might have yielded different outcomes. However, Android’s Google Play Store was also chosen due to its popularity.

Future Research

A possible expansion of the research could include investigating those diabetes apps that were excluded from this research, either because they belonged to nonrelevant categories or because the developed module did not search for the root words in the apps’ description field. Future research could also focus on analyzing the taxonomy of app categories and match them to officially recognized and standardized clinical categories, such as the Systematized Nomenclature of Medicine Clinical Terms or Medical Subject Headings. Related to that, there is a new trend emerging toward the creation of machine learning approaches to identify privacy issues in mobile apps [56,57]. However, to the best of our knowledge, those methods have unfortunately not yet been applied to health apps. Further, there is a need for homogenous approaches for the assessment of privacy in health apps, as was highlighted recently in a scoping review addressing the issue [58].

Finally, from a legal perspective, although many diabetes apps are available worldwide, their privacy policies usually only comply with the specific national data protection regulations of the developers’ country or region of origin. For instance, the BeatO SMART Diabetes Management app claims that both its privacy policy and its terms of use fully adhere to Indian law, but if this app were to be used in the Middle East or the European Union, it would be unclear whether it would also comply with data protection laws in the country or region of use. This could indeed be another matter of study.

Conclusions

If privacy issues in diabetes mobile apps are not dealt with carefully, users may unwillingly and unknowingly share very sensitive private data. Therefore, it is crucial that all stakeholders are involved in the development of diabetes apps from the very beginning of the process in order to ensure apps’ absolute compliance with data protection regulations and user privacy.

As the economic value of personal data increases [59], a completely new business model for apps has emerged: users pay for the usage of an app with their data, which is then sold to third parties, such as advertising clients [60]. The lesson to be learned is that there is a price to pay in exchange for free apps, usually at the expense of privacy. Consequently, new control measures are needed to enable users to decide which personal information they are willing to disclose in return for a certain service [61].
The importance of personal data protection laws and their endorsement are of utmost importance. Well-designed privacy policies may protect individuals by requiring consent for the collection, use, disclosure, or retention of sensitive personal and health information, and they may regulate the use of these extremely sensitive data, allowing users to modify their information as well as to revoke their previous consent. Therefore, we recommend proper training for users, enforcement of strict data protection laws by governments and regulatory bodies, much tougher security policies and protocols in both Android apps and the Google Play Store, and the implication and supervision of all stakeholders in the app development process.

Authors' Contributions
JIF-S was the principal investigator. He designed the majority of the work, supervised the research, and took over most of the data interpretation and writing of the manuscript. In addition, he was responsible for developing the software module for extracting apps’ metadata. MH and AA-A significantly contributed to the results and discussion sections of the paper. JV-A contributed to the overall manuscript and study by providing a clinical perspective. LF-L conceived the original research idea and greatly assisted with the design of the methodology and with the discussion section. Finally, CLS-B’s contribution to the analysis and interpretation of the results was fundamental. All of the authors contributed to and approved the manuscript.

Conflicts of Interest
LF-L is co-founder of Adhera Health Inc (USA), a digital health company that provides digital therapeutic solutions for people with chronic conditions

Multimedia Appendix 1
Qualitative results of case studies.
[DOCX File, 5315 KB - diabetes_v6i1e16146_app1.docx]

Multimedia Appendix 2
Top 10 Android’s dangerous permissions identified.
[DOCX File, 16 KB - diabetes_v6i1e16146_app2.docx]

Multimedia Appendix 3
Comma-separated values files.
[DOCX File, 14 KB - diabetes_v6i1e16146_app3.docx]

References
Nov;60(11):2129-2138. [doi: 10.1007/s00125-017-4360-x] [Medline: 28725914]
5. Klonoff DC. The current status of mHealth for diabetes: will it be the next big thing? J Diabetes Sci Technol 2013 May
8. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014 Apr 09;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]


Abbreviations

API: application programming interface
CSV: comma-separated values
DM: diabetes mellitus
mHealth: mobile health

©José Javier Flors-Sidro, Mowafa Househ, Alaa Abd-Alrazaq, Josep Vidal-Alaball, Luis Fernandez-Luque, Carlos Luis Sanchez-Bocanegra. Originally published in JMIR Diabetes (http://diabetes.jmir.org), 13.01.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), 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.
Using Virtual Reality to Improve Health Care Providers’ Cultural Self-Efficacy and Diabetes Attitudes: Pilot Questionnaire Study

Elizabeth Ann Beverly1*, PhD; Carrie Love2*, MFA; Matthew Love2*, MFA; Eric Williams2*, MFA; John Bowditch2*, MFA

1Department of Primary Care, Heritage College of Osteopathic Medicine, Ohio University, Athens, OH, United States
2Game Research and Immersive Design Lab, J Warren McClure School of Emerging Communication Technologies, Ohio University, Athens, OH, United States
*all authors contributed equally

Corresponding Author:
Elizabeth Ann Beverly, PhD
Department of Primary Care
Heritage College of Osteopathic Medicine
Ohio University
1 Ohio University
Grosvenor Hall 357
Athens, OH, 45701
United States
Phone: 1 17405934616
Fax: 1 7405932205
Email: beverle1@ohio.edu

Abstract

Background: In southeastern Appalachian Ohio, the prevalence of diabetes is 19.9%, nearly double that of the national average of 10.5%. Here, people with diabetes are more likely to have a delayed diagnosis, limited access to health care, and lower health literacy. Despite the high rates of diabetes in the region, the availability of endocrinologists and certified diabetes care and education specialists is limited. Therefore, innovative strategies to address the growing diabetes care demands are needed. One approach is to train the primary care workforce in new and emerging therapies for type 2 diabetes to meet the increasing demands and complexity of diabetes care.

Objective: The aim of this study was to assess the effectiveness of a virtual reality training program designed to improve cultural self-efficacy and diabetes attitudes.

Methods: Health care providers and administrators were recruited from large health care systems, private practices, university-owned hospitals or clinics, Federally Qualified Health Centers, local health departments, and AmeriCorps. Providers and administrators participated in a 3-hour virtual reality training program consisting of 360-degree videos produced in a professional, cinematic manner; this technique is called virtual reality cinema (cine-VR). Questionnaires measuring cultural self-efficacy, diabetes attitudes, and presence in cine-VR were administered to providers and administrators before and after the program.

Results: A total of 69 participants completed the study. The mean age of the sample was 42.2 years (SD 13.7), 86% (59/69) identified as female, 83% (57/69) identified as White, 86% (59/69) identified as providers, and 25% (17/69) identified as nurses. Following the training program, we observed positive improvements in all three of the cultural self-efficacy subscales: 
- Cognitive (mean change –1.29; t_{65} = –9.309; P < .001),
- Practical (mean change –1.85; t_{65} = –9.319; P < .001), and
- Affective (mean change –0.75; t_{65} = –7.067; P < .001). We observed the largest magnitude of change with the subscale, with a Cohen d of 1.16 indicating a very large effect. In addition, we observed positive improvements in all five of the diabetes attitude subscales: Need for special training (mean change –0.21; t_{67} = –6.154; P < .001), Seriousness of type 2 diabetes (mean change –0.34; t_{67} = –8.114; P < .001), Value of tight glucose control (mean change –0.13; t_{67} = –3.029; P = .001), Psychosocial impact of diabetes (mean change –0.33; t_{67} = –6.610; P < .001), and Attitude toward patient autonomy (mean change –0.17; t_{67} = –3.889; P < .001). We observed the largest magnitude of change with the Psychosocial impact of diabetes subscale, with a Cohen d of 0.87 indicating a large effect. We observed only...
one significant correlation between presence in cine-VR (ie, Interface Quality) and a positive change score (ie, Affective self-efficacy) \( r = .285; P = .03 \).

**Conclusions:** Our findings support the notion that cine-VR education is an innovative approach to improve cultural self-efficacy and diabetes attitudes among health care providers and administrators. The long-term impact of cine-VR education on cultural self-efficacy and diabetes attitudes needs to be determined.

*(JMIR Diabetes 2021;6(1):e23708) doi:10.2196/23708*

**KEYWORDS**

virtual reality; diabetes attitudes; cultural self-efficacy; health care providers; VR; diabetes; training

**Introduction**

Appalachia is a 205,000-square-mile region that encompasses 420 counties in 13 US states from Mississippi to New York. Ohio’s Appalachian region encompasses 32 counties [1], of which 16 are designated as economically at risk or distressed [2]. Here, 17.2% of the population live below the poverty line as compared to 14.4% for the rest of the state [3], and the counties with the highest poverty rates, ranging from 22.5% to 30.2%, are Appalachian [3]. People who live in Appalachian Ohio are more likely to be unemployed, have lower educational achievement, and limited access to transportation [4]. These social determinants of health contribute to the health disparities observed among people living in this region [5].

One health disparity disproportionately affecting people in Appalachian Ohio is diabetes [5]. An alarming 19.9% of adults in southeastern Ohio have diabetes [6], which is nearly double the national average of 10.5% [7]. In this region, people are more likely to have a delayed diabetes diagnosis, limited access to health care, lower health literacy, and lower empowerment [8,9]. For these reasons, people here are more likely to have macrovascular and microvascular complications, lower limb amputations, and depression [9-11]. Despite the high rates of diabetes in the region, the availability of endocrinologists and certified diabetes care and education specialists in Appalachian Ohio is limited [12]. Therefore, innovative strategies to address the growing diabetes care demands are needed.

One approach is to train the primary care workforce in new and emerging therapies for type 2 diabetes to meet the increasing demands and complexity of diabetes care. Primary care providers deliver more than 90% of the clinical care to people with type 2 diabetes in the United States [13]. This is even more pertinent in rural America where family physicians comprise a greater proportion of the workforce and provide comprehensive and irreplaceable care to the community [14]. Therefore, tailored continuing education for rural primary care providers and their staff is critical. Continuing education should address standards of medical care for diabetes as well as cultural competency and attitudes toward diabetes. Studies show that health care providers’ attitudes toward diabetes influence their approach to care (eg, paternalistic vs patient-centered care) and how they interact with people with diabetes [15-18]. Furthermore, continuing education that recognizes the unique cultural contributions of regions like Appalachian Ohio is necessary to improve providers’ ability to care for people from different backgrounds [19,20]. People from Appalachia share common language, behaviors, dietary habits, and value systems. Health care providers who understand their patients’ cultural backgrounds are more likely to observe improvements in diabetes outcomes and patient satisfaction [21,22]. Thus, tailoring continuing education to address diabetes attitudes and Appalachian culture is critical to improve the quality of care to an ever-increasing number of people with diabetes in Appalachian Ohio.

Virtual reality cinema (cine-VR) is an innovative educational technique that has the potential to transform the delivery and content of continuing medical education. Cine-VR is dynamic, accessible, and adaptable to providers’ needs and preferences [23]. Cine-VR gives providers access to life-like medical encounters without risk or harm to the patient. Further, cine-VR offers providers a glimpse into the lives of patients and culture of the region. These qualities are invaluable to geographically and culturally distinct regions like Appalachian Ohio.

For this study, we developed a 3-hour cine-VR training program designed to educate providers and administrators about diabetes, social determinants of health, and Appalachian culture. The aim of the study was to assess the effectiveness of cine-VR training in improving health care providers’ and administrators’ cultural sensitivity and diabetes attitudes. We hypothesized that cine-VR training would improve cultural self-efficacy and diabetes attitudes.

The following are our hypotheses:

1. Levels of cultural self-efficacy will increase after the 3-hour cine-VR training program.
2. Diabetes attitudes will improve after the 3-hour cine-VR training program.
3. Positive changes in cultural self-efficacy will be associated with increased presence in the cine-VR scenarios.
4. Positive changes in diabetes attitudes will be associated with increased presence in the cine-VR scenarios.

**Methods**

**Overview**

The purpose of this pilot study was to call attention to social determinants of health and Appalachian culture and to delineate their relationship to diabetes via 360-degree cine-VR simulations. Specifically, we administered questionnaires to providers and administrators before and after a cine-VR training program in order to (1) assess changes in cultural self-efficacy pre- and posttraining, (2) assess changes in diabetes attitudes pre- and posttraining, and (3) examine the relationship between changes in cultural self-efficacy and diabetes attitudes and
presence in cine-VR. The Ohio University Office of Research Compliance approved the protocol (Institutional Review Board No. 19-X-99) and all recruitment procedures and materials.

Recruitment

Providers and administrators were recruited from large health care systems, private practices, university-owned hospitals or clinics, Federally Qualified Health Centers, local health departments, and AmeriCorps. In Appalachian Ohio, the majority of providers practiced at large health care systems and Federally Qualified Health Centers. Specifically, participants were recruited via emails from the Ohio University Diabetes Institute listserve and Area Health Education Center listserve, advertisements in social media, flyers in the community, and brief announcements at educational events. Participants included physicians, nurse practitioners, registered nurses, pharmacists, dietitians, certified diabetes educators, physical therapists, dentists, community health workers, and health care administrators and staff (eg, health department employees, free clinic directors, and AmeriCorps service members). The majority of providers specialized in primary care. Health care administrators were recruited given their role in health care–related decisions and their impact on quality of care. Additionally, administrators play a significant role in the assimilation of evidence-based management and training, and cine-VR has the potential to be an evidence-based educational training model.

Power Analysis

We conducted an a priori power analysis using Statulator [24], an online statistical calculator, which determined that a total sample size of 34 participants was estimated to achieve 80% power at a 5% significance level (P<.05) and to detect an effect size of 0.30.

Cinematic 360-Degree Virtual Reality Simulations

We hosted nine 3-hour training programs in Athens, Ohio. These training programs utilized 360-degree, virtual reality, professionally produced video in a cinematic manner to educate providers and administrators about diabetes, social determinants of health, and Appalachian culture. In the Using Virtual Reality to Visualize Diabetes in Appalachia program, participants watched 10 cine-VR simulations and two traditional films and observed interactions among the main character and her primary care physician, pharmacist, family, and community [25]. The main character in the simulations is Lula Mae, a 72-year-old woman with type 2 diabetes living in Appalachian Ohio. She is a widow; her husband died 27 years ago from a heart attack. She has three adult children and seven grandchildren. She cares full time for her adult son who suffered a traumatic brain injury from serving in the US Army. Lula Mae and her adult son live in an old house originally belonging to her grandparents. Her two adult daughters and grandchildren live on the same family land in their separate homes. Lula Mae is a source of care and support for her entire family, from her own children to her grandchildren. In doing so, her own health care needs come second to the daily needs of the people she loves. Despite Lula Mae’s struggles, we learn about the strengths of Appalachian culture and the resiliency one person can have if providers invest the time to connect with her one-on-one.

Training Program Curriculum

The Ohio University team developed a detailed curriculum taught synchronously with the cine-VR simulations. The curriculum included 12 modules that addressed the following content: (1) diabetes burnout, (2) food insecurity, (3) strengths of Appalachian culture, (4) rural transportation barriers, (5) elements of an effective patient-provider relationship, (6) diabetes and psychosocial issues, (7) high cost of diabetes medications, (8) gender roles in Appalachia, (9) cultural values in Appalachia, (10) diabetes complications, (11) diabetes comorbidities, and (12) patient-provider communication. An experienced behavioral diabetes researcher (EB) trained in interactive lecturing delivered all nine training sessions. The participants were encouraged to interact with each other and the lecturer. The lecturer incorporated straightforward and rhetorical questions to engage the participants. The simulations and curriculum were designed to increase cultural self-efficacy, improve diabetes attitudes, and increase presence in cine-VR. We provided 3.0 continuing medical education or continuing education credits for health care providers at no cost. Integrity of the education was ensured via a written curriculum, preapproved educational materials, and investigator observation of the training sessions.

Virtual Reality Technology

Working with the Ohio University’s Game Research and Immersive Design Lab, we leveraged a coalition of experts from Ohio University’s Diabetes Institute and the medical school, school of nursing, social work program, nutrition program, communication sciences and disorders program, school of film, theater program, and visual communication school. The interdisciplinary team consisted of one physician, three nurses, one social worker, one clinical psychologist, one audiologist, one registered dietitian, one health behaviorist, five filmmakers, four scriptwriters, and two website developers. This collaboration allowed us to create educational content that was not only medically accurate but emotionally powerful and visually stunning. Each series began with a traditionally shot short film to set the stage between Lula Mae and her relationship with a provider. This was followed by three cine-VR simulations that opened narrative windows into her daily life, her world, and her struggles. The fifth and sixth simulations of each series were guided simulations, a cine-VR face-to-face conversation with Lula Mae’s provider and Lula Mae herself. This six-video pattern was repeated twice, once covering Lula Mae’s relationship with her primary care provider and once covering her relationship with her local pharmacist.

The cine-VR simulations narratively demonstrated how Lula Mae’s social determinants of health and environment shaped her behaviors. Capturing those moments with camera systems that allow the audience to see a full 360-degree sphere created opportunities to present information in ways not possible with traditional filming methods. For example, when inside Lula Mae’s home, we saw the disorganization and chaos that resulted from a lack of social support. When the family car was stranded on the side of a remote road, we saw the transportation barriers
and isolation that families face in rural areas without public transportation. As a result of the 360-degree filming techniques employed, the team was able to present much more information about Lula Mae’s life and the factors affecting her diabetes.

The simulations were screened in an Oculus Go (Facebook Technologies) head-mounted display so that participants could turn their head and body in any direction and gather relevant information, much as if they were present in the actual location. Observant participants could notice subtle details, such as her surroundings, the condition of her home, or other activities co-occurring in the space. With traditionally shot films, this information would be presented in a close-up or with camera movement to call a viewer’s attention to relevant information, resulting in a more passive and guided viewing experience. Presenting the content in cine-VR creates an active viewing experience, with the viewer choosing what they want to watch and pay attention to, which increases immersion and encourages intellectual and emotional engagement. Viewers feel a sense of accomplishment as they notice subtle details planted by the filmmaking team, heightening the experience.

The fifth and sixth simulations of each series were what we called guided simulations, a prerecorded, cine-VR face-to-face conversation with Lula Mae’s provider and Lula Mae herself. Screened in a headset, these normally awkward, high-stakes conversations give the participants a chance to practice without the pressures of being watched or failing. Participants are encouraged to speak predetermined dialogue to a character in the headset and hear them respond. All of the cine-VR simulations were initiated simultaneously from a central computer, urging everyone in the room to say the same words at the same time, thereby reducing the potential for users to feel awkward about speaking aloud in public.

**Measures**

In addition to sociodemographic factors (ie, age, sex, race or ethnicity, occupation, years in practice, health care sector, percentage of Medicaid patients, and type of Medicaid patients), participants completed the following measures.

**Transcultural Self-Efficacy Tool–Multidisciplinary Healthcare Provider**

The Transcultural Self-Efficacy Tool–Multidisciplinary Healthcare Provider (TSET-MHP) is an 83-item scale that assesses changes in self-efficacy for cultural knowledge, cultural practical skills, and cultural awareness [26]. This scale yields three subscales: (1) Cognitive, (2) Practical, and (3) Affective [27]. All three subscales are rated on a 10-point scale, ranging from 1 (not confident) to 10 (totally confident). The Cognitive subscale asks participants to rate their level of confidence in their knowledge of the ways cultural factors influence health care for people belonging to different cultural backgrounds. The Practical subscale asks participants to rate their level of confidence in interviewing people of different cultural backgrounds to learn about their values, beliefs, and social determinants of health. Lastly, the Affective subscale asks participants to rate their level of confidence in acceptance of similarities and differences among cultural groups. These subscales demonstrate excellent internal consistency (Cronbach α ranging from .92 to .98) [27].

**Diabetes Attitude Scale-3**

The Diabetes Attitude Scale-3 (DAS-3) [17] is a 33-item scale that measures diabetes-related attitudes with five discrete subscales: (1) Need for special training (Cronbach α=.67), (2) Seriousness of type 2 diabetes (Cronbach α=.80), (3) Value of tight glucose control (Cronbach α=.72), (4) Psychosocial impact of diabetes (Cronbach α=.65), and (5) Attitude toward patient autonomy (Cronbach α=.76). Health care professionals are asked to rate their level of agreement on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The scale demonstrates good internal consistency and high content validity [17].

**Presence Questionnaire**

The 32-item Presence Questionnaire [28] measures the subjective experience of being in a virtual environment when a person is physically situated in another. Items are rated on a 7-point scale, ranging from 1 (not at all) to 4 (somewhat) to 7 (completely). We used a subset of 15 questions from the Wittmer-Singer questionnaire and removed 17 questions that measured haptic (ie, the use of technology that simulates touch) factors because the cine-VR simulations did not involve interaction with the simulated environment. For example, we removed questions that asked participants about their ability to touch objects in the virtual environment or move around in the virtual environment (eg, “How closely were you able to examine objects?” or “How compelling was your sense of moving around inside the virtual environment?”). This revised questionnaire had four subscales: (1) Involvement (Cronbach α=.83), (2) Sensory Fidelity (Cronbach α=.75), (3) Adaptation and Immersion (Cronbach α=.46), and (4) Interface Quality (Cronbach α=.53). In addition, the research team added three questions to assess presence in the virtual environment; we labeled this fifth subscale Presence (Cronbach α=.78). We calculated our own internal consistency for each subscale using a reliability analysis. The revised 18-item questionnaire demonstrated internal consistency ranging from poor to very good.

**Data Collection**

At the training program, participants received a packet that included two copies of the informed consent form, a preassessment packet, and a postassessment packet. The principal investigator read the informed consent form to all attendees of the training program. Individuals interested in participating signed the informed consent form and placed it in the packet. The informed consent form emphasized the voluntary nature of participation and reminded participants that the study was not related to their participation in the overall training program. Participants completed a brief demographic form and the two preassessment questionnaires via pen and paper; this session lasted approximately 15 minutes. All questionnaires were prelabeled with an identification number prior to the start of the study. At the completion of the training program, participants completed three postassessment questionnaires via pen and paper; this session lasted approximately 15 minutes.
Participants with questions about the study were directed to email or call the principal investigator (EB).

**Statistical Analysis**

We assessed demographic factors using descriptive statistics and presented them as means and standard deviations or sample sizes and percentages. Chi-square tests, Fisher exact tests, independent-samples $t$ tests, and one-way analyses of variance were conducted to examine differences by age, gender, race, provider status, or percentage of Medicaid (ie, limited income and resources) patients. We performed paired $t$ tests to examine changes in TSET-MHP subscale scores and DAS-3 subscale scores before and after the cine-VR training program to assess changes in cultural self-efficacy and diabetes attitudes. In addition, we determined effect sizes using Cohen $d$ by calculating the mean difference between the pre- and postassessment responses divided by the pooled standard deviation. Finally, we calculated mean change scores for TSET-MHP subscales and DAS-3 subscales. Then, we conducted Pearson correlations with the mean change scores for each subscale and the mean subscale scores of the Presence Questionnaire. We defined statistical significance as a $P$ value less than .05 and conducted analyses in SPSS Statistics for Windows, version 26.0 (IBM Corp).

**Results**

**Overview**

A total of 76 individuals consented to participate in the study; however, 7 participants did not complete postsurveys. The final sample included 69 participants out of 76 (91% completion rate). The mean age of participants was 42.2 years (SD 13.7), 86% (59/69) identified as female, 83% (57/69) identified as White, 25% (17/69) were nurses, and 86% (59/69) were health care providers (see Table 1). Among health care providers, 72% (36/50) served more than 30% of patients with limited income and resources (ie, Medicaid) in their practice. The majority of providers cared for adult Medicaid patients (44/47, 94%), followed by 77% (30/39) who cared for older adults with Medicaid, and 69% (24/35) who cared for children with Medicaid.
Table 1. Participant demographic characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participants (N=69)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>42.2 (13.7)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>59 (86)</td>
</tr>
<tr>
<td>Male</td>
<td>10 (14)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Asian Indian</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Black</td>
<td>4 (6)</td>
</tr>
<tr>
<td>Chinese</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Hispanic or Latinx</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Other Asian</td>
<td>2 (3)</td>
</tr>
<tr>
<td>White (non-Hispanic)</td>
<td>57 (83)</td>
</tr>
<tr>
<td><strong>Occupation, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Community health worker</td>
<td>16 (23)</td>
</tr>
<tr>
<td>Dentist</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Dietitian</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Exercise physiologist</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Health care administrator or staff</td>
<td>10 (14)</td>
</tr>
<tr>
<td>Nurse</td>
<td>17 (25)</td>
</tr>
<tr>
<td>Physician</td>
<td>12 (17)</td>
</tr>
<tr>
<td>Nurse practitioner</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Pharmacist</td>
<td>4 (6)</td>
</tr>
<tr>
<td>Physical therapist</td>
<td>1 (1)</td>
</tr>
<tr>
<td><strong>Years in health care, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>7 (10)</td>
</tr>
<tr>
<td>1-5</td>
<td>15 (22)</td>
</tr>
<tr>
<td>6-10</td>
<td>6 (9)</td>
</tr>
<tr>
<td>11-15</td>
<td>3 (4)</td>
</tr>
<tr>
<td>16-20</td>
<td>5 (7)</td>
</tr>
<tr>
<td>21-25</td>
<td>14 (20)</td>
</tr>
<tr>
<td>26-30</td>
<td>4 (6)</td>
</tr>
<tr>
<td>≥31</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Not applicable</td>
<td>10 (14)</td>
</tr>
<tr>
<td><strong>Health care sector, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Health care system–affiliated clinic</td>
<td>15 (22)</td>
</tr>
<tr>
<td>Hospital</td>
<td>6 (9)</td>
</tr>
<tr>
<td>Private practice</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Federally Qualified Health Center</td>
<td>4 (6)</td>
</tr>
<tr>
<td>Other</td>
<td>42 (61)</td>
</tr>
<tr>
<td><strong>Percentage of Medicaid patients served (n=50(^a)), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>≤30%</td>
<td>9 (18)</td>
</tr>
<tr>
<td>&gt;30%</td>
<td>36 (72)</td>
</tr>
</tbody>
</table>
Cultural Self-Efficacy

Mean subscale scores for the TSET-MHP are presented in Table 2. Pretraining mean scores showed that the participants had the most confidence in their Affective cultural self-efficacy (mean 8.09, SD 1.19). Prior to the training, cultural self-efficacy scores did not differ by age, gender, race, provider status, or percent of Medicaid patients.

As hypothesized, we observed positive improvements in all three of the cultural self-efficacy subscales (see Table 2): Cognitive (mean change −1.29; \( t_{65} = −9.309; P < .001 \)), Practical (mean change −1.85; \( t_{65} = −7.067; P < .001 \)). We observed the largest magnitude of change with the Practical subscale, with a Cohen \( d \) of 1.16 indicating a very large effect. Following the training program, the cultural self-efficacy subscale scores did not differ by age, gender, race, provider status, or percent of Medicaid patients except for postassessment Cognitive scores. Participants who self-identified as non-White reported greater increases than White participants in postassessment Cognitive subscale scores (mean difference −0.8447; \( t_{65} = −2.021; P = .047 \)).

Table 2. Mean differences between Transcultural Self-Efficacy Tool–Multidisciplinary Healthcare Provider (TSET-MHP) subscale scores before and after the training program.

<table>
<thead>
<tr>
<th>TSET-MHP subscale</th>
<th>Presurvey score(^a), mean (SD)</th>
<th>Postsurvey score(^a), mean (SD)</th>
<th>( P ) value</th>
<th>Cohen ( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive (n=66)</td>
<td>6.77 (1.63)</td>
<td>8.06 (1.30)</td>
<td>&lt;.001</td>
<td>0.87</td>
</tr>
<tr>
<td>Practical (n=66)</td>
<td>6.15 (1.78)</td>
<td>8.00 (1.38)</td>
<td>&lt;.001</td>
<td>1.16</td>
</tr>
<tr>
<td>Affective (n=67)</td>
<td>8.09 (1.19)</td>
<td>8.82 (1.05)</td>
<td>&lt;.001</td>
<td>0.66</td>
</tr>
</tbody>
</table>

\(^a\)Items are rated on a 10-point scale, ranging from 1 (not confident) to 10 (totally confident).

Diabetes Attitudes

Mean scores for the five DAS-3 subscales are presented in Table 3. Pretraining mean scores showed that participants generally agreed with the Need for special training (mean 4.59, SD 0.38), the Seriousness of type 2 diabetes (mean 4.23, SD 0.49), the Value of tight glucose control (mean 4.10, SD 0.40), the Psychosocial impact of diabetes (mean 4.43, SD 0.43), and the Attitude toward patient autonomy (mean 4.09, SD 0.46). No differences were observed in diabetes attitudes based on age, gender, race, provider status, or percent of Medicaid patients pretraining.

As hypothesized, we observed positive improvements in all five of the diabetes attitude subscales (see Table 3): Need for special training (mean change −0.21; \( t_{67} = −6.154; P < .001 \)), Seriousness of type 2 diabetes (mean change −0.34; \( t_{67} = −8.114; P < .001 \)), Value of tight glucose control (mean change −0.13; \( t_{67} = −3.029; P = .001 \)), Psychosocial impact of diabetes (mean change −0.33; \( t_{67} = −6.610; P < .001 \)), and Attitude toward patient autonomy (mean change −0.17; \( t_{67} = −3.889; P < .001 \)). We observed the largest magnitude of change with the Psychosocial impact of diabetes subscale, with a Cohen \( d \) of 0.87 indicating a large effect. Similar to the pretraining assessment, diabetes attitudes did not differ based on age, gender, race, provider status, or percent of Medicaid patients postraining.

Table 3. Mean differences between Diabetes Attitude Scale-3 (DAS-3) subscale scores before and after the training program (n=68).

<table>
<thead>
<tr>
<th>DAS-3 subscale</th>
<th>Presurvey score(^a), mean (SD)</th>
<th>Postsurvey score(^a), mean (SD)</th>
<th>( P ) value</th>
<th>Cohen ( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for special training</td>
<td>4.59 (0.38)</td>
<td>4.81 (0.27)</td>
<td>&lt;.001</td>
<td>0.65</td>
</tr>
<tr>
<td>Seriousness of type 2 diabetes</td>
<td>4.23 (0.49)</td>
<td>4.57 (0.39)</td>
<td>&lt;.001</td>
<td>0.78</td>
</tr>
<tr>
<td>Value of tight glucose control</td>
<td>4.10 (0.40)</td>
<td>4.24 (0.43)</td>
<td>.001</td>
<td>0.32</td>
</tr>
<tr>
<td>Psychosocial impact of diabetes</td>
<td>4.43 (0.43)</td>
<td>4.75 (0.31)</td>
<td>&lt;.001</td>
<td>0.87</td>
</tr>
<tr>
<td>Attitude toward patient autonomy</td>
<td>4.09 (0.46)</td>
<td>4.26 (0.48)</td>
<td>&lt;.001</td>
<td>0.38</td>
</tr>
</tbody>
</table>

\(^a\)Items are rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).
Presence in Cinematic Virtual Reality

Following the training program, we observed mean scores greater than or equal to 5.9, out of a maximum score of 7, for all five subscales: Involvement (mean 6.22, SD 0.59), Sensory Fidelity (mean 5.90, SD 0.81), Adaptation and Immersion (mean 6.22, SD 0.61), Interface Quality (mean 5.92, SD 1.31), and Presence (mean 6.28, SD 0.70). The high subscale scores demonstrate favorable perceptions of the technology and strength of presence in the cine-VR simulations. Presence in subscale scores did not differ based on age, gender, race, provider status, or percent of Medicaid patients.

Posttraining, change scores in cultural self-efficacy and diabetes attitudes were correlated with the mean subscale scores of presence. We observed only one significant correlation between the change score in Affective self-efficacy and the Interface Quality subscale score ($r = .285$, $P = .03$). No other significant correlations were observed between presence in cine-VR subscales and cultural self-efficacy subscale scores or diabetes attitude subscale scores (see Multimedia Appendix 1). These findings did not support the hypotheses that stated that increased presence in cine-VR would be associated with positive changes in cultural self-efficacy subscales and diabetes attitude subscales.

Discussion

Principal Findings

In this pilot study, we assessed health care providers’ and administrators’ cultural self-efficacy and diabetes attitudes before and after a 360-degree cine-VR training program. Following the training program, we observed statistically significant improvements in all three cultural self-efficacy subscales: (1) Cognitive, (2) Practical, and (3) Affective. The largest magnitude of effect was observed with the Practical subscale, which corresponds to confidence in interviewing patients about social determinants of health. In addition, all five diabetes attitude subscales improved significantly posttraining: (1) Need for special training, (2) Seriousness of type 2 diabetes, (3) Value of tight glucose control, (4) Psychosocial impact of diabetes, and (5) Attitude toward patient autonomy, with the largest magnitude of change observed in Psychosocial impact of diabetes. Lastly, we observed high scores for presence in cine-VR, indicating favorable perceptions of the technology and immersion in the 360-degree virtual environment. Contrary to expectations, only one positive change score in Affective self-efficacy was correlated with increased presence in cine-VR.

Comparison With Prior Work

Effective cine-VR simulations provide a platform to practice and acquire skills that will later translate to clinical outcomes concerning patient care; in addition, they afford participants the opportunity to practice clinical judgment and apply problem-solving skills in a risk-free, replicable clinical environment [29,30]. Cine-VR technology offers new opportunities for clinical assessment and intervention. Advances in virtual reality technologies can now support the creation of low-cost, yet sophisticated, immersive simulations, capable of running on consumer-level computing devices [31]. Compared to traditional video training, the immersive qualities of cine-VR affect the participant’s ability to more strongly retrieve the experience from memory, suggesting that cine-VR experiences become part of an autobiographical associative network, whereas a conventional video experience remains an isolated episodic event [32].

Existing research in narrative health promotion demonstrates the power of culturally tailored stories as engaging content to positively affect attitudes, beliefs, and behaviors. Qualitative results show that the digital storytelling more positively affects participants than traditional face-to-face training on its own, specifically in four growth areas: truth-telling, sense-making, social support, and feeling valued [33]. Research concerning digital storytelling and its uses within health care are only in their infancy in terms of discovering applications and uses. However, recent studies demonstrate that digital stories allow for a deeper understanding of an experience rather than simply hearing an explanation of that experience [34]. Our research supports this finding. Our findings suggest that this innovative cine-VR program can be used to educate providers about type 2 diabetes, social determinants of health, and Appalachian culture, which, in turn, may enhance the delivery of high-quality, evidence-based diabetes care in rural Appalachian Ohio. Additional research is needed to determine the impact of the training on patient care and health outcomes.

Finally, presence describes the extent to which a participant feels present or immersed in a virtual environment [35,36] and is commonly regarded as a necessary mediator that allows real emotions to be activated [37,38]. We hypothesized that higher levels of presence would be associated with positive changes in cultural self-efficacy and diabetes attitudes. We observed only one significant correlation between the change score in Affective self-efficacy and the Interface Quality subscale score. This finding suggests that participants who felt less distracted by the headset or experienced fewer delays with the simulations showed a greater improvement in the Affective self-efficacy scores posttraining. We observed no other significant correlations between positive change scores and presence. This may be explained by the limited variability in presence subscale scores and the overall high level of presence measured in the study. The strength of this 360-degree cine-VR simulation training program is the realism afforded by providing the participant access to the whole environment as compared to traditional virtual reality (eg, animated environments and characters), which has been criticized as being too unrealistic [39].

Limitations

Limitations of this study include the small homogeneous sample, selection bias, social desirability bias, and lack of a control group. While a final sample of 69 participants is small, our a priori power analysis determined that a sample size of 34 paired participants was sufficient to achieve 80% power and a level of significance of $P < .05$. We successfully doubled the required sample size estimate. However, data from 69 providers and administrators from one geographic region limits the generalizability of the findings to other providers. Further, the predominantly White study sample limits the generalizability to all providers; however, the racial and ethnic distribution of...
the study sample (83% White) is reflective of the racial and ethnic distribution in southeastern Ohio (95% White) [40]. Next, our findings may be susceptible to selection bias, as individuals who volunteered to participate may have been more willing or motivated to participate in a novel educational program about diabetes, social determinants of health, and Appalachian culture. In addition, the responses may be susceptible to selection bias given the participants may have felt undue pressure to provide positive feedback on the training session. A similar susceptibility to selection bias may be prescribed to the use of new technology encouraging people to provide positive feedback. Finally, this study presents findings from a 3-hour cine-VR training program on type 2 diabetes in rural Appalachia. We did not include a control condition as a comparison group. Future research should use a randomized controlled design to assess the impact of two different educational interventions on providers’ and administrators’ cultural self-efficacy and diabetes attitudes.

**Conclusions**

Continuing medical education is an important component of clinical care for all providers. Health care providers and administrators need ongoing and repeated training to help them improve and maintain their knowledge. Stay current with the latest developments, address real-world challenges, and learn effective team management skills. Our findings support the notion that 360-degree cine-VR education is an innovative approach to improve cultural self-efficacy and diabetes attitudes among health care providers and administrators. The long-term impact of cine-VR education on cultural self-efficacy and diabetes attitudes needs to be determined.

**Acknowledgments**

This study was part of the Medicaid Equity Simulation Project funded by the Ohio Department of Medicaid and administered by the Ohio Colleges of Medicine Government Resource Center. The views expressed in this publication about the cine-VR simulations are solely those of the creators and do not represent the views of the state of Ohio or federal Medicaid programs.

**Conflicts of Interest**

None declared.

Multimedia Appendix 1

Correlations among subscale scores of presence in virtual reality, change scores in cultural self-efficacy, and Diabetes Attitude Scale-3 (DAS-3) subscales (n=65).

[DOCX File, 14 KB - diabetes_v61e23708_app1.docx ]

**References**


Abbreviations

cine-VR: virtual reality cinema
DAS-3: Diabetes Attitude Scale-3
TSET-MHP: Transcultural Self-Efficacy Tool–Multidisciplinary Healthcare Provider

©Elizabeth Ann Beverly, Carrie Love, Matthew Love, Eric Williams, John Bowditch. Originally published in JMIR Diabetes (http://diabetes.jmir.org), 27.01.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), 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.
Exchanges in a Virtual Environment for Diabetes Self-Management Education and Support: Social Network Analysis

Carlos A Pérez-Aldana1, MSc; Allison A Lewinski2,3, PhD, MPH, RN; Constance M Johnson1,4, PhD, RN, FAAN; Allison A Vorderstrasse5, DNSc, APRN, FAAN; Sahiti Myneni1, PhD, MSE

1School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, United States
2Durham Veterans Affairs Medical Center, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham, NC, United States
3Duke University School of Nursing, Durham, NC, United States
4Cizik School of Nursing, The University of Texas Health Science Center at Houston, Houston, TX, United States
5Rory Meyers School of Nursing, New York University, New York, NY, United States

Corresponding Author:
Carlos A Pérez-Aldana, MSc
School of Biomedical Informatics
The University of Texas Health Science Center at Houston
Suite 600
7000 Fannin
Houston, TX, 77030
United States
Phone: 1 713 500 3900
Email: Carlos.A.PerezAldana@uth.tmc.edu

Abstract

Background: Diabetes remains a major health problem in the United States, affecting an estimated 10.5% of the population. Diabetes self-management interventions improve diabetes knowledge, self-management behaviors, and clinical outcomes. Widespread internet connectivity facilitates the use of eHealth interventions, which positively impacts knowledge, social support, and clinical and behavioral outcomes. In particular, diabetes interventions based on virtual environments have the potential to improve diabetes self-efficacy and support, while being highly feasible and usable. However, little is known about the patterns of social interactions and support taking place within type 2 diabetes–specific virtual communities.

Objective: The objective of this study was to examine social support exchanges from a type 2 diabetes self-management education and support intervention that was delivered via a virtual environment.

Methods: Data comprised virtual environment–mediated synchronous interactions among participants and between participants and providers from an intervention for type 2 diabetes self-management education and support. Network data derived from such social interactions were used to create networks to analyze patterns of social support exchange with the lens of social network analysis. Additionally, network correlations were used to explore associations between social support networks.

Results: The findings revealed structural differences between support networks, as well as key network characteristics of supportive interactions facilitated by the intervention. Emotional and appraisal support networks are the larger, most centralized, and most active networks, suggesting that virtual communities can be good sources for these types of support. In addition, appraisal and instrumental support networks are more connected, suggesting that members of virtual communities are more likely to engage in larger group interactions where these types of support can be exchanged. Lastly, network correlations suggest that participants who exchange emotional support are likely to exchange appraisal or instrumental support, and participants who exchange appraisal support are likely to exchange instrumental support.

Conclusions: Social interaction patterns from disease-specific virtual environments can be studied using a social network analysis approach to better understand the exchange of social support. Network data can provide valuable insights into the design of novel and effective eHealth interventions given the unique opportunity virtual environments have facilitating realistic environments that are effective and sustainable, where social interactions can be leveraged to achieve diverse health goals.

(JMIR Diabetes 2021;6(1):e21611) doi:10.2196/21611
KEYWORDS
type 2 diabetes; diabetes education; self-management; social support; virtual environments; social network analysis

Introduction

Overview
Diabetes remains a major health problem in the United States, affecting an estimated 34.2 million people of all ages (about 10.5% of the country’s population) [1]. Data show that type 2 diabetes (T2D) accounts for the most diabetes burden (between 90% and 95%), and its prevalence will continue to increase [1,2]. Diabetes is a challenging chronic illness because self-management is critical to reduce and delay the onset of complications and mortality [3-6]. Several evidence-based strategies, such as diabetes self-management education (DSME) and ongoing self-management support by peers and providers, have been shown to be effective in the management of T2D [7-9]. In particular, self-management is important in T2D given that patients manage 99% of their own care [10,11]. Moreover, diabetes self-management interventions improve diabetes knowledge and self-management behaviors, in addition to clinical outcomes [12]. Despite these benefits, less than 60% of people with diabetes attend DSME and only about 7% of newly diagnosed patients with diabetes attend DSME within 12 months following their diagnosis [13-16], indicating a pressing need for the delivery of accessible DSME and ongoing self-management support interventions.

Widespread internet connectivity provides new opportunities for wider web technology access and use by patients. Internet-based interventions, also known as eHealth, can connect patients to both peers and providers to facilitate support as well as access to evidence-based information [17]. Research suggests that T2D interventions incorporating interactive, individualized, and frequent interactions among patients, educators, and providers are among the most effective approaches [9]. eHealth interventions can provide such interactions in an effective and accessible way, which otherwise would be costly and unsustainable [12]. In addition, eHealth interventions have shown positive impacts on knowledge, social support, and clinical and behavioral outcomes [18]. Johnson et al have highlighted the benefits of eHealth interventions on T2D management, such as increased support, self-efficacy, and knowledge; improvements in glycemic levels and self-management behaviors; and efficient use of primary care services [12]. Furthermore, successful eHealth programs focused on DSME provided relevant content, engaging interactive elements, personalized learning experiences, and self-assessment tools for monitoring and feedback [17-20]. However, in spite of the potential benefits eHealth offers for DSME, eHealth interventions have been mostly based on traditional website formats. Such website formats generally lack realistic simulated environments where DSME actually takes place, such as patient community places (eg, grocery stores and restaurants) [7,21].

Virtual Environments and Diabetes Self-Management Education and Support

Virtual environments offer an effective way to provide patients with realistic settings for the acquisition and application of knowledge in community settings where daily T2D self-management takes place, while addressing barriers such as transportation, cost, time, and scheduling issues [22]. In addition, virtual environments have started to show a potential to improve diabetes self-efficacy and social support, while being highly feasible and usable [12]. Second Life (Linden Lab), a highly popular virtual world, has been shown to be an effective tool that can lead to “significant learning gains” [23]. Second Life allows users to socialize and behave in a similar way as they would naturally do in normal settings through virtual human representations known as avatars [24]. Furthermore, virtual environments, such as Second Life, offer the potential for users to perform behaviors within realistic scenarios by providing them with presence, immersion, and social interaction, while facilitating communication between patients, educators, and providers [12,24]. While virtual environments have been used to deliver health information, education, social support, and social networking, most Second Life–based health sites to date have focused on disseminating information and offering support groups [24].

Self-management diabetes interventions based on virtual environments enable diabetes education, the development of new skills, and the exchange of peer support in synchronous and asynchronous ways [7]. The Second Life Impacts Diabetes Education & Self-Management (SLIDES) virtual community was among the first interventions aimed at providing DSME and support using Second Life [24]. The results of SLIDES showed improvements in diabetes self-efficacy, social support, and foot care, as well as trends toward improvements in diet, weight loss, and clinical outcomes, while being highly feasible and usable [12]. The development of the SLIDES platform, as well as its preliminary effects, is described elsewhere [12,24]. Virtual environments, such as SLIDES, are innovative ways to provide accessible DSME and ongoing self-management support. A key characteristic of these environments is the potential for participants to develop real-world skills via simulation and rehearsal within the virtual environment that can be transferable and thus affect behaviors in the real world [12].

Another significant characteristic of virtual environments is the facilitation of social support among participants [12,24]. Social support is generally described as “an exchange of resources between at least two persons aimed at increasing the wellbeing of the receiver” [25-27]. Social support is recognized as a key component of diabetes self-management, in addition to adequate skills and behavioral development [22,28,29]. Studies have shown that social support is commonly provided through social interactions to achieve health outcomes [30,31]. Moreover, research suggests that people with T2D can benefit from frequent and sustained social interactions among peers and providers by obtaining education and support [28,32-34]. In addition, T2D interventions that are based on virtual environments can provide realistic, personalized, and ongoing interaction and support that assist participants in health care decision making [7,12,34-36]. SLIDES showed that virtual
environment-mediated interactions resemble physical ones; therefore, patients with T2D are presented with the possibility of greatly improving their access to social support [12,34]. However, the social networks highlighting the patterns of interactions within T2D-specific virtual communities, such as SLIDES, have not been studied. While the prominent effects of social relationships on health decisions and related behavior changes have been established [37,38], little is known about social interactions and the exchange of support in disease-specific virtual environments.

**Social Network Analysis and Online Health Communities**

The study of social networks provides researchers with a unique opportunity to get an in-depth view and a better understanding of the structure of online communities [38,39]. Social network research has shown that social connections (i.e., peers, family members, etc) disseminate health information, provide social support, and influence health behaviors [38,39]. Social network analysis (SNA) has been used to study the ways in which social connections can influence individuals’ attitudes, beliefs, and behaviors. Such network influences can be caused by the network environment, the position an individual occupies in the network, or structural or network-level properties [38,39]. For example, being central in a social network determines a high importance for information dissemination. Similarly, individuals located on a network’s periphery, known as peripheral individuals, can act as bridges connecting otherwise disconnected groups, thus enabling collective actions. Peripheral individuals are characterized by having one or few connections on the outside of a network and thus participating infrequently. Moreover, peripheral individuals are usually free from social norms and constraints, and thus, innovation can occur [38,39]. Furthermore, network structural properties, such as clustering, can help to identify highly connected groups of individuals, where behavior change can be accelerated. Lastly, densely connected networks have been shown to generate faster diffusion and increased coordinated action [38,39].

SNA is increasingly becoming useful to the study of online health communities owing to the exponential growth in the use of electronic communications [40]. The massive amounts of social interactions taking place within online communities today are providing researchers with valuable network data. Research has focused on the analysis of online social interactions from both general purpose social media platforms (e.g., Twitter and YouTube) and health care–specific platforms (e.g., American Diabetes Association online community) [41-44]. Often, qualitative analysis and computational text analysis are used to analyze social media interactions [41-43]. Studies have shown that SNA provides insights into social influence, information dissemination, and behavioral diffusion [39,40,45,46]. On one hand, communication structure (who communicates with whom) is key for the study of peer influence on health behaviors [40]. On the other hand, analyses of the structures of online peer-to-peer communications provide valuable insights into opinion leaders [40,45,47]. Both approaches have the potential to help researchers model effective network data–based interventions [40]. Similarly, social support exchange patterns within disease-specific virtual communities, such as SLIDES, can be studied using a SNA approach, which would allow the visualization and description of communication structures, peer influences, and behavioral diffusion, as well as the impact on health outcomes, such as blood glucose levels, for patients with diabetes [45-50]. However, despite the benefits SNA offers, to our knowledge, social interactions occurring within virtual environments have not been studied using this approach. In this study, a secondary data analysis of SLIDES social interactions through the SNA lens was carried out to examine social support exchange patterns between participants and providers [12,24,34].

**Research Aims**

The overall goal of our study was to examine social support exchanges from a T2D self-management education and support intervention (SLIDES) that was delivered via a virtual environment. The specific aims of our study were as follows: (1) to examine patterns of social interaction and support of the SLIDES intervention by creating network structures for different types of social supports and assessing these support networks using quantitative network measures; (2) to explore the associations between social support network structures by correlating them with each other using the quadratic assignment procedure (QAP); and (3) to provide insights into the exchange of social support within a disease-specific virtual environment.

**Methods**

**SNA Methodology**

**Social Network Data**

SLIDES social interaction data were used for our study [34]. SLIDES included a total sample of 24 individuals, with 20 participants and 4 providers (including diabetes educators and moderators). Detailed participant demographics are described elsewhere [12]. SLIDES facilitated virtual interactions among participants with T2D and providers in the following two types of sessions: education and support. Education sessions were held twice a week, and support sessions were held weekly. SLIDES social interactions consisted mostly of synchronous naturalistic conversations that took place throughout different locations within the virtual environment (e.g., bookstore, restaurant, and classroom) [12,24]. These conversations enabled the exchange of social support among participants and between participants and providers, and were continuously recorded and transcribed [12,24]. These transcriptions provided the data set from which network data were derived for our analysis. Detailed information on the SLIDES study site, theoretical framework, sample, measures, and outcomes have been published elsewhere [12,24]. Our analysis focused on interactions where social support was exchanged among participants and between participants and providers during a 6-month study enrollment period [34]. Study participants could log into SLIDES and participate as much or as little as they wanted and engage in synchronous conversations. Social support was defined as “personal informal advice and knowledge that help individuals initiate and sustain T2D self-management behaviors, thus increasing adherence” [22,25,27,30,34]. Social support types included emotional, instrumental, informational, and appraisal [22,25-27,29,34]. SLIDES social interactions, which were
previously characterized by the aforementioned types of social support [34,51], were used to create network structures in order to analyze social support exchange patterns at the group level (ie, participants/providers who interacted in a conversation by either listening or engaging directly, where a certain type of support was exchanged, were all linked together for that particular conversation). Thus, the unit of analysis included the tie among participants and between participants and providers who interacted via synchronous conversations, as well as the types of social support exchanged in each transcribed conversation as previously characterized [34,51].

Network Structures and Measures

Network structures were created for each type of social support by representing participants and providers as nodes and representing interactions where social support was exchanged as edges (interconnections between nodes). For each type of social support network, all edges indicating who participated in a conversation were included (ie, who interacted with whom during a virtual conversation in which social support was exchanged). Quantitative network measures were used to assess network structures across all types of social support. Network measures explain structural differences (eg, density and cohesion), as well as node importance within a network (eg, centrality) [38,39]. The following network measures were used: average degree (average number of connections of all nodes; a higher average degree number means that members of a network interacted with a higher number of members via synchronous conversations, either on a one-to-one basis or at a group level); graph density (proportion of connections relative to the total number of possible connections; ranging from 0 to 1; a higher graph density means that members of a network most likely engaged in conversations involving a higher number of members, ie, larger groups); average path length (average distance between all node dyads; the distance of a dyad is 1, which means a direct interaction between two members of the network; a higher average path length is associated with a higher distance or number of steps required for two network members to interact with each other, resulting in a less efficient network); average clustering coefficient (average measure of the interconnectivity of the node neighborhood; ranging from 0 to 1; a higher average clustering coefficient means that node neighborhoods are more interconnected, indicating conversations among a larger number of members for larger node neighborhoods); and modularity (the level of development of subcommunities within a network; ranging from −1 to 1; higher modularity values indicate higher levels of subcommunity development within a network) [38,39].

Network Statistical Analysis

Once network structures were created, we correlated them with each other to explore associations between social support network structures. The QAP was used to test network correlations. QAP is a nonparametric method based on permutations that allows testing structural similarities (correlations) between social network structures [52]. We used Gephi version 0.9.2 and UCINET version 6.685 (Analytic Technologies) to create network structures and to calculate network measures, as well as to perform correlation analysis [53,54].

Results

Network Structures

Figure 1 shows a network structure depicting all SLIDES social interactions where all types of social support were exchanged among participants and between participants and providers. Network structures for each type of social support exchanged by SLIDES participants are shown in Figure 2.
Figure 2. Network structures of Second Life Impacts Diabetes Education & Self-Management (SLIDES) social support interactions by the type of support. Node size indicates degree and node color indicates the existence of subcommunities, where larger subcommunities are shown in orange and smaller subcommunities are shown in purple and grey.

In addition, Table 1 summarizes the network measures for each social support network. As seen in Figure 2, the emotional and appraisal support networks were the most populous, with the former comprising 24 nodes and 1219 edges and the latter comprising 20 nodes and 737 edges. Moreover, the emotional and appraisal support networks had the highest average degrees (9.08 and 9.5, respectively) compared with the instrumental and informational support networks (6.0 and 3.2, respectively). This indicates that each member of these support networks interacted on average with nine other members via synchronous conversations, either on a one-to-one basis or at a group level, thus making them the most active networks. Additionally, assessment of degree at a node level showed that all support networks were somewhat centralized around a few nodes, suggesting that some members were more popular. Furthermore, the appraisal (0.5) and instrumental (0.43) support networks were the densest, suggesting that members of these networks most likely engaged in conversations involving a higher number of members (ie, larger groups), where some participants directly exchanged appraisal and/or instrumental support, while other members of the group had a latent exposure to this support.
Table 1. Summary of social network metrics for Second Life Impacts Diabetes Education & Self-Management (SLIDES) social support networks.

<table>
<thead>
<tr>
<th>Social support network</th>
<th>Average degree</th>
<th>Graph density</th>
<th>Average path length</th>
<th>Clustering coefficient</th>
<th>Modularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional</td>
<td>9.08</td>
<td>0.39</td>
<td>1.74</td>
<td>0.73</td>
<td>0.11</td>
</tr>
<tr>
<td>Instrumental</td>
<td>6.0</td>
<td>0.43</td>
<td>1.62</td>
<td>0.76</td>
<td>0.12</td>
</tr>
<tr>
<td>Informational</td>
<td>3.2</td>
<td>0.35</td>
<td>1.98</td>
<td>0.57</td>
<td>0.46</td>
</tr>
<tr>
<td>Appraisal</td>
<td>9.5</td>
<td>0.5</td>
<td>1.52</td>
<td>0.72</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Additionally, no substantial differences were observed between all average path length values. However, the appraisal (1.52) and instrumental (1.62) support networks had a slightly lower average path length compared with the emotional (1.74) and informational (1.98) support networks. This indicates that the distance or number of steps needed for members of these networks to interact with each other required on average fewer steps to exchange the supports, thus making these networks more efficient. In terms of network structure and community development, on one hand, the instrumental, emotional, and appraisal support networks had higher average clustering coefficients (76%, 73%, and 72%, respectively) compared with the informational support network (57%). These results indicate high levels of interconnectivity within these support networks. On the other hand, the modularity values of the emotional (0.11), appraisal (0.12), and instrumental (0.12) support networks were lower compared with that of the informational (0.46) support network. This indicates that subcommunities of network members exchanging informational support reached higher levels of development in comparison with subcommunities from all other support networks.

Lastly, Figure 3 illustrates a two-mode network representing the affiliation between participants and providers, and the types of social support exchanged via social interactions. As seen in Figure 3, according to degree, the two-mode network is centralized around emotional and appraisal support, indicating that a higher number of participants and providers participated in interactions where these types of support were exchanged (either directly or indirectly having a latent exposure as previously discussed). Moreover, a subgroup of participants and providers engaged more frequently in interactions where emotional support and appraisal support were exchanged, which are represented by thicker edges.

Figure 3. Two-mode network structure of social interactions for all types of support. The shape of the nodes distinguishes two sets of nodes as follows: squares represent participants and providers, and circles represent types of social support. In addition, the color of the circles represents each type of social support (orange, purple, yellow, and blue representing emotional, appraisal, informational, and instrumental support, respectively). Finally, the size of the circles indicates degree, and edge thickness represents the frequency of participants’ interactions within each type of support.

Network Statistical Analysis

Table 2 shows network correlation scores obtained by QAP analysis. All social support networks were correlated with one another. QAP correlation scores between the emotional and appraisal, instrumental and appraisal, and instrumental and emotional support networks were much stronger when compared with the correlations between the informational and appraisal, informational and emotional, and instrumental and informational support networks. The stronger correlation scores suggest that considerable similarities exist between the aforementioned social support networks.
Table 2. Network correlation test results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Appraisal</th>
<th>Emotional</th>
<th>Informational</th>
<th>Instrumental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>1</td>
<td>0.974</td>
<td>0.344</td>
<td>0.833</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>.004</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Emotional</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>0.974</td>
<td>1</td>
<td>0.318</td>
<td>0.818</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;.001</td>
<td>—</td>
<td>.003</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Informational</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>0.344</td>
<td>0.318</td>
<td>1</td>
<td>0.204</td>
</tr>
<tr>
<td>P value</td>
<td>.004</td>
<td>.003</td>
<td>—</td>
<td>.02</td>
</tr>
<tr>
<td>Instrumental</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>0.833</td>
<td>0.818</td>
<td>0.204</td>
<td>1</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>.02</td>
<td>—</td>
</tr>
</tbody>
</table>

*Not applicable.

Discussion

Principal Findings

In this study, we used SNA to examine patterns of social interactions and support of SLIDES, an intervention for T2D self-management education and support that was delivered via a virtual environment [12,24]. To the best of our knowledge, this study is among the first to explore the patterns of social interactions of a disease-specific virtual environment. This novel approach provided insights into the exchange of social support within the SLIDES virtual community. Our findings indicate that emotional and appraisal support networks were the largest, most centralized, and most active, indicating that a virtual community with a larger number of members can be more supportive. Moreover, a higher centralization indicated that some network members were more active, which suggests that a virtual community benefits from having active members, such as educators and moderators, because they can help engage the community. This is important for the design of interventions based on virtual environments. For example, interventions could recruit diabetes moderators or leaders to act as peer influencers or change agents. Moreover, appraisal and instrumental support networks are more connected than emotional and informational support networks. This suggests that more members are likely to engage in larger group synchronous conversations, thus indicating that well-connected networks can facilitate the exchange of appraisal and instrumental support within virtual communities. This finding could be leveraged when designing interventions that facilitate the exchange of appraisal and/or instrumental support.

An analysis of the structures of the support networks revealed higher levels of interconnectivity within the instrumental, emotional, and appraisal support networks, as indicated by their higher average clustering coefficients. Clustering can accelerate information and behavior spread [38,39], thus suggesting that interventions based on virtual environments can leverage this characteristic to accelerate the exchange of social support. Despite high degrees of clustering, instrumental, emotional, and appraisal support networks had low modularity values, indicating low levels of subcommunity development. In contrast, the informational support network showed a higher level of subcommunity development. From an intervention’s perspective, subcommunities or groups within informational support networks can be leveraged to spread resources and behaviors, in addition to providing informational support. Studies have shown that groups have norms and exert social pressure, enabling behavior change, as well as more opportunities to access information, resources, and support [39].

Our findings also show that a higher number of participants and providers participated in interactions where emotional support and appraisal support were exchanged, and they did so more frequently. These findings diverge from a previous analysis by Lewinski et al, where informational support and emotional support were the most commonly exchanged types of support among participants and between participants and providers, and appraisal support exchange was lower [34]. Their analysis focused on support exchanges at a dyadic level in order to characterize interactions. In contrast, our analysis focused on support exchanges at a group level, as previously indicated. In other words, a dyadic analysis for two participants who interact in a group conversation would identify the frequency of support exchanged between those two participants. On the other hand, our network approach to this same scenario would take into account the connections between all participants who engaged in the conversation, including those who actively engaged one another to exchange support, as well as the other participants who engaged passively and had a latent exposure. Taking this into account, we hypothesize that a higher and more frequent engagement in interactions where emotional and appraisal support were exchanged was caused by the role providers, specifically diabetes educators, played assisting in the self-management of diabetes.

Lastly, network correlations showed that all social support networks were correlated with one another. Specifically, stronger...
correlation scores for emotional and appraisal, instrumental and appraisal, and instrumental and emotional support networks indicate that considerable similarities exist between these networks. These results suggest that SLIDES participants who exchanged emotional support were likely to exchange appraisal or instrumental support. Likewise, participants who exchanged appraisal support were likely to exchange instrumental support. From an intervention’s perspective, educators and moderators from virtual communities can leverage interactions where a certain type of support is exchanged in order to maximize the provision of advice and support among members of such communities. For example, by promoting interactions between members where emotional support is exchanged, further discussion and opportunities could be created that would most likely prompt exchange of appraisal or instrumental support [34,55,56]. As a result, a higher number of supportive relationships would be fostered among participants and providers, increasing the effectiveness of support networks and thus substantiating the value of virtual communities for diabetes self-management and other health goals.

Limitations

There are several limitations in this study. The small sample size of the SLIDES study (N=24) created a small virtual community, which consequently resulted in a small community. The social dynamics resulting from a small community might differ from larger ones, which suggests that our findings should be interpreted with caution. The creation of social networks from interactions, where some type of social support was exchanged, was considered at a group conversational level and not at a dyadic level. This resulted in group identification of social support interactions, meaning that a type of social support was assigned to all group participants interacting in a conversation where social support occurred during a particular conversation. Future studies could improve network creation by analyzing participants’ interactions at a dyadic level so that social support exchanges describe social ties at a dyadic level, thus providing more accurate social support dynamics. Despite these limitations, we consider these findings valuable because of the insights provided into social support exchanges within disease-specific virtual environments.

Conclusions

This study described the utility of SNA to examine social support in a DSME virtual environment. Our findings have revealed structural differences between support networks, as well as key network characteristics of supportive interactions facilitated by the virtual community, with emotional and appraisal networks being large, centralized, and most active, thus emphasizing the value of virtual environments as sources of these two support types for T2D patients. In addition, support networks have highlighted the benefits central members, such as educators and moderators, can contribute by facilitating community engagement. Specifically, educators and moderators from the SLIDES intervention have facilitated community engagement by leading weekly synchronous group meetings that include educational sessions, focusing on core American Diabetes Association/American Association of Diabetes Education self-management curriculum, as well as support sessions [12].

Furthermore, our appraisal and instrumental support networks suggest that members of virtual communities are more likely to engage in larger group interactions where these types of support can be exchanged, with the caveat that some members can engage one another to actively exchange support, while the other members engage passively and have a latent exposure to support exchange. Lastly, our network correlation analysis has shown that participants who exchange emotional support are likely to exchange appraisal or instrumental support, and participants who exchange appraisal support are likely to exchange instrumental support. These associations suggest that interactions, where a certain type of support is exchanged, could be leveraged to maximize the provision of advice and support among network members, thus increasing the effectiveness of support networks enabled by virtual communities.

Network data can provide valuable insights into the design of novel and effective digital health interventions given the unique opportunity disease-specific virtual environments have facilitating realistic environments that are effective and sustainable, where social interactions can be leveraged to achieve diverse health goals.

Acknowledgments

Data in this study were obtained in the following grants: F31-NR016622-01 (principal investigator [PI]: Lewinski) funded by the National Institutes of Health, National Institute for Nursing Research and 1R21LM010727-01 (PI: Johnson) funded by the National Library of Medicine. Support for Dr Lewinski was provided by the VA Office of Academic Affiliations (TPH 21-000), and publication support was provided by Durham VA Health Services Research Center of Innovation funding (CIN 13-410). Part of the research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health under Award Number 1R01LM012974-01A1. The findings and conclusions in this document are those of the authors who are responsible for its contents and do not represent the views of the Department of Veterans Affairs or the National Institutes of Health; therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs.

Conflicts of Interest

AAL reports receiving funds from PhRMA Foundation and Otsuka. Other authors have no conflicts to declare.

References

http://diabetes.jmir.org/2021/1/e21611/


54. Borgatti SP, Everett MG, Freeman LC. Ucinet for Windows: Software for Social Network Analysis. 2002 Dec 01. URL: https://sites.google.com/site/ucinetsoftware/home [accessed 2020-12-01]


**Abbreviations**

- **DSME**: diabetes self-management education
- **QAP**: quadratic assignment procedure
- **SLIDES**: Second Life Impacts Diabetes Education & Self-Management
- **SNA**: social network analysis
- **T2D**: type 2 diabetes

**Please cite as:**


URL: http://diabetes.jmir.org/2021/1/e21611/

doi:10.2196/21611

PMID:33492236

©Carlos A Pérez-Aldana, Allison A Lewinski, Constance M Johnson, Allison A Vorderstrasse, Sahiti Myneni. Originally published in JMIR Diabetes (http://diabetes.jmir.org), 25.01.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), 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.
Ability of Current Machine Learning Algorithms to Predict and Detect Hypoglycemia in Patients With Diabetes Mellitus: Meta-analysis

Abstract

Background: Machine learning (ML) algorithms have been widely introduced to diabetes research including those for the identification of hypoglycemia.

Objective: The objective of this meta-analysis is to assess the current ability of ML algorithms to detect hypoglycemia (ie, alert to hypoglycemia coinciding with its symptoms) or predict hypoglycemia (ie, alert to hypoglycemia before its symptoms have occurred).

Methods: Electronic literature searches (from January 1, 1950, to September 14, 2020) were conducted using the Dialog platform that covers 96 databases of peer-reviewed literature. Included studies had to train the ML algorithm in order to build a model to detect or predict hypoglycemia and test its performance. The set of 2×2 data (ie, number of true positives, false positives, true negatives, and false negatives) was pooled with a hierarchical summary receiver operating characteristic model.

Results: A total of 33 studies (14 studies for detecting hypoglycemia and 19 studies for predicting hypoglycemia) were eligible. For detection of hypoglycemia, pooled estimates (95% CI) of sensitivity, specificity, positive likelihood ratio (PLR), and negative likelihood ratio (NLR) were 0.79 (0.75-0.83), 0.80 (0.64-0.91), 8.05 (4.79-13.51), and 0.18 (0.12-0.27), respectively. For prediction of hypoglycemia, pooled estimates (95% CI) were 0.80 (0.72-0.86) for sensitivity, 0.92 (0.87-0.96) for specificity, 10.42 (5.82-18.65) for PLR, and 0.22 (0.15-0.31) for NLR.

Conclusions: Current ML algorithms have insufficient ability to detect ongoing hypoglycemia and considerate ability to predict impeding hypoglycemia in patients with diabetes mellitus using hypoglycemic drugs with regard to diagnostic tests in accordance with the Users’ Guide to Medical Literature (PLR should be ≥5 and NLR should be ≤0.2 for moderate reliability). However, it should be emphasized that the clinical applicability of these ML algorithms should be evaluated according to patients’ risk profiles such as for hypoglycemia and its associated complications (eg, arrhythmia, neuroglycopenia) as well as the average ability of the ML algorithms. Continued research is required to develop more accurate ML algorithms than those that currently exist and to enhance the feasibility of applying ML in clinical settings.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020163682; http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42020163682

https://doi.org/10.2196/22458
KEYWORDS
machine learning; hypoglycemia; meta-analysis

Introduction

Hypoglycemia is a major barrier to achieving the tight glycemic control in patients with diabetes mellitus (DM) that is required to delay the progression of late DM-related complications. Although many patients exhibit symptoms of hypoglycemia such as anxiety, heart palpitations, and confusion, a significant number have diminished ability to recognize these hypoglycemic symptoms [1,2], which is defined as “impaired awareness of hypoglycemia” [3]. This impaired awareness can lead to severe hypoglycemia, which is associated with seizures, coma, and death. Real-time glucose monitoring can help patients maintain optimal glycemic control while avoiding symptomatic or asymptomatic hypoglycemia [4]. However, the traditional monitoring method, intermittent glucose monitoring by finger stick, provides only a limited number of readings and is unlikely to detect hypoglycemia of a short duration. Continuous glucose monitoring (CGM) typically produces a reading every 5 minutes and can alert the patient to not only the occurrence of hypoglycemia but also impending hypoglycemia [5]. Accuracy of CGM has progressively improved, with overall measurement errors reduced by twofold than in the first commercially available CGM devices introduced in 2000 [5].

However, even if CGM advancements enabled patients to continuously track their subcutaneous glucose levels, the statistical disadvantage of the CGM data stream would remain as a major limitation. The autocorrelation of the CGM reading vanishes after 30 minutes, meaning that the projection of blood glucose levels more than 30 minutes ahead would be inaccurate [6]. This finding suggests that the algorithm for identifying hypoglycemia should consider a patient’s contextual information such as diet, physical activity, and medications (including insulin) as well as various features of the CGM trend arrow [7].

Machine learning (ML) algorithms have been widely introduced to diabetes research including those for identification of hypoglycemia. The growing use of mobile health (mHealth) apps, sensors, wearables, and other point-of-care devices, including CGM sensors for self-monitoring and management of DM, have made possible the generation of automated and continuous diabetes-related data and created the opportunity for applying ML to automated decision support systems [8]. Combining ML-based decision support systems with the abundance of generated data has the potential to identify hypoglycemia with greater accuracy.

Conventionally, ML has been applied to detect abnormalities in blood glucose levels using physiological parameters that are highly correlated with hypoglycemia (eg, changes in brain or cardiac electrical activities) [7]. Recently, in addition to the detection of hypoglycemia, ML-based decision support systems have been proposed for predicting hypoglycemia by using various historical data (eg, series of blood glucose data, other laboratory and demographic data, verbal data in medical records, or secure messages suggesting occurrence of hypoglycemic events) [8]. Despite many reports of ML algorithms for detecting or preventing hypoglycemia, their abilities have not been comprehensively or quantitatively assessed. This meta-analysis aims to assess the current ability of ML algorithms to detect or predict hypoglycemia in patients with DM.

Methods

Protocol Registration

The study protocol has been registered in the international prospective register of systematic reviews (PROSPERO; Registration ID: CRD42020163682).

Literature Searches

We used Dialog to perform the electronic literature searches. The platform allows users to access and search 96 databases of peer-reviewed literature. Publication dates ranged from January 1, 1950, to September 14, 2020. Search terms consisted of 2 elements: (1) thesaurus and text words related to ML and (2) text words related to hypoglycemia and thesaurus terms related to glucose monitoring or blood glucose. The use of the thesaurus term was limited to 2 databases: EMBASE (EMTREE terms) and MEDLINE (MeSH terms). The above 2 elements were combined using the BOOLEAN operator “AND” (Multimedia Appendix 1). Manual searches were added to review reference lists in relevant studies. If eligible studies were obtained from the reference lists, the reference lists in those studies were also examined. Manual searches were continued until no eligible study was found in the references lists.

Study inclusion criteria were (1) all participants had DM; (2) study endpoint was hypoglycemia; (3) researchers clarified that they originally trained the ML algorithm using training data to build a model for detecting or predicting hypoglycemia or the same researchers trained the ML algorithm in a previous study; (4) the model’s performance was tested using the test data; and (5) sensitivity and specificity for detection or prediction of hypoglycemia were presented or could be calculated.

Exclusion criteria were (1) an event-based study (ie, specificity could not be estimated because nonhypoglycemia data were not included in the test data), (2) a case study (ie, training and test data were derived from only 1 patient), and (3) a 2 × 2 contingency table consisting of the number of true positives, false positives, false negatives, and false positives could not be reproduced. If studies met all of the inclusion criteria but did not allow the reproduction of a 2 × 2 contingency table, we asked the corresponding author of these studies for the total number of test data sets (N-total) and events (N-hypo) so that we could reproduce the 2 × 2 table. If the same test data were shared by 2 or more eligible studies, we chose the most updated study in which the ML algorithm was considered to show the best performance.

The outcome of meta-analyses of diagnostic or prognostic tests is the extent of consistency between an index test and a reference standard. The index test is defined as a new test that is proposed when the method for perfectly diagnosing a target condition in...
all individuals does not exist or cannot be used. In this meta-analysis, it corresponded to an ML algorithm that classified the input data as either hypoglycemia or nonhypoglycemia. The reference standard is defined by a procedure that is considered the best available method for categorizing participants into having or not having a target condition. In this meta-analysis, it corresponded to methods for diagnosing hypoglycemia in clinical practice, which included measurement of glucose levels, the International Classification of Diseases (ICD) code for hypoglycemia, or experts’ subjective judgment. Evaluation item was the ability of ML algorithms to detect hypoglycemia (ie, alert to hypoglycemia coinciding with its symptoms) or the ability to predict hypoglycemia (ie, alert to hypoglycemia before its symptoms have occurred). In studies that assessed the ability for detection, data used for the index test (ie, the ML algorithm) and data used for a reference standard (ie, diagnosing hypoglycemia) had to be examined at the same time. In studies assessing predictive ability, the data input into the ML algorithm had to be examined before the diagnosis of hypoglycemia.

Data Extraction
Data were extracted by two authors (SK and KF). Disagreements were resolved by discussion with a third author (HiS). We fundamentally selected 1 datum if there were 2 or more extractable data for a set of test data in an individual study. If an individual study tested 2 or more ML classification methods or 2 or more models for 1 ML classifier, we extracted the datum related to the classifier or model that the study proposed as the best. If 2 or more different results were presented for the same model depending on the prediction window or horizon, we extracted data on the result in relation to the longest prediction window or horizon.

The following study characteristics were extracted: first author, publication year, evaluated item (ie, detecting or predicting hypoglycemia), country, type of DM (ie, type 1 or type 2), number of study participants, N-total, N-hypo, mean or range of the patients’ age, time of day of hypoglycemic events, place of supposed hypoglycemic episode (ie, experimental, in-hospital, and out-of-hospital), ML algorithm used for classification into hypoglycemia and nonhypoglycemia, threshold of glucose level for hypoglycemia, method for diagnosing hypoglycemia, method for separating the database into training and test data, and profiling data that were input into ML algorithms for performance testing.

Study Quality
To evaluate study quality, we used a revised tool to assess diagnostic accuracy of studies (QUADAS-2). The QUADAS-2 consists of 4 domains: selection of participants, index test, reference standard, and flow and timing. All 4 domains were used for assessment of risk of bias and the first 3 domains were used to assess the consensus of applicability. Each domain has 1 query in relation to the risk of bias or applicability consisting of 7 questions (Multimedia Appendix 2) [9]. A “Yes” answer was assigned 1 point.

Data Synthesis
The ability of ML algorithms to detect hypoglycemia and predict hypoglycemia was independently assessed. For data that were used to test the model’s performance, the number of true positives, false positives, true negatives, and false negatives was calculated. The set of 4 data was pooled with a hierarchical summary receiver operating characteristic (HSROC) model [10]. Indicators for the model’s performance included sensitivity, specificity, positive likelihood ratio (PLR), which is calculated as (sensitivity/[1–specificity]), and negative likelihood ratio (NLR), which is calculated as ([1–sensitivity]/specificity). Study heterogeneity was assessed by calculating I² values for PLR and NLR based on a multivariate random-effects meta-regression that considered within- and between-study correlations [11] and classifying them into quartiles (0% to <25%, low; 25% to <50%, low-to-moderate; 50% to <75%, moderate-to-high; >75%, high) [12]. Publication bias was statistically assessed as proposed by Deeks et al [13], wherein the logarithm of the diagnostic odds ratio is regressed against its corresponding inverse of the square root of the effective sample size.

Sensitivity analyses were added, and the analysis was limited to studies that shared similar characteristics in terms of the type of DM, time of day when hypoglycemia occurred, place of supposed hypoglycemic events, and the profiling data input into the ML algorithm. It is of note that at least four data sets are necessary to perform these sensitivity analyses because the HSROC model has 4 parameters: sensitivity, specificity, accuracy, and threshold. A two-sided P-value <.05 was considered statistically significant. All statistical analyses were performed using Stata 16 (StataCorp).

Results

Literature Searches
Multimedia Appendix 3 shows the flow chart of the procedure for selecting studies. Using prespecified search terms, 1226 articles were retrieved; 61 databases published at least one of the retrieved articles (Multimedia Appendix 4). Of these 1226 articles, 150 studies were selected for further review. Manual searches resulted in the addition of 32 studies for further review, making a total of 182 studies. Of these, 149 studies were subsequently excluded for various reasons. Specifically, 12 studies [14-25] presented insufficient data to allow reproduction of the 2 × 2 contingency table, although data on sensitivity and specificity were presented. We asked the authors of these studies to provide N-totals and N-hypos so that we could calculate the number of true positives, false positives, true negatives, and false negatives. However, only the author of 2 studies responded to our communication [15,25], and therefore the remaining 10 studies with insufficient data had to be excluded from the meta-analysis. Finally, 33 studies [15,20,25-55] were eligible.

Data Extraction of Study Characteristics
Table 1 shows the summary of study characteristics. Of the 33 studies, 19 studies (58%) [26-31,33,35,36,38,42,44,47,54] predicted hypoglycemia, and the remaining 14 studies (42%) detected hypoglycemia [15,20,25,32,34,37,43,48-53,55]. As much as 25 of the 33 included studies (76%) [15,20,25,27,29,30,32,35,36,38,39,41-44,46,53,55] specified type 1 as the type of DM. Type 2 DM was specified in only 3...
of these studies (9%) [28,31,45] and the remaining 5 studies [33,34,37,40,54] did not specify the type of DM.

Regarding the time of day when hypoglycemic events occurred, nocturnal hypoglycemia was the most frequently reported (14 studies of the 33 included studies; 42%) [15,20,26,30,32,35,36,41,44,49-53]). As to the place of the supposed hypoglycemic episode, 16 of the 19 studies that predicted hypoglycemia (84%) [26-30,35,36,38-42,44-47] supposed the event took place in an out-of-hospital setting. The remaining 3 studies (16%) [31,33,54] supposed hypoglycemia occurring in an in-hospital setting. Of the 14 studies that detected hypoglycemia, 11 studies (79%) [15,20,25,32,43,48-52,55] detected hypoglycemia in an experimental setting, where hypoglycemia was induced by a hypoglycemic clamp procedure. In 20 of the 33 included studies (61%) [15,20,25,27,29,31,32,35,36,38,41,43-45,49-52,54,55], a hold-out method was used to separate the information in the database according to training and test data. 

Multimedia Appendix 5 shows the profiling data input into the ML algorithm for testing its performance in detecting or predicting hypoglycemia. In the majority of the 19 studies for predicting hypoglycemia (13 studies; 68%) [26-30,35,36,38,40-42,46,47], historical CGM data were input into the ML algorithm while the remaining 6 studies (32%) [31,33,39,44,45,54] did not use CGM. Of the 14 studies that detected hypoglycemia using ML, 7 studies (50%) [20,25,32,49,50,52,55] used information from electroencephalograms (EEGs) and 4 studies (29%) [15,43,51,53] used results of electrocardiography (ECG).
Table 1. Study characteristics of the 33 included studies to assess the ability of machine learning to detect or predict hypoglycemia.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dave et al [27]</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>112</td>
<td>637,735</td>
<td>18,233</td>
<td>13</td>
<td>N/S</td>
<td>Out[^s]</td>
<td>RF[^w]</td>
<td>3.9</td>
<td>CGM</td>
<td>HO[^pp]</td>
</tr>
<tr>
<td>Elhadd et al [28]</td>
<td>Pre</td>
<td>Qatar</td>
<td>T2D[^p]</td>
<td>13</td>
<td>3918</td>
<td>172</td>
<td>51</td>
<td>N/S</td>
<td>Out</td>
<td>XG-Boost</td>
<td>Unclear</td>
<td>CGM</td>
<td>nCV</td>
</tr>
<tr>
<td>Marcus et al [29]</td>
<td>Pre</td>
<td>Israel</td>
<td>T1D</td>
<td>11</td>
<td>43,533</td>
<td>5264</td>
<td>18-39</td>
<td>N/S</td>
<td>Out</td>
<td>KRR[^k]</td>
<td>3.9</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Mosquera-Lopez et al [30], Test 1</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>10</td>
<td>117</td>
<td>17</td>
<td>34</td>
<td>Noc</td>
<td>Out</td>
<td>SVM</td>
<td>3.9</td>
<td>CGM</td>
<td>ExV</td>
</tr>
<tr>
<td>Mosquera-Lopez et al [30], Test 2</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>20</td>
<td>2706</td>
<td>258</td>
<td>35</td>
<td>Noc</td>
<td>Out</td>
<td>SVM</td>
<td>3.9</td>
<td>CGM</td>
<td>ExV</td>
</tr>
<tr>
<td>Mueller et al [31]</td>
<td>Pre</td>
<td>USA</td>
<td>T2D</td>
<td>453,487</td>
<td>90,687</td>
<td>2580</td>
<td>66</td>
<td>N/S</td>
<td>In[^i]</td>
<td>REFS</td>
<td>3.9</td>
<td>Blood/ICD</td>
<td>HO</td>
</tr>
<tr>
<td>Ruan et al [33]</td>
<td>Pre</td>
<td>UK</td>
<td>N/S[^o]</td>
<td>17,658</td>
<td>3276</td>
<td>703</td>
<td>66</td>
<td>N/S</td>
<td>In</td>
<td>XG-Boost</td>
<td>3.9</td>
<td>Blood</td>
<td>nCV</td>
</tr>
<tr>
<td>Rubega et al [25]</td>
<td>Dec</td>
<td>Italy</td>
<td>T1D</td>
<td>34</td>
<td>2516</td>
<td>1258</td>
<td>55</td>
<td>N/S</td>
<td>Exp[^u]</td>
<td>NN[^z]</td>
<td>3.9</td>
<td>Blood</td>
<td>HO</td>
</tr>
<tr>
<td>Chen et al [34]</td>
<td>Dec</td>
<td>USA</td>
<td>N/S</td>
<td>No data</td>
<td>300</td>
<td>11</td>
<td>No data</td>
<td>N/S</td>
<td>In</td>
<td>LR[^aa]</td>
<td>N/A[^lk]</td>
<td>Experts[^mm]</td>
<td>nCV</td>
</tr>
<tr>
<td>Guemes et al [35]</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>6</td>
<td>55</td>
<td>6</td>
<td>40-60</td>
<td>Noc</td>
<td>Out</td>
<td>SVM</td>
<td>3.9</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Jensen et al [36]</td>
<td>Pre</td>
<td>Denmark</td>
<td>T1D</td>
<td>463</td>
<td>921</td>
<td>79</td>
<td>43</td>
<td>Noc</td>
<td>Out</td>
<td>LDA[^bb]</td>
<td>3</td>
<td>Blood</td>
<td>HO</td>
</tr>
<tr>
<td>Jin et al [37]</td>
<td>Dec</td>
<td>USA</td>
<td>N/S</td>
<td>No data</td>
<td>4104</td>
<td>132</td>
<td>No data</td>
<td>N/S</td>
<td>In</td>
<td>SVM</td>
<td>N/A</td>
<td>ICD[^mn]</td>
<td>nCV</td>
</tr>
<tr>
<td>Oviedo et al [38]</td>
<td>Pre</td>
<td>Spain</td>
<td>T1D</td>
<td>10</td>
<td>1447</td>
<td>420</td>
<td>41</td>
<td>Pos[^q]</td>
<td>Out</td>
<td>SVM</td>
<td>3.9</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Reddy et al [39]</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>55</td>
<td>90</td>
<td>29</td>
<td>33</td>
<td>Ex</td>
<td>Out</td>
<td>RF</td>
<td>3.9</td>
<td>Blood</td>
<td>ExV</td>
</tr>
<tr>
<td>Seo et al [40]</td>
<td>Pre</td>
<td>Korea</td>
<td>N/S</td>
<td>104</td>
<td>7052</td>
<td>412</td>
<td>52</td>
<td>Pos</td>
<td>Out</td>
<td>RF</td>
<td>3.9</td>
<td>CGM</td>
<td>nCV</td>
</tr>
<tr>
<td>Study source</td>
<td>Assessment</td>
<td>Country</td>
<td>Type of DM</td>
<td>Patients, n</td>
<td>N-Total</td>
<td>N-Hypo</td>
<td>Mean or range of age (years)</td>
<td>Time</td>
<td>Place</td>
<td>Machine Learning</td>
<td>Threshold of Hypo detection (mmol/L)</td>
<td>Method of separation</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>---------</td>
<td>------------</td>
<td>-------------</td>
<td>---------</td>
<td>--------</td>
<td>-------------------------------</td>
<td>------</td>
<td>-------</td>
<td>------------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Arthur et al [41]</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>6</td>
<td>51</td>
<td>6</td>
<td>40-60</td>
<td>Noc</td>
<td>Out</td>
<td>ANN&lt;sup&gt;cc&lt;/sup&gt;</td>
<td>3.9</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Tofanin et al [42]</td>
<td>Pre</td>
<td>Italy</td>
<td>T1D</td>
<td>20</td>
<td>7096</td>
<td>36</td>
<td>46</td>
<td>N/S</td>
<td>Out</td>
<td>I- MPC&lt;sup&gt;dd&lt;/sup&gt;</td>
<td>3.9</td>
<td>CGM</td>
<td>ExV</td>
</tr>
<tr>
<td>Ling et al [43]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>16</td>
<td>269</td>
<td>55</td>
<td>15</td>
<td>N/S</td>
<td>Exp</td>
<td>FNN&lt;sup&gt;e&lt;/sup&gt;</td>
<td>3.3</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Sampath et al [44], DIA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Pre</td>
<td>Ukraine</td>
<td>T1D</td>
<td>34</td>
<td>150</td>
<td>40</td>
<td>18-65</td>
<td>Noc</td>
<td>Out</td>
<td>RA</td>
<td>3.9</td>
<td>Blood</td>
<td>HO</td>
</tr>
<tr>
<td>Sampath et al [44], Child&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Pre</td>
<td>Ukraine</td>
<td>T1D</td>
<td>179</td>
<td>476</td>
<td>222</td>
<td>3-16</td>
<td>Noc</td>
<td>Out</td>
<td>RA&lt;sup&gt;ff&lt;/sup&gt;</td>
<td>3.9</td>
<td>Blood</td>
<td>ExV</td>
</tr>
<tr>
<td>Sudharsan et al [45]</td>
<td>Pre</td>
<td>USA</td>
<td>T2D</td>
<td>Unclear</td>
<td>839</td>
<td>428</td>
<td>No data</td>
<td>N/S</td>
<td>Out</td>
<td>RF</td>
<td>3.9</td>
<td>Blood</td>
<td>HO</td>
</tr>
<tr>
<td>Eljil [46]</td>
<td>Pre</td>
<td>UAE</td>
<td>T1D</td>
<td>10</td>
<td>667</td>
<td>100</td>
<td>25</td>
<td>N/S</td>
<td>Out</td>
<td>BAG&lt;sup&gt;gg&lt;/sup&gt;</td>
<td>3.3</td>
<td>CGM</td>
<td>nCV</td>
</tr>
<tr>
<td>Plis et al [47]</td>
<td>Pre</td>
<td>USA</td>
<td>T1D</td>
<td>2</td>
<td>5816</td>
<td>152</td>
<td>No data</td>
<td>N/S</td>
<td>Out</td>
<td>SVM</td>
<td>3.9</td>
<td>CGM</td>
<td>ExV</td>
</tr>
<tr>
<td>Jensen et al [48]</td>
<td>Dec</td>
<td>Denmark</td>
<td>T1D</td>
<td>10</td>
<td>1267</td>
<td>160</td>
<td>44</td>
<td>N/S</td>
<td>Exp</td>
<td>SEP-COR&lt;sup&gt;hh&lt;/sup&gt;</td>
<td>3.9</td>
<td>Blood</td>
<td>LOO&lt;sup&gt;ll&lt;/sup&gt;</td>
</tr>
<tr>
<td>Jensen et al [48]</td>
<td>Dec</td>
<td>Denmark</td>
<td>T1D</td>
<td>10</td>
<td>1267</td>
<td>160</td>
<td>44</td>
<td>N/S</td>
<td>Exp</td>
<td>+ SVM</td>
<td>3.9</td>
<td>Blood</td>
<td>LOO</td>
</tr>
<tr>
<td>Nguyen et al [49]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>5</td>
<td>144</td>
<td>76</td>
<td>12-18</td>
<td>Noc</td>
<td>Exp</td>
<td>FNN</td>
<td>3.3</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Nguyen et al [50]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>5</td>
<td>44</td>
<td>20</td>
<td>12-18</td>
<td>Noc</td>
<td>Exp</td>
<td>ANN</td>
<td>3.3</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Nuryani et al [51]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>5</td>
<td>575</td>
<td>133</td>
<td>16</td>
<td>Noc</td>
<td>Exp</td>
<td>PSO&lt;sup&gt;jj&lt;/sup&gt; + SVM</td>
<td>Unclear</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Chan et al [15]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>16</td>
<td>100</td>
<td>52</td>
<td>15</td>
<td>Noc</td>
<td>Exp</td>
<td>FNN</td>
<td>3.3</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Ling et al [52]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>5</td>
<td>27</td>
<td>8</td>
<td>16</td>
<td>Noc</td>
<td>Exp</td>
<td>Fuzzy SVM</td>
<td>3.3</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Nguyen and Jones [20]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>6</td>
<td>79</td>
<td>27</td>
<td>12-18</td>
<td>Noc</td>
<td>Exp</td>
<td>BNN</td>
<td>3.3</td>
<td>Blood</td>
<td>HO</td>
</tr>
<tr>
<td>Skladnev et al [53]</td>
<td>Dec</td>
<td>Australia</td>
<td>T1D</td>
<td>52</td>
<td>52</td>
<td>11</td>
<td>16</td>
<td>Noc</td>
<td>In</td>
<td>FNN</td>
<td>3.9</td>
<td>Blood</td>
<td>ExV</td>
</tr>
<tr>
<td>Zhang et al [54]</td>
<td>Pre</td>
<td>USA</td>
<td>N/S</td>
<td>1004</td>
<td>1114</td>
<td>556</td>
<td>No data</td>
<td>N/S</td>
<td>In</td>
<td>DT&lt;sup&gt;jj&lt;/sup&gt;</td>
<td>3.3</td>
<td>CGM</td>
<td>HO</td>
</tr>
<tr>
<td>Study source</td>
<td>Assessment&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Country</td>
<td>Type of DM</td>
<td>Patients, n</td>
<td>N-total&lt;sup&gt;b&lt;/sup&gt;</td>
<td>N-hypo&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Mean or range of age (years)</td>
<td>Time&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Place&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Machine learning</td>
<td>Threshold of Hypo&lt;sup&gt;f&lt;/sup&gt; (mmol/L)</td>
<td>Method of Hypo detection&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Method of separation&lt;sup&gt;h&lt;/sup&gt;</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>---------</td>
<td>------------</td>
<td>-------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Iaione and Marques [55]</td>
<td>Dec</td>
<td>Brazil</td>
<td>T1D</td>
<td>8</td>
<td>1990</td>
<td>995</td>
<td>35</td>
<td>Mor&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Exp</td>
<td>ANN</td>
<td>3.3</td>
<td>Blood</td>
<td>HO</td>
</tr>
</tbody>
</table>

<sup>a</sup>Ability for which the machine learning algorithm was assessed.
<sup>b</sup>N-total: total number of data included in test data.
<sup>c</sup>N-hypo: total number of hypoglycemic episodes included in the test data.
<sup>d</sup>Time of day when hypoglycemia occurred.
<sup>e</sup>Place of supposed hypoglycemic episode.
<sup>f</sup>Threshold of glucose level that was used to diagnose hypoglycemia.
<sup>g</sup>Method for separating training and test data.
<sup>h</sup>Method used for diagnosing hypoglycemia.
<sup>i</sup>DIA: DIAadvisor.
<sup>j</sup>Child: ChildrenData.
<sup>k</sup>Pre: predicting hypoglycemia.
<sup>l</sup>Dec: detecting hypoglycemia.
<sup>m</sup>T1D: type 1 diabetes mellitus.
<sup>n</sup>T2D: type 2 diabetes mellitus.
<sup;o</sup>N/S: not specified.
<sup>p</sup>NOC: nocturnal hypoglycemia.
<sup>q</sup>Pos: postprandial.
<sup>r</sup>Mor: hypoglycemia during morning.
<sup>s</sup>Out: out-of-hospital setting.
<sup>t</sup>In: in-hospital setting.
<sup>u</sup>Exp: experimental setting (ie, hypoglycemia is induced by injection of insulin. Exercise or drug intervention is included in out of hospital setting).
<sup>v</sup>SVM: support vector machine.
<sup>w</sup>RF: random forest.
<sup>x</sup>KRR: Kernel Ridge Regression.
<sup>y</sup>BNN: Bayesian neural network.
<sup>z</sup>NN: neural network.
<sup>a</sup>LR: logistic regression.
<sup>b</sup>LDA: linear discriminant analysis.
<sup>c</sup>ANN: artificial neural network.
<sup>d</sup>MPC: individual model-based predictive control.
<sup>e</sup>FNN: fuzzy neural network.
<sup>f</sup>RA: ranking aggregation algorithms.
<sup>g</sup>BAG: bagging (bootstrap aggregating).
<sup>h</sup>SEPCOR: separability and correlation analysis.
<sup>i</sup>PSO: particle swarm optimization.
<sup>j</sup>DT: decision tree.
<sup>k</sup>N/A: Not applicable.
<sup>l</sup>CGM: continuous glucose monitoring.
<sup>m</sup>Experts’ subjective judgment.
<sup>n</sup>ICD: International Classification of Diseases.
<sup>o</sup>nCV: n-fold cross-validation.
<sup>p</sup>HO: hold-out method.
<sup>q</sup>LOO: leave-one-out cross-validation.
Assessment of Study Quality

Multimedia Appendix 6 shows the results of study quality assessments using QUADAS-2. Mean score (SD) was 5.6 (1.1), which corresponded to 80% of full marks (=7). The applicability of the reference test was evaluated to be low in 61% of the 33 included studies (20 studies) because hypoglycemia was not diagnosed by measuring blood glucose levels or ICD codes but by CGM (ie, glucose levels in blood are indirectly estimated from those in interstitial tissue) (19 studies) [15,26-30,35,38-43,46,47,49-52,54] or experts’ subjective judgement (1 study) [34]. The 2 factors were mainly responsible for lowering the study quality. We considered that the threshold of hypoglycemia in the index test was not specified in 7 studies, which used the cross-validation method [26,28,33,37,40,46], and 1 study, which used the leave-one-out method to separate test data from training data [48].

Data Synthesis

Ability for Detection of Hypoglycemia Using ML Algorithms

Figure 1 shows the HSROC curve and pooled estimates of sensitivity and specificity based on the 14 studies that assessed the ability of the ML algorithm to detect hypoglycemia. The pooled estimates (95% CI) were 0.79 (0.75-0.83) for sensitivity and 0.80 (0.64-0.91) for specificity. The pooled estimates (95% CI) of PLR and NLR were 2.20 (1.46-3.32) and 0.37 (0.28-0.49), respectively. Between-study heterogeneity expressed as I² was high both for PLR (98%; 95% CI 95%-99%) and NLR (80%; 95% CI 50%-90%). Statistically significant publication bias was detected (P=.15).

Figure 1. Hierarchical summary receiver-operating characteristic (HSROC) curve for detection of hypoglycemia using machine learning algorithms. Circles indicate study-specific sensitivity and specificity for each of the 14 included studies. The size of each circle is proportional to study sample size. The pooled point estimates of sensitivity and specificity are plotted in a filled square.

We conducted several sensitivity analyses using a portion of the above 14 studies that had 1 study characteristic in common. It was not apparent that any of the sensitivity analyses showed results different from the overall analysis. Limiting the analyses to 12 studies [15,20,25,32,43,48-53,55] that specified type 1 as the DM type, pooled sensitivity, specificity, PLR, and NLR were 0.78 (95% CI 0.73-0.82), 0.71 (95% CI 0.60-0.79), 2.65 (95% CI 1.88-3.72), and 0.26 (95% CI 0.19-0.36), respectively. When analyses were limited to the 7 studies that detected nocturnal hypoglycemia using ML algorithms [15,20,49-53], the pooled estimates (95% CI) were 0.75 (0.70-0.80) for sensitivity, 0.65 (0.55-0.74) for specificity, 2.14 (1.67-2.76) for PLR, and 0.38 (0.30-0.48) for NLR. With analyses of the 11 studies that detected hypoglycemia in an experimental setting, pooled sensitivity, specificity, PLR, and NLR were 0.78 (95% CI 0.73-0.82), 0.71 (95% CI 0.60-0.80), 2.66 (95% CI 1.84-3.85), and 0.31 (0.24-0.41), respectively. The pooled estimate (95% CI) was 0.78 (0.71-0.84) for sensitivity, 0.67 (0.55-0.77) for specificity, 2.39 (1.63-3.50) for PLR, and 0.33 (0.22-0.48) for NLR when the analysis was limited to 7 studies that used EEG abnormalities for detecting hypoglycemia. These estimations were similar when limited to 4 studies that used ECG abnormalities for detection of hypoglycemia: pooled estimate (95% CI) was 0.76 (0.67-0.82) for sensitivity; 0.67 (0.54-0.78) for specificity; 2.31 (1.65-3.23) for PLR; and 0.36 (0.28-0.47) for NLR.

Ability to Predict Hypoglycemia Using ML Algorithms

Figure 2 shows the HSROC curve for predicting hypoglycemia based on the 19 studies that assessed the predictive ability for
hypoglycemia. The point estimates (95% CI) were 0.80 (0.72-0.86) for sensitivity, 0.92 (0.87-0.96) for specificity, 10.42 (5.82-18.65) for PLR, and 0.22 (0.15-0.31) for NLR. Extremely high between-study heterogeneity was observed for both PLR (I² [95% CI] 100% [100%-100%]) and NLR (I² [95% CI] 99% [98%-100%]). Publication bias was not statistically significant (P=.68).

**Figure 2.** Hierarchical summary receiver-operating characteristic (HSROC) curve for prediction of hypoglycemia using machine learning algorithms. Circles indicate study-specific sensitivity and specificity for each of the 19 included studies. The size of each circle is proportional to study sample size. The pooled point estimates of sensitivity and specificity are plotted in a filled square.

When the analyses were limited to 13 studies that specified type 1 as the DM type [26,27,29,30,35,36,38,39,41,42,44,46,47], the pooled estimates (95% CI) were 0.77 (0.67-0.85) for sensitivity, 0.92 (0.84-0.96) for specificity, 9.82 (4.58-21.04) for PLR, and 0.25 (0.16-0.38) for NLR. In the analyses of 7 studies that specified night as the time of hypoglycemic events [26,30,31,35,36,41,44], the predictive ability was low compared with that of the overall analysis—pooled estimate (95% CI): 0.74 (0.65-0.82) for sensitivity, 0.81 (0.72-0.88) for specificity, 3.98 (2.64-6.00) for PLR, and 0.31 (0.23-0.43) for NLR. Relatively high sensitivity and low NLR were observed in the 13 studies that used CGM historical data for predicting hypoglycemia—pooled estimate (95% CI): 0.82 (0.71-0.90) for sensitivity, 0.92 (0.83-0.97) for specificity, 10.41 (4.52-24.01) for PLR, and 0.19 (0.12-0.32) for NLR—compared with 6 studies that did not use CGM—pooled estimate (95% CI): 0.76 (0.66-0.84) for sensitivity, 0.92 (0.88-0.95) for specificity, 10.14 (6.13-16.77) for PLR, and 0.26 (0.17-0.38) for NLR. After excluding 3 studies [31,33,54] that showed that the supposed hypoglycemic events occurred in-hospital, the pooled estimates (95% CI) of the 16 studies with such events occurring in an out-of-hospital setting were 0.82 (0.74-0.88) for sensitivity, 0.92 (0.85-0.96) for specificity, 10.58 (5.44-20.55) for PLR, and 0.20 (0.13-0.39) for NLR.

**Discussion**

**Principal Findings**

Overall, the PLR and NLR of ML algorithms for detecting hypoglycemia were 4.05 and 0.26, respectively. These estimates were almost unchanged throughout several sensitivity analyses that were limited to studies that shared 1 characteristic in common. According to the Users’ Guide to Medical Literature with regard to diagnostic tests [56], the PLR should be 5 or more to moderately increase the probability of persons having or developing a disease and the NLR should be 0.2 or less to moderately decrease the probability of having or developing a disease after taking the index test. In summary, the current ML algorithms had insufficient ability to detect the occurrence of hypoglycemia. However, that would not mean that ECG or EEG monitoring in combination with ML, which was the case with 79% (11/14) of the included studies, was useless in detecting hypoglycemia. For example, for patients with both DM and high cardiovascular risk, in particular, those who are vulnerable to cardiac arrhythmias, using ECGs for detecting hypoglycemia is useful considering that a hypoglycemia-induced arrhythmia could contribute to increased cardiovascular mortality [57]. Similarly, for patients with repeated episodes of hypoglycemia, the combination of ML and EEG was indicated to be beneficial to prevent hypoglycemia-induced neuroglycopenia resulting in cognitive impairment and ultimately death, because blood glucose levels alone do not appear to predict that condition [58].
Thus, the clinical applicability of these devices should be evaluated by the individual’s risk of hypoglycemia and its related arrhythmia and neuroglycopenia as well as the overall ability of algorithms for ML.

The overall sensitivity, specificity, PLR, and NLR for predicting hypoglycemia were 0.80, 0.92, 10.42, and 0.22, respectively. Applying the above described guidelines for diagnostic tests to these results, it is worth considering the use of current ML algorithms as a tool for alerting patients to impending hypoglycemic events. In addition, it is considered that a test with a PLR over 10 has a particularly strong power to alter posttest probability of the targeted disease compared with pretest probability [56]. If a positive test result were to be received, patients with DM who are administered hypoglycemic treatments would be strongly recommended to pay more attention to the possibility of impending hypoglycemic events than they would before receiving the predictive test for hypoglycemia. However, considering that the PLR and NLR values indicate relative risk (ie, risk of disease at posttest compared with that at pretest), the accuracy of predictive ability depends on patients’ risk of hypoglycemia in daily life. For example, even a less than 10% false-positive rate (8% in this meta-analysis) may be acceptable in patients at high risk of hypoglycemia but not in low-risk individuals due to too frequent false alarms. In such a case, there is fear that these patients will ignore the alarms and therefore miss the opportunity to take corrective action when the alarm is indeed true [59]. It is emphasized that the utility of ML algorithm depends on the extent of the patient’s risk of hypoglycemia. In addition, as indicated in the “Results” section, there was high between-study heterogeneity among studies. Specifically, when limiting analyses to the studies that predicted nocturnal hypoglycemia, the predictive ability was insufficient (pooled estimate: 3.98 for PLR; 0.31 for NLR). Considering that nocturnal hypoglycemia is the most common type of hypoglycemia among all hypoglycemic episodes [60], continued research is needed for further development of ML algorithms to predict hypoglycemia.

Several limitations of this meta-analysis should be addressed. First, the principal major limitation is the pooling of studies among which there was much variability in the type of DM, profiling data for detecting or predicting hypoglycemia, time of day when hypoglycemic events occurred, setting of supposed hypoglycemic events, and ML classification methods. In particular, although the ability for predicting hypoglycemia depended largely on the ML classification methods [33], this meta-analysis did not consider the difference in the test performance among various ML methods. Instead, the meta-analysis focused on ML’s comprehensive ability across studies using data in relation to the best model in each study, if 2 or more models existed, rather than comparisons among 2 or more models within 1 study. Given that generalization of evidence is among the most important roles in all meta-analyses, the issue of the variation in ML methods, in particular, the difference between old and new ML techniques, might be beyond the scope of this meta-analysis. Nevertheless, it should be emphasized that successful application of ML lies in the correct understanding of the advantages and disadvantages of different ML methods. Second, only 3 studies exclusively targeted patients with type 2 DM. With the increasing use of insulin to treat type 2 DM in the elderly, the prevalence of hypoglycemia is likely to escalate. In addition, the response to hypoglycemia is different between type 1 and type 2 DM [61]. Future studies should aim to develop and validate ML algorithms for detecting or predicting hypoglycemia in type 2 DM. Third, in most of the included studies, the ML classification models were developed in an experimental setting or by using previously recorded data as training and testing data instead of live data. Future studies need to train and test the algorithm on data from DM patients in everyday clinical practice to determine feasibility.

Conclusion
Overall, current ML algorithms have insufficient ability to detect ongoing hypoglycemia and considerable ability to predict hypoglycemia in patients with DM receiving hypoglycemic treatments. However, the clinical applicability of these ML algorithms should be evaluated according to patients’ risk profiles such as for hypoglycemia and its associated complications (eg, arrhythmia, neuroglycopenia) as well as the average ability of the ML algorithm. Continued research is required to further develop ML algorithms to enhance their feasibility, considering the inaccuracy of CGM in the hypoglycemic range, the increased prevalence of hypoglycemia in the elderly, and increasing evidence for the effectiveness of tight glycemic control in preventing microvascular complications [62].

Acknowledgments
All authors thank Ms Haga and Ms Chino in Niigata University for their excellent secretarial work. SK was financially supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS) (ID: 19K12840). The sponsor had no influence over the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Search strategy in this meta-analysis.
Multimedia Appendix 2
Study quality assessment using the quality assessment of diagnostic accuracy studies (QUADUS-2).

Multimedia Appendix 3
Study flow in this meta-analysis.

Multimedia Appendix 4
Databases which published articles that were retrieved by the search terms (see Appendix 1).

Multimedia Appendix 5
Profiling data input into ML algorithm for testing its performance.

Multimedia Appendix 6
Results of assessing study quality using revised tool for the quality assessment of diagnostic accuracy studies (QUADUS-2). The criterion corresponding to each domain (D) and signaling question (SQ) is indicated in Appendix 2.

References


43. Ling SH, San PP, Nguyen HT. Non-invasive hypoglycemia monitoring system using extreme learning machine for Type 1 diabetes. ISA Trans 2016 Sep;64:440-446. [doi: 10.1016/j.isatra.2016.05.008] [Medline: 27311357]


Abbreviations

CGM: continuous glucose monitoring
DM: diabetes mellitus
HSROC: hierarchical summary receiver operating characteristic
ICD: International Classification of Diseases
ML: machine learning
N-hypo: total number of events
NLR: negative likelihood ratio
N-total: total number of test data sets
PLR: positive likelihood ratio

©Satoru Kodama, Kazuya Fujihara, Haruka Shiozaki, Chika Horikawa, Mayuko Harada Yamada, Takaaki Sato, Yuta Yaguchi, Masahiko Yamamoto, Masaru Kitazawa, Midori Iwanaga, Yasuhiro Matsubayashi, Hirohito Sone. Originally published in JMIR Diabetes (http://diabetes.jmir.org), 29.01.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), 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.
Public Perspectives on Anti-Diabetic Drugs: Exploratory Analysis of Twitter Posts

Su Golder¹, BSc, MSc, PhD; Millie Bach¹, MSc; Karen O’Connor², MSc; Robert Gross³, MD, MSCE; Sean Hennessy², PharmD, PhD; Graciela Gonzalez Hernandez², PhD

¹Department of Health Sciences, University of York, York, United Kingdom
²Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Pennsylvania, PA, United States
³Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Pennsylvania, PA, United States

Corresponding Author:
Su Golder, BSc, MSc, PhD
Department of Health Sciences
University of York
Heslington
York, YO10 5DD
United Kingdom
Phone: 44 01904321904
Email: su.golder@york.ac.uk

Abstract

Background: Diabetes mellitus is a major global public health issue where self-management is critical to reducing disease burden. Social media has been a powerful tool to understand public perceptions. Public perception of the drugs used for the treatment of diabetes may be useful for orienting interventions to increase adherence.

Objective: The aim of this study was to explore the public perceptions of anti-diabetic drugs through the analysis of health-related tweets mentioning such medications.

Methods: This study uses an infoveillance social listening approach to monitor public discourse using Twitter data. We coded 4000 tweets from January 1, 2019 to October 1, 2019 containing key terms related to anti-diabetic drugs by using qualitative content analysis. Tweets were coded for whether they were truly about an anti-diabetic drug and whether they were health-related. Health-related tweets were further coded based on who was tweeting, which anti-diabetic drug was being tweeted about, and the content discussed in the tweet. The main outcome of the analysis was the themes identified by analyzing the content of health-related tweets on anti-diabetic drugs.

Results: We identified 1664 health-related tweets on 33 anti-diabetic drugs. A quarter (415/1664) of the tweets were confirmed to have been from people with diabetes, 17.9% (298/1664) from people posting about someone else, and 2.7% (45/1664) from health care professionals. However, the role of the tweeter was unidentifiable in two-thirds of the tweets. We identified 13 themes, with the health consequences of the cost of anti-diabetic drugs being the most extensively discussed, followed by the efficacy and availability. We also identified issues that patients may conceal from health care professionals, such as purchasing medications from unofficial sources.

Conclusions: This study uses an infoveillance approach using Twitter data to explore public perceptions related to anti-diabetic drugs. This analysis gives an insight into the real-life issues that an individual faces when taking anti-diabetic drugs, and such findings may be incorporated into health policies to improve compliance and efficacy. This study suggests that there is a fear of not having access to anti-diabetic drugs due to cost or physical availability and highlights the impact of the sacrifices made to access anti-diabetic drugs. Along with screening for diabetes-related health issues, health care professionals should also ask their patients about any non–health-related concerns regarding their anti-diabetic drugs. The positive tweets about dietary changes indicate that people with type 2 diabetes may be more open to self-management than what the health care professionals believe.

(JMIR Diabetes 2021;6(1):e24681) doi:10.2196/24681

KEYWORDS
diabetes; insulin; Twitter; social media; infodemiology; infoveillance; social listening; cost; rationing
In 2016, 4.2 million diabetes-related deaths were reported worldwide [1], which makes diabetes the seventh leading cause of mortality [2]. For both type 1 and type 2 diabetes, treatment and management aim to achieve adequate glycemic control [3]. Medication nonadherence is reported to be high for insulin and even higher for noninsulin anti-diabetic drugs [4,5]. Patients’ beliefs about medications, such as whether they are perceived to be essential or whether they have side effects, can influence both adherence and self-management behaviors [6]. The odds of nonadherence is 3.4 times as high in those who believe that anti-diabetic drugs have serious side effects and 14.3 times as high in people who believe that diabetes treatment regimens are too complex [7].

Given social media’s ability to connect large numbers of people and thereby generate large volumes of data, it has become a novel area for health research and a powerful tool to understand public perceptions. This study uses a particular social media site, that is, Twitter. As a popular social media outlet, Twitter is both a microblogging site and a social networking platform [8]. Since its conception in 2006 [9], Twitter’s popularity has grown to a reported 330 million monthly active users in 2019 [10]. The utilization of Twitter as a data collection platform is increasing and it is the most commonly utilized social media platform within health research [11]. Simmenberg et al [12] demonstrated that the number of health-related studies harnessing Twitter in 2015 was over 10 times higher than that in 2010, and their systematic review of 137 studies identified many ways in which Twitter data can be used. The most common Twitter analyses identified by the authors were content analyses, wherein the words, pictures, or sentiment of tweets are analyzed. The monitoring of vocabulary within tweets for pharmacovigilance purposes is an expanding area of research [13], while the exploration of tweets discussing perceptions of medications can help understand compliance and therapeutic decision making [14]. With regard to diabetes, studies have examined changing sentiments in Tweets on diabetes since the COVID-19 outbreak [15], and public perceptions have been examined on Twitter in detail for diseases such as Ebola virus disease [16] and cancer [17] and products such as e-cigarettes [18].

In this study, we sought to identify perceptions held by people discussing anti-diabetic drugs on Twitter. In particular, we sought to assess 3 questions: (1) Who discusses anti-diabetic drugs on Twitter? (2) Which anti-diabetic drugs are the most frequently discussed on Twitter? and (3) What are the most common health-related topics discussed on Twitter regarding anti-diabetic drugs?

Publicly available tweets posted between January 1, 2019 and October 1, 2019 were retrieved by the University of Pennsylvania’s Health Language Processing Center [19] from a large publicly available data set curated by the Internet Archive. The Internet Archive is a nonprofit organization that builds digital libraries of internet sites and provides free access to the data to researchers. We removed retweets from the collection. We selected this time scale in order to account for any seasonal or newsworthy variations in the tweets posted. Search terms associated with anti-diabetic drugs, including generic names, brand names, and common misspellings (Multimedia Appendix 1) were used to retrieve 10,308 tweets (Figure 1). After removing 515 duplicates, 92.9% (9107/9793) of the medication-related tweets were found to be about insulin. We, therefore, constructed a purposive sample of all tweets about noninsulin anti-diabetic drugs (n=686) so as to not lose any potential valuable information and a random sample about insulin (n=3314).

Qualitative studies traditionally have small sample sizes [20], but social media analyses are associated with qualitative data on a quantitative scale [21]. Consequently, qualitative Twitter analyses often use a sample of tweets rather than the full sampling frame [22]: sample sizes range from a few hundred [23] to thousands of tweets [12]. Guided by previous research, we initially began with 4000 random tweets (4000/9793 or 40.8% of our total sample), with additional samples to be analyzed if code saturation and meaning saturation were not met. Code saturation can be defined as the point at which all codes have been identified, while meaning saturation is the point at which all codes are understood [24]. After coding all 4000 tweets, code saturation and meaning saturation appeared to have been met [24] and a further sample was not necessary. Codes are labels for assigning units of meaning [25]. In qualitative content analysis, the use of codes results in the generation of themes that can be used to interpret the meaning of the text [26]. Health-related tweets were coded based on the perception expressed in the tweet. This used the conventional content analysis inductive framework proposed by Hsieh and Shannon [27] to explore both the manifest and latent meanings of the tweets and ensured that the codes arose from the data itself rather than being predefined. An inductive approach was particularly useful as there is little theory on anti-diabetic drug perceptions discussed via Twitter on which to base any assumptions and there is no particular framework to work from. Inductive approaches on Twitter data are also commonplace in the scientific literature [16]. Initial codes were given to each tweet, and upon reflection of the whole data set, similar or linked codes were clustered into themes. Some similar themes were further combined to form subthemes under an overarching theme.
The themes identified at this stage formed the basis of the coding scheme. We created a manual containing the coding scheme and instructions with examples on how to correctly assign codes. We filtered the Internet Archive data set by matching the keywords list, which includes all anti-diabetic drugs and their variants in the tweets. Only tweets in English and those that were not retweets were retrieved. The output file created contains all tweets where a match was found and included the user ID, tweet ID, tweet text, data created, and the keyword that matched in separate columns in an Excel. The keyword column helped ascertain the drug mention; however, the themes were hand-coded from scratch [28].

Two researchers independently coded 231 tweets by using the coding scheme. A random sample of 231 tweets was found to be sufficient to measure agreement and to stimulate discussion on the coding scheme as all codes were represented multiple times in this sample size. Because the initial kappa coefficient was 0.67, disagreements were discussed, and the coding instructions adapted accordingly. A further 169 tweets were then coded independently by 2 reviewers, producing a satisfactory kappa score of 0.73 [29]. Each of the remaining tweets was then coded by one of the two researchers, with all codes checked by the other reviewer and any disagreements resolved by discussion. First, tweets were coded for whether they truly were anti-diabetic drug–related. Second, any anti-diabetic drug–related tweets were coded as either health-related or non–health-related. Health-related tweets were further coded. Tweeters were categorized as (1) those who used the drug themselves, (2) people who knew someone who takes the drug, (3) health care providers, or (4) unclear, that is, the relationship between the tweeter and the anti-diabetic drug was unclear. Figure 2 shows a theoretical tweet, which has been coded, to show how coding was performed.

The availability of social media data means that it is relatively easy to trace quotations back to the user; therefore, there is a risk of deductive disclosure [30]. This makes reporting direct quotations problematic. Subtle changes to tweets are at odds with the Twitter display requirements, which prevent the alteration of tweets [31]. We, therefore, undertook a descriptive approach through paraphrasing tweets and by only directly quoting commonly used terms so that they cannot be traced back to an individual tweet. All data used in this study were collected according to the Twitter terms of use and were publicly available at the time of collection and analysis. We have an institutional review board certificate of exemption from the University of Pennsylvania. Each theme was explored regardless of how often it occurred.

Figure 1. Flowchart summarizing the tweet selection process.
Results

Tweeter Description

The results of this study are based on the 1664 health-related tweets (Table 1). A quarter (415/1664, 24.9%) of the tweets were by patients with diabetes taking anti-diabetic drugs, or who had taken the anti-diabetic drug in the past or who might initiate the anti-diabetic drug in the future; 87 (21.1%) of these self-identified as having type 1 diabetes, 61 (14.6%) as having type 2 diabetes, 2 (0.5%) as having gestational diabetes, and 2 (0.5%) as having secondary diabetes. The type of diabetes could not be classified for two-thirds of the tweeters; 17.9% (298/1664) of the tweets were second-person accounts, often about a family member or a person in a news story, and 2.7% (45/1664) of the tweets were from health care professionals. We could not establish the relationship between the tweeter and the anti-diabetic drug for the remaining 54.4% (906/1664) of the tweets.

Table 1. Proportions of the types of tweets and tweeters.

<table>
<thead>
<tr>
<th>Type of tweet/type of tweeter</th>
<th>Explanation</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Irrelevant tweets (n=2336)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non–health-related</td>
<td>Tweets that mention an anti-diabetic drug but are not directly related to health, for example, jokes, advertisements.</td>
<td>1556 (66.6)^a</td>
</tr>
<tr>
<td>Not a drug</td>
<td>Key term is used but is not in reference to a drug, for example, using the term “insulin” to mean the endogenous hormone rather than the exogenous anti-diabetic drug.</td>
<td>693 (29.6)</td>
</tr>
<tr>
<td>Not in English</td>
<td>The majority of the tweets were not in English.</td>
<td>7 (0.3)</td>
</tr>
<tr>
<td>Not related to diabetes</td>
<td>Tweet refers to drug being used for a purpose other than diabetes.</td>
<td>80 (3.4)</td>
</tr>
<tr>
<td><strong>Health-related tweets (n=1664)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-person report</td>
<td>Tweet from a diabetic person—uses phrases like “my drug…”</td>
<td>415 (24.9)</td>
</tr>
<tr>
<td>Second-person report</td>
<td>Tweets from someone who is not diabetic but is about a diabetic person—uses phrases like “my daughter’s drug…”</td>
<td>298 (17.9)</td>
</tr>
<tr>
<td>Health care professional</td>
<td>Tweet is from a health care professional—uses phrases like “my patient’s drug”</td>
<td>45 (2.7)</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>There is insufficient context to determine who is sending the tweet.</td>
<td>906 (54.4)</td>
</tr>
</tbody>
</table>

^aOf these, 920 (59.1%) tweets were on cost.

Anti-Diabetic Drugs Under Discussion

Tweets related to 33 anti-diabetic drugs across 11 drug classes were identified: insulin (1281 tweets), biguanides (194), SGLT2 inhibitors (102), DDP4 inhibitors (33), GLP1 agonists (97), sulfonlureas (11), thiazolidinediones (16), metformin (2), α-glucosidase inhibitors (1), meglitinides (1), and amylase analogues. People tweeted using both generic and brand names.

Common Perceptions

We identified 13 themes (Table 2). In most cases, we could not determine if the tweet was about type 1 or type 2 diabetes. Cost and efficacy dominated type 1 diabetes posts and other treatments, and adverse drug reactions dominated type 2 diabetes tweets. Type 1 diabetes tweets were also more likely to discuss more than one topic (Figure 3).
Table 2. Themes of the health-related tweet categories (n=1664).

<table>
<thead>
<tr>
<th>Theme</th>
<th>Explanation</th>
<th>Subthemes</th>
<th>n (%)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Tweet discusses the cost of an anti-diabetic drug in relation to health issues.</td>
<td>How much do anti-diabetic drugs cost? Attitudes toward cost, insurance problems, health consequences, social consequences, managing cost</td>
<td>669 (40.2)</td>
<td></td>
</tr>
<tr>
<td>Efficacy</td>
<td>Tweet discusses efficacy of the drug, both positive and negative. This includes tweets about the necessity of the drug and tweets that state that death will occur if the anti-diabetic drug is not taken.</td>
<td>Positive and negative</td>
<td>465 (27.9)</td>
<td></td>
</tr>
<tr>
<td>Information resource</td>
<td>Tweet provides information about the anti-diabetic drugs. These tweets reference research articles or clinical guidelines rather than someone’s belief about the anti-diabetic drugs.</td>
<td>Links and information summaries</td>
<td>371 (22.2)</td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td>Tweet discusses the availability of or access to anti-diabetic drugs.</td>
<td>Nationwide availability, personal availability, ensuring availability</td>
<td>158 (9.5)</td>
<td></td>
</tr>
<tr>
<td>Nonadherence</td>
<td>Tweet discusses someone not following the recommendation for taking the anti-diabetic drugs.</td>
<td>Taking too much, taking too little, consequences of nonadherence</td>
<td>124 (7.5)</td>
<td></td>
</tr>
<tr>
<td>Personal opinion</td>
<td>Tweet discusses a personal belief about anti-diabetic drugs.</td>
<td>Preferences, opinions of people without diabetes, opinions of people with diabetes</td>
<td>94 (5.6)</td>
<td></td>
</tr>
<tr>
<td>Other treatment options</td>
<td>Tweet compares an anti-diabetic drug to another management option for diabetes.</td>
<td>Other management options, effect on anti-diabetic drug, attitudes toward other treatments</td>
<td>54 (3.2)</td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Tweet is being used to seek advice or to challenge others.</td>
<td>Advice from others, educational tool</td>
<td>41 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Changes to treatment</td>
<td>Tweet discusses starting, stopping, or changing to another anti-diabetic drug.</td>
<td>Starting a medication, stopping a medication, changing insulin delivery</td>
<td>31 (1.8)</td>
<td></td>
</tr>
<tr>
<td>Stigma</td>
<td>Tweet discusses stigma surrounding anti-diabetic drugs.</td>
<td>Specific situations associated with insulin delivery, reducing stigma, opinions of people without diabetes</td>
<td>29 (1.7)</td>
<td></td>
</tr>
<tr>
<td>Dose</td>
<td>Tweet discusses dosing of anti-diabetic drugs. This includes stating the dose, saying how it is taken, or general statements about having to change the dose.</td>
<td>Stating the dose and calculating doses</td>
<td>28 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Adverse drug reaction</td>
<td>Tweet is about an experience of an adverse drug reaction. These should be tweets about adverse drug reactions that have actually happened, rather than beliefs about the potential side effects of an anti-diabetic drug.</td>
<td>Specific side effects, general side effects, associated with insulin delivery</td>
<td>21 (1.3)</td>
<td></td>
</tr>
<tr>
<td>Abuse</td>
<td>Tweet discusses taking the anti-diabetic drug for nonmedical reasons.</td>
<td>Intent to kill or for fun</td>
<td>10 (0.6)</td>
<td></td>
</tr>
<tr>
<td>Nonclassifiable</td>
<td>Some tweets did not provide enough context to determine what it was about.</td>
<td>Too short or incomprehensible</td>
<td>85 (5.1)</td>
<td></td>
</tr>
</tbody>
</table>
Anti-Diabetic Drugs Are Too Expensive

The cost of insulin was the most common topic. Some tweeters listed the cost while others described it as “too expensive” (669/1664, 40.2%). Tweeters also remarked that the cost had “skyrocketed.” Health care practitioners were aware of the impact. They described how prices had increased during their time and how they tried to prescribe low-cost anti-diabetic drugs. Cost was an issue for both those with and without health insurance coverage. Certain insurance plans cover certain drugs but not insulin. Younger people expressed fears about aging out of their parents’ insurance.

It was generally felt that high costs were unfair and the profit margin too great. Many believed that anti-diabetic drugs should be free. This was fueled by comparisons of the costs outside the United States or comparisons to other medications. The health consequences of being unable to afford anti-diabetic drugs were extensively discussed. Tweeters expressed difficulty in achieving blood glucose level targets, which they reported resulted in long-term repercussions such as losing limbs, going blind, renal failure, and strokes. Diabetic ketoacidosis was mentioned as a specific concern, and the worst-case scenario was death. There were also economic and social consequences such as bankruptcy and homelessness. Some tweeters had made lifestyle decisions based solely on their need for anti-diabetic drugs such as taking a job with insurance rather than a preferred job. Tweeters were open in discussing ways of affording anti-diabetic drugs, including asking other tweeters for money, selling their belongings, or working more than one job. Alternative options were buying cheaper anti-diabetic drugs from abroad, buying over-the-counter medicines, or turning to the black market. Large-scale approaches to making anti-diabetic drugs more affordable included using Twitter to promote campaigns such as the #InsulinForAll movement (a campaign launched in the lead up to World Diabetes Day in 2014 by The Pendsey Trust and T1 International) and to contact people in power, with tweets being sent to the US President and pharmaceutical companies.
Anti-Diabetic Drugs Have Varying Efficacy

There was an agreement that insulin was lifesaving. Short-term benefits such as glucose control were noted, as well as generally feeling better. Some tweeters reported issues with their insulin such as insufficient blood glucose reductions, and there were concerns about “Walmart insulin,” with some posts claiming that it is ineffective and caused hypoglycemia. Noninsulin anti-diabetic drugs were perceived to have different levels of efficacy (465/1664, 27.9%). For instance, exenatide and empagliflozin were viewed as effective in reducing weight, which was viewed favorably. Another SGLT2 inhibitor, canagliflozin, was reported to prevent microvascular complications. Metformin had mixed reviews; some felt it worked while others did not.

Wealth of Information on Anti-Diabetic Drugs

Information was mostly tweeted as links to or summaries of journal articles (371/1664, 22.2%). Articles varied from laboratory studies to efficacy evaluations. Studies exploring alternative methods of insulin delivery and the use of noninsulin anti-diabetic drugs as adjunct therapies in type 1 diabetes were considered particularly important. Information also came in the form of videos and links to reports on drug approvals and safety published by regulatory bodies.

Anti-Diabetic Drugs Are Not Always Available

Problems in availability included delays in mail orders, stolen, or lost medication (158/1664, 9.5%). There were posts calling for wider availability of nonprescription insulin. Some tweeters reported use of nonofficial outlets, and Twitter was used to find, sell, or give away extra supplies. Others discussed anti-diabetic drug availability on a national scale. The main topic concerning the United Kingdom was the impact of leaving the European Union. Additional barriers in the United States were the government shutdown from December 22, 2018 to January 25, 2019 [32], which caused financial and logistic issues, impaired access for deported immigrants, and US sanctions on Venezuela. Tweeters were proactive in discussing ways to ensure their anti-diabetic drug supply, such as stockpiling in the United Kingdom or traveling to Canada or Mexico from the United States. However, there were concerns over stockpiling due to storage issues and insulin’s shelf-life and a strong sense that people should not need to travel abroad to receive life-saving medications.

Adherence Can Be Difficult

The majority of tweeters reporting nonadherence mentioned missing doses (124/1664, 7.5%). Those mentioning metformin or liraglutide simply stated they had missed a dose, while insulin users provided more detailed reasons. Some forgot to take their insulin or had equipment problems; others deliberately choose not to take it. Reasons for this included dislike of needles, reactions to news stories condemning insulin, diabulimia with tweeters restricting their insulin intake to control their weight, and incorrectly following advice (this included injecting insulin through clothes or failing to take bolus insulin if not eating due to illness). The most commonly cited reason for nonadherence was cost (85/124, 68.5%), which led to rationing either by taking less insulin per injection or by omitting injections. Some who were not then rationing expressed fears about having to in the future. Insulin overdoses were less commonly discussed, with causes including misreading the dose volume or accidentally taking 2 injections. The only issue reported by tweeters who took an overdose was hypoglycemia.

Tweeters Hold a Range of Personal Beliefs

Some Tweeters stated preferences for particular anti-diabetic drugs that had no scientific evidence for the mechanism of action (94/1664, 5.6%). For instance, there was a perception that insulin makes type 2 diabetes worse. Tweeters with diabetes were mostly negative about being on anti-diabetic drugs, expressing that anti-diabetic drugs make life difficult. Some of these negative attitudes centered around equipment, including not liking the “huge” exenatide needles or the hassle of changing cartridges in prefilled insulin pens.

Anti-Diabetic Drugs Are Considered Alongside Other Treatments

Anti-diabetic drugs were discussed alongside lifestyle changes, particularly diet changes and specific diets, including the ketogenic diet or a vegan lifestyle (54/1664, 3.2%). Mentions of herbal treatments centered around a news story about the death of a person with type 1 diabetes whose herbalist advised the person to stop his/her insulin. Those using alternative or supplementary treatments were happy to do so, and many expressed annoyance at being offered anti-diabetic drugs with no option of management through lifestyle changes. Subsequently, these alternative treatments were discovered through social media or personal research rather than being initiated by a health care provider. The only alternative treatments that health care providers tweeted support for were exercise and ketogenic diets. Those with type 1 diabetes expressed frustration at being told to try nondrug treatments, particularly diet changes. Although they recognized that reducing carbohydrate intake can reduce insulin requirements, some felt the need to state that type 1 diabetes requires insulin, regardless of diet.

Anti-Diabetic Drugs Generate Questions

Those struggling to adjust their anti-diabetic drugs to adequately control their blood glucose levels sought advice from others, and there were questions about where to source “cheap” insulin (41/1664, 2.5%). Health care professionals asked their peers questions, including on the correct anti-diabetic drug, on theoretic scenarios, or interpretation of study findings.

Anti-Diabetic Drug Regimens Can Change

Tweeters with type 2 diabetes actively tried to avoid starting insulin. Similarly, stopping insulin was seen as an achievement. Those who had previously managed with only lifestyle changes felt apprehensive about initiating medications. Some tweeters completely stopped their anti-diabetic drugs, usually with guidance from health care providers and changing to a nondrug therapy. Insulin users reported changing to different types of insulin or administration method rather than a different class of anti-diabetic drugs. These data were captured from 1.8% (31/1664) of the tweets.
Anti-Diabetic Drugs Are Associated With Stigma

Taking insulin injections in the public resulted in perceptions of being judged or objection to the practice. Those wearing an insulin device or with scars and bruising due to needles felt these drew unwanted attention. Stigma was greater at airport checkpoints, work, or school. These data were captured from 1.7% of the tweets (29/1664). Some tweets discussed a reduction in stigma. This included restaurants providing carbohydrate content information to facilitate insulin dosing and the sense of togetherness when an individual saw other patients with diabetes taking injections. Some tweeters who did not have diabetes believed that there was no stigma for patients with diabetes, arguing that, “patients with diabetes are not judged for using insulin; so, why should people with depression be judged for taking antidepressants?”

Dosing Varies Based on the Anti-Diabetic Drug

Dosing based on meal-time carbohydrate or protein intake was noted to be difficult. Some tweeters shared their calculations. Some tweeters admitted to guessing their doses but that was not effective. For tweeters on noninsulin anti-diabetic drugs, doses were decided upon by health care providers. These data were captured from 1.7% of the tweets (28/1664).

Anti-Diabetic Drugs Can Cause Adverse Drug Reactions

The explicitness of the descriptions of the adverse drug reactions varied. Gastrointestinal issues, including vomiting or stomach aches, were mentioned for metformin and empagliflozin. Insulin and pioglitazone were both reported to cause weight issues. Other adverse drug reactions included allergic reactions to insulin, cognitive issues with metformin, and blood count changes with empagliflozin. Some adverse reactions were specific to the mode of insulin delivery, including local skin reactions to injections and scar tissue formation following the use of pumps. Other tweeters stated they had an adverse reaction but did not explain further. Tweeters discussed ways to cope, such as by spreading out the doses. The only adverse reaction that seemed to cause cessation was near-death experiences in 3 cases. These data were captured from 1.6% of the tweets (28/1664).

Anti-Diabetic Drugs Can Be Abused

There were first-person reports of deliberately taking too much insulin for the thrill of trying to restabilize blood glucose levels. Insulin was recognized as potentially deadly—there were tweets about people trying to kill themselves or someone else by administering insulin. These data were captured from 0.6% of the tweets (10/1664).

Non–Health-Related Tweets

While this study’s primary focus was the exploration of health-related tweets, it became evident that trends within the non–health-related tweets were also important (1556/1664). Though some non–health-related tweets were jokes or advertisements, 59.1% (920/1556) of the tweets were on the cost of anti-diabetic drugs—these raised similar issues to the health-related cost tweets without discussing the health implications.

Discussion

Overview

This study explored public perceptions of anti-diabetic drugs via the analysis of health-related tweets. We found that the issue of cost dominated both health and non–health-related tweets regarding insulin and overwhelmed our results, with implications for other identified themes such as availability, adherence (via rationing), and safety of cheaper versions. We found a similar proportion of health-related tweets in our sample (1664/4000, 41.6%) when compared to that in our study on statins (5201/11,852, 43.8%) [33]. However, the excluded non–health-related tweets differed from those on statins. People tweeting on the non–health-related aspects of anti-diabetic drugs often referred to cost or unfair pricing, while non–health-related tweets on statins were often cultural references, jokes, financial or news reports, or web-based pharmacies.

Within our health-related tweets, it was possible to identify whether the person tweeting was discussing their own diabetes in 24.9% of the cases (415/1664), someone known to them with diabetes in 17.9% of the cases (298/1664), or if they were in a health care profession (45/1664, 2.7%). Interestingly, with those tweeting on statins [33], it was possible to identify whether the person tweeting was taking statins in 32.8% of the cases (1707/5201), someone they know taking statins in 6.6% of the cases (346/5201), or whether the person was a health care professional (325/5201, 6.2%). The much higher proportion of people discussing someone known to them with diabetes may be because of the large scale concern for people with diabetes not being able to afford their insulin.

While type 2 diabetes makes up 90% of the global cases of diabetes [1], for those tweets where we could decipher the type of diabetes more were from people with type 1 than from people with type 2 diabetes and in line with this, insulin was by far the most discussed drug (9107/9793, 92.9% of the tweets). When considering that 44.7% of the people with type 1 diabetes are younger than 40 years compared to just 4% of the people with type 2 diabetes [34] and two-thirds of Twitter users are younger than 35 years [35], a possible partial explanation is that the Twitter demographic is more aligned with the younger demographic with type 1 diabetes. Another explanation is the high proportion of people discussing the injustice of the high cost of insulin for type 1 diabetes.

The implications of high-cost insulin were far reaching. While tweets reporting bankruptcy, stealing, and homelessness associated with the cost of insulin may seem like extreme subjects to post on a public platform, a study in 2020 with individuals with type 1 diabetes in the United States corroborated these stories [36]. Approximately 39.2% of the patients struggling to afford their insulin do not tell their health care professionals [37], making Twitter a potential way of identifying patients in need. Tweets about the increasing cost of insulin reflect the general trend in the United States. The price of insulin glargine—the most commonly prescribed insulin in the United States [38]—increased by 117% over 7 years [39]. Even for those who have a Medicare insurance plan, diabetes-related out-of-pocket spending increased by 10% per
year between 2006 and 2013 [40]. This is despite the average spending for other prescription medications only increasing by 2.8% over the same period [40]. An analysis of the tweets about statins found that only 3.5% (182/5201) of the tweets mentioned cost [33] compared to 40.2% (669/1664) of the tweets in this study. This may be because the cost of a month’s supply of statins, on average, is only one-third of the price of a month’s supply of anti-diabetic drugs [41].

A relationship between cost and availability, adherence, safety and efficacy was apparent from the tweets. Twitter appeared to be an informal marketplace for trading anti-diabetic drugs, although we did not confirm actual transactions. The overall sentiment of the tweets is that the lack of affordable anti-diabetic drugs is unfair and detrimental to health, which is in agreement with the findings of Litchman et al [42], who reported that those giving away their extra anti-diabetic drugs did so out of altruism and frustration at the lack of pricing regulations rather than the need to profit. Some tweeters travelled abroad to purchase their anti-diabetic drugs; these tweeters are among the estimated 2.3 million US individuals who buy their medications abroad [43]. Although this analysis cannot quantify how many individuals do this, it does give an insight into the reasons specific to anti-diabetic drugs. Prior research has found that those without health insurance are most likely to purchase prescription medications abroad [43], and this was reflected in the tweets. Of note, Hong et al [43] inferred that those seeking health information on the internet or using web-based chat groups were twice as likely to purchase medications abroad; therefore, given that this is a Twitter analysis, there may be an overrepresentation of individuals who purchase their anti-diabetic drugs in this way. It is currently illegal to purchase insulin abroad and import it into the United States for personal use [44]; therefore, the fear of being caught may explain why there has been little mention of this method in previous studies. In July 2019, the Food and Drug Administration proposed the Safe Importation Action Plan, intending to facilitate the import of medications from Canada [45]. Despite the tweet collection covering this period, there were no tweets related to this, questioning how far this announcement spread. The tweet collection period coincided with several delays to the date the United Kingdom was due to leave the European Union. Tweets related to this highlighted the importance of protecting medication imports. The worries about imports are supported by Holtt et al [46], who noted that only animal insulin is manufactured in the United Kingdom, with Novo Nordisk, Eli Lilly, and Sanofi having to import their insulins.

This study indicates the potential impact of high-cost insulin and concerns about availability, leading to rationing. This is in line with the results of a global survey of 1478 individuals with type 1 diabetes, and their care providers reported that 25.9% of the respondents from the United States had rationed their insulin at some point in the last year [47]. Rationing is deeply problematic and there was a little debate regarding insulin’s effectiveness, with powerful descriptions of how it is lifesaving. Participants with type 1 diabetes in a previous study described insulin as “life or death” for them [36], but this analysis shows that the general public also appreciates the life-saving nature of insulin. We found little evidence of the stigma associated with being on insulin among people with type 1 diabetes, which has been reported in previous studies [48]. The growing empathy for people with type 1 diabetes because of the high prices of insulin may be interconnected with a decline in the stigma.

Opinions on the efficacy of anti-diabetic drugs to treat type 2 diabetes were more varied; many tweeters expressed their desire to stop their medication, and tweets discussing other treatment options for type 2 diabetes seemed to favor dietary changes. Other studies have also indicated poor adherence in type 2 diabetes [49]. With respect to type 2 diabetes, people experience more stigma when on insulin than when on a noninsulin anti-diabetic drug [50]. A qualitative systematic review found that health care providers often doubt their patients’ ability to self-manage their diabetes, consequently preferring a paternalistic approach [51]. This is reflected in the sense of annoyance among the tweeters at not being given the option to manage type 2 diabetes by lifestyle changes alone.

There has been interest in using Twitter as a source for collecting anecdotal accounts of adverse drug reactions [13]. In our analysis of statins [33], we identified 6.8% (353/5201) of the tweets to be about adverse reactions compared to just 1.3% (21/1664) in this study. This was unexpected, given that dose-related serious adverse effects with drugs to treat diabetes are considered to be among the adverse drug effects with the highest public health impact [52], while statins have a much higher degree of safety. The cheap version ReliOn (Walmart insulin) was the only type of insulin that had its efficacy and safety questioned.

A major source of criticism of social media is the high volume of misinformation. Misinformation on social media can have detrimental effects on health behaviors, and they are difficult to correct once they gain acceptance [53]. We found little evidence of misinformation among our tweets, and in line with the literature, no misinformation was shared by health care professionals [53]. Broadly, there were 2 ways individuals used Twitter to discuss anti-diabetic drugs. The first was as a microblogging site for recording day-to-day experiences such as trying to afford their insulin, rationing, side effects, and incidences involving stigma. These tweets may provide a useful introduction into what life is like while taking anti-diabetic drugs, which could influence the support provided by health care professionals. Alternatively, Twitter was used as a tool that was intended to bring about change, with tweeters discussing complex social issues. This is pertinent to policymakers as it highlights the issues that both patients and the public consider most pressing.

**Strengths and Limitations**

The large volume of Twitter data from a mix of tweeters with and without diabetes allowed an insight into a broad range of perspectives. Manual coding was used during the tweet analysis, which is considered the gold standard method [28]. While the use of automated computer programs may be quicker and can allow large data sets to be coded, they are associated with lower accuracy [22]. These findings represent the perspectives of the Twitter-using population but not necessarily the general population [54]. As an illustration, in the United States, the average tweeter is likely to be White, young, well-educated,
and a Democrat [54]. As this study did not collect demographic
data, it is hard to appreciate which population this study does
reflect. Since Twitter is available worldwide, this study planned
to take a global approach to anti-diabetic drug perceptions, but
upon analysis, it became evident that a large burden of the tweets
centered around issues in the United States. It was only after
the research process began that Patel et al [55] published their
analysis of 50,286 diabetes-related tweets, indicating that 43.6%
of the tweets came from the United States, followed by 14.9%
from the United Kingdom. Despite the large volume of tweets,
we only identified issues relevant to a few countries and were
unable to compare differences among countries, as we did not
collect the geolocations of the Twitter users. Future work could
address this. The limited non-US issues collected may, in part,
be because of the search terms we used and that we only used
a single social media platform. Other platforms may be used
to explore perceptions from a wider population and in other
countries. Our analysis does not go beyond content analysis.
We did not record any user engagement metrics or interactions.
We were also unable to verify any of the claims made, and
people may post things on the internet that they would not say
in person. However, the fact that information shared on social
media is expressed spontaneously in an open digital space with
a flat role hierarchy is a major advantage for capturing
perceptions that otherwise would not be reported [56]. Finally,
we were unable to distinguish whether posts were referring to
type 1 or type 2 diabetes in the majority of the tweets. Issues
with anti-diabetic drugs are likely to be dependent on the type
of diabetes. This limitation may be generalizable to other
medications studied on social media, which are used for more
than one indication.

Conclusion
The use of Twitter has provided an insight into the immediate
perceptions of anti-diabetic drugs outside of a clinical setting,
thereby giving a unique perspective. Not only does this study
support the findings already established in the current literature,
but it has also provided an appreciation of the struggles of people
taking anti-diabetic drugs, particularly in light of the high cost
of insulin. This study has also shown that the public is aware
of these issues and are waiting for governments and health care
systems to make changes.

Acknowledgments
This work was supported by National Institutes of Health (NIH) National Library of Medicine under grant number NIH NLM
1R01. NIH National Library of Medicine funded this research but were not involved in the design and conduct of the study;
collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision
to submit the manuscript for publication. SG and KO had full access to all the data in the study and take responsibility for the
integrity of the data and the accuracy of the data analysis.

Conflicts of Interest
Sean Hennessy has received grant support and has consulted for numerous pharmaceutical companies. All other authors report
no conflicts of interest.

Multimedia Appendix 1
Key terms used for the search.

References
[accessed 2021-01-16]
204871/9789241565257_eng.pdf?sequence=1 [accessed 2020-01-16]
4. Peyrot M, Barnett AH, Meneghini LF, Schumm-Draeger PM. Insulin adherence behaviours and barriers in the multinational
Text] [doi: 10.1111/j.1464-5491.2012.03605.x] [Medline: 22313123]
between non-adherence behaviors, patients' experience with healthcare and beliefs in medications: a survey of patients with
31580168]
7. Mann DM, Ponieman D, Leventhal H, Halm EA. Predictors of adherence to diabetes medications: the role of disease and


19. Health language processing lab at the Institute for Biomedical Informatics. URL: https://healthlanguageprocessing.org [accessed 2020-09-16]


41. Langreth R, Migliozzi B, Gokhale K. The U.S. pays a lot more for top drugs than other countries. The U.S. pays a lot more for top drugs than other countries.: Bloomberg URL: https://www.bloomberg.com/graphics/2015-drug-prices/ [accessed 2020-07-18]


44. United States Food and Drug Administration. FDA Safe Importation Action Plan. Plan outlines pathways to lower prices and reduce out of pocket costs through safe importation of certain prescription drugs. URL: https://www.fda.gov/about-fda/reports/fda-safe-importation-action-plan [accessed 2021-01-16]


47. T1 International. Costs and rationing of insulin and diabetes supplies: findings from the 2018 T1 International patient survey. URL: https://pdfs.semanticscholar.org/ab82/46c5d107cb5f81e49df1e6477b99767ceb.pdf?_g=2.15937203.488900161.1594582198.1434567109.1594038301 [accessed 2020-07-13]


© Su Golder, Millie Bach, Karen O’Connor, Robert Gross, Sean Hennessy, Graciela Gonzalez Hernandez. Originally published in JMIR Diabetes (http://diabetes.jmir.org), 26.01.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), 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.
Diabetes Distress and Glycemic Control in Type 2 Diabetes: Mediator and Moderator Analysis of a Peer Support Intervention

Kara Mizokami-Stout\textsuperscript{1,2,3}, MSc, MD; Hwajung Choi\textsuperscript{4,5}, PhD; Caroline R Richardson\textsuperscript{6}, MD; Gretchen Piatt\textsuperscript{7,8}, MPH, PhD; Michele Heisler\textsuperscript{3,4,8}, MPA, MD

\textsuperscript{1}National Clinician Scholars Program, Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States
\textsuperscript{2}Division of Metabolism, Endocrinology and Diabetes, University of Michigan, Ann Arbor, MI, United States
\textsuperscript{3}Ann Arbor Veteran Affairs Hospital, Ann Arbor, MI, United States
\textsuperscript{4}Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States
\textsuperscript{5}Department of Health Management and Policy, University of Michigan, Ann Arbor, MI, United States
\textsuperscript{6}Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States
\textsuperscript{7}Department of Learning Health Sciences, University of Michigan, Ann Arbor, MI, United States
\textsuperscript{8}Department of Health Behavior and Health Education, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Kara Mizokami-Stout, MSc, MD
Division of Metabolism, Endocrinology and Diabetes
University of Michigan
1000 Wall Street
5100 Brehm Tower
Ann Arbor, MI,
United States
Phone: 1 734 232 1269
Email: kmizokam@med.umich.edu

Abstract

Background: High levels of psychosocial distress are correlated with worse glycemic control as measured by glycosylated hemoglobin levels (HbA\textsubscript{1c}). Some interventions specifically targeting diabetes distress have been shown to lead to lower HbA\textsubscript{1c} values, but the underlying mechanisms mediating this improvement are unknown. In addition, while type 2 diabetes mellitus (T2D) disproportionately affects low-income racial and ethnic minority populations, it is unclear whether interventions targeting distress are differentially effective depending on participants’ baseline characteristics.

Objective: Our objective was to evaluate the mediators and moderators that would inform interventions for improvements in both glycemic control and diabetes distress.

Methods: Our target population included 290 Veterans Affairs patients with T2D enrolled in a comparative effectiveness trial of peer support alone versus technology-enhanced peer support with primary and secondary outcomes including HbA\textsubscript{1c} and diabetes distress at 6 months. Participants in both arms had significant improvements in both HbA\textsubscript{1c} and diabetes distress at 6 months, so the arms were pooled for all analyses. Goal setting, perceived competence, intrinsic motivation, and decisional conflict were evaluated as possible mediators of improvements in both diabetes distress and HbA\textsubscript{1c}. Baseline patient characteristics evaluated as potential moderators included age, race, highest level of education attained, employment status, income, health literacy, duration of diabetes, insulin use, baseline HbA\textsubscript{1c}, diabetes-specific social support, and depression.

Results: Among the primarily African American male veterans with T2D, the median age was 63 (SD 10.2) years with a baseline mean HbA\textsubscript{1c} of 9.1% (SD 1.7%). Improvements in diabetes distress were correlated with improvements in HbA\textsubscript{1c} in both bivariate and multivariable models adjusted for age, race, health literacy, duration of diabetes, and baseline HbA\textsubscript{1c}. Improved goal setting and perceived competence were found to mediate both the improvements in diabetes distress and in HbA\textsubscript{1c}. Baseline patient characteristics evaluated as potential moderators included age, race, highest level of education attained, employment status, income, health literacy, duration of diabetes, insulin use, baseline HbA\textsubscript{1c}, diabetes-specific social support, and depression.

Conclusions: Prior studies have demonstrated that some but not all interventions that improve diabetes distress can lead to improved glycemic control. This study found that both improved goal setting and perceived competence over the course of the
peer support intervention mediated both improved diabetes distress and improved HbA1c. This suggests that future interventions targeting diabetes distress should also incorporate elements to increase goal setting and perceived competence. The intervention effect of improvements in diabetes distress on glycemic control in peer support may be more pronounced among White and insulin-dependent populations. Additional research is needed to understand how to better target diabetes distress and glycemic control in other vulnerable populations.

(JMIR Diabetes 2021;6(1):e21400) doi:10.2196/21400

KEYWORDS
diabetes mellitus; diabetes distress; health behavior; peer support

Introduction

Diabetes distress, or the negative emotional and behavioral responses that can occur as a result of having a demanding chronic illness like diabetes, is an increasingly recognized psychosocial factor influencing diabetes self-management [1]. The prevalence of at least moderate levels of diabetes distress is up to 45% in adults with type 2 diabetes (T2D) [2], and high levels of diabetes distress lead to poor medication adherence, higher glycosylated hemoglobin A1c (HbA1c) values, and, ultimately, poor quality of life [2-4].

While the link between high levels of diabetes distress and higher HbA1c has been well established [1], a number of evaluated interventions specifically targeting diabetes distress lead to improvements in glycemic control [5]. Examples of such interventions include educational, psychosocial, or psychological programs (including cognitive behavioral therapy, motivational interviewing, and mindfulness-based interventions). Prior RCTs and systematic reviews have elucidated that psychosocial and psychological interventions, particularly those that are tailored specifically for diabetes and have a patient empowerment or motivational interviewing component, are more successful at improving glycemic outcomes in addition to reducing diabetes distress [5-9]. The exact mechanisms behind this relationship are not clear, but drawing on well-established behavioral theories may help to clarify this link. Perceived competence and self-efficacy, or the belief in an individual’s ability to complete a task, is a key feature of social cognitive theory [10], and it has been found to be consistently negatively correlated with distress and is in the mechanistic pathway between diabetes distress and self-management behaviors in T2D [11,12]. It is therefore likely that improving [2] perceived competence is an important element of interventions that improve both diabetes distress and glycemic control. Similarly, self-determination theory postulates that autonomy support, defined as the provision of social support in a way that respects the patient’s values, autonomy, and choice, is an important motivator for patients with chronic disease such as diabetes [13]. As such, autonomy support has also been shown to be an important buffer against the effects of diabetes distress on glycemic outcomes [14]. However, beyond this, there is not a consistent strategic approach common among interventions that improves both diabetes distress and glycemic control. Further elucidation is thus needed to ensure that effective intervention components that improve these constructs are incorporated into future interventions for diabetes mellitus.

Equally important is understanding the characteristics of participants who benefit the most from these interventions. Prior studies have found that patients who are younger, female, have longer duration of diabetes, and are of ethnic minority status, particularly African Americans, have higher diabetes distress levels [15-17]. Interventions targeting specific ethnic minority populations who experience disproportionate diabetes burden and elevated diabetes distress levels have shown mixed findings. These studies, however, are limited by small sample sizes and do not allow comparisons of effects across participants of different ethnicities [18]. Similarly, diabetes-specific characteristics of those who respond to interventions specifically for distress are unknown. As may be anticipated, high diabetes distress levels are associated with fear of insulin use in insulin-naïve patients [19], but it is unclear whether interventions targeting distress are as effective in insulin users as in noninsulin users.

Peer support interventions, in which an individual with prior experience or knowledge who has been successful in their own self-management behaviors serves as a supportive mentor for a target population of patients with similar ethnic or socioeconomic background, are emerging as an important tool for patients with diabetes mellitus, particularly for vulnerable patient populations [14]. Peer support interventions have been successful in improving both glycemic outcomes and psychosocial outcomes, including diabetes distress, and are an attractive, low-cost approach for health care systems [20-22]. A recently published randomized controlled trial (RCT) of peer support versus technology-enhanced peer support for primarily African American veterans with T2D who receive care at an urban Veterans Affairs (VA) health center published by Heisler et al [23] demonstrated that the peer coach model they evaluated, both with and without technology enhancement, was effective at improving glycemic control and reducing diabetes distress over the 6-month intervention period.

In this trial, participants were randomized to peer coaches without any additional eHealth tools or to peer coaches using an individually tailored, web-based educational tool (iDecide) over the course of 6 months. This tool had interactive features to allow participants to understand their personal diabetes risk profile as well as explore options for medications based on cost, effectiveness, and side effects [23]. Peer coaches all received training in motivational interviewing [23]. In this trial, both arms achieved statistically and clinically significant improvements in both diabetes distress and HbA1c without any significant difference between the two intervention arms [23]. This successful trial thus presents an opportunity to explore the...
psychosocial mechanisms that lead to improvements in glycemic control when diabetes distress is reduced as well as the participant baseline characteristics that may predict responsiveness to such an intervention. The objectives of this study were therefore to evaluate mediators and moderators in the relationship between change in diabetes distress and change in glycemic control over a 6-month period in response to a peer support intervention.

Methods

Conceptual Model for Mediator and Moderator Analysis

A mediator analysis is one method to explore the psychosocial mechanisms that link diabetes distress and glycemic control. In such an analysis, a conceptual model is created that hypothesizes potential targets, or mediators, along the mechanistic pathway that an intervention must include in order to be successful in achieving the desired outcome. In the previously mentioned RCT by Heisler et al [23], participants had at least weekly contact with a fellow patient with T2D who had received a 2-hour training session with a focus on motivational interviewing, including active listening skills, rolling with resistance, enhancing change talk, goal setting, and action planning. During these sessions, peer coaches helped participants develop and follow up on weekly action steps to meet the participants’ defined behavioral goals. In order to ensure fidelity and help further strengthen the peer coach’s motivational interviewing skills, we held monthly hour-long booster sessions to provide reinforcement and additional training to coaches throughout the intervention period. Based on self-determination theory, which postulates that patients with diabetes who experience more autonomy supportiveness by their health care providers and supporters are more motivated and perceive themselves to be more competent in diabetes self-management, we hypothesized that both intrinsic motivation and perceived competence are important targets in the mechanistic pathway between diabetes distress and glycemic control [24]. Similarly, based on prior studies demonstrating the importance of goal setting and decisional conflict, we hypothesized that both are crucial elements of self-management support interventions to improve both diabetes distress and glycemic control [25]. Our full mediation model is demonstrated in Figure 1 with the pathway through relationship a and relationship b demonstrating the fully mediated model through our hypothesized mediators of goal setting, perceived competence, intrinsic motivation, and decisional conflict.

Figure 1. Conceptual model for hypothesized mediators and moderators of improved glycemic control in a peer coaching intervention.

A moderator analysis can be used to evaluate the characteristics of participants who benefited the most from the peer support intervention of reducing diabetes distress to improve glycemic outcomes. These characteristics are called moderators as they help inform differential effects in the relationship between an independent and dependent variable and hence identify potential modifiers and/or target population for the intervention. In our conceptual model shown in Figure 1, we hypothesized that...
potential moderators include baseline patient characteristics (age, race, education, employment, and health literacy), certain diabetes characteristics (duration of diabetes, HbA₁c, and insulin use), diabetes-specific social support, and comorbid depression. Our specific questions were as follows:

- In an intervention that improves both diabetes distress and glycemic control, are improvements in diabetes distress correlated with improvements in HbA₁c (main effect)?
- Do goal setting, perceived competence, intrinsic motivation, and decisional conflict work individually or in combination to mediate the relationship between diabetes distress and glycemic control (mediating effect)?
- Does age, race, education, employment, health literacy, duration of diabetes, HbA₁c, insulin use, diabetes-specific social support, or depression moderate the relationship between diabetes distress and glycemic control (moderating effect)?

### Setting, Recruitment, Intervention, and Measures

The target population for this study included veterans with T2D and high baseline HbA₁c values enrolled in a comparative effectiveness RCT of peer support versus technology-enhanced peer support. The description of recruitment, intervention, outcomes, and results of this RCT have been described previously [23]. Glycemic control was measured using HbA₁c at baseline and 6 months. Diabetes distress and potential mediators were measured using validated surveys at baseline and 6 months, which were then scaled from 0 to 100, with higher numbers indicating more positive outcomes (eg, lower diabetes distress, higher goal setting). Specifically, the following scales were used (see Multimedia Appendix 1 for further details):

- **Diabetes distress:** Measured, analyzed, and reported using the 2-item validated Diabetes Distress Scale–2, which assesses feelings that living with diabetes is overwhelming and/or that the participant is failing in their diabetes management [26,27].
- **Goal setting:** Measured, analyzed, and reported using the 3-item goal setting subscale of the Patient Assessment of Chronic Illness Care, which assesses whether participants were aided in setting goals for self-management and, if so, whether an action plan was developed [28].
- **Perceived competence:** Measured, analyzed, and reported using the 4-item validated Perceived Competence scale, which assesses the extent to which a participant feels confident and capable of meeting the challenges of diabetes self-management [13].
- **Intrinsic motivation:** Measured, analyzed, and reported using the intrinsic motivation subscale of the Treatment Self-Regulation Questionnaire, which assesses the extent to which participants feel self-motivated to improve their health behaviors [13].
- **Decisional conflict:** Measured, analyzed, and reported using the 1-item validated Decisional Conflict Scale, which assess the extent to which a participant is satisfied with their medication options for diabetes [29].

In the RCT, both arms demonstrated improved diabetes distress and HbA₁c values at 6 months. Therefore, in this study, participants in both arms were combined to investigate goal setting, perceived competence, intrinsic motivation, and decisional conflict as potential mediators, as shown in Figure 1. Additionally, baseline characteristics were evaluated as moderators of improvement in both diabetes distress and glycemic control, also shown in Figure 1.

### Statistical Analysis

Descriptive statistics were used to evaluate frequencies and means of baseline participant characteristics, and paired t tests were used to evaluate the change in means from baseline to 6 months for the independent variable, dependent variable (HbA₁c), and hypothesized mediator variables (goal setting, perceived competence, intrinsic motivation, and decisional conflict). Bivariate and multivariable linear regressions were used to assess whether the change in diabetes distress at 6 months (independent variable) is associated with the change in HbA₁c at 6 months (dependent variable). Covariates include age, race, health literacy, duration of diabetes, and baseline HbA₁c.

We next assessed the role of goal setting, perceived competence, intrinsic motivation, and decisional conflict as mediators between the change in diabetes distress and the change in HbA₁c at 6 months. Multivariable linear regression models were used with the covariate adjustments of age, race, health literacy, duration of diabetes, and baseline HbA₁c. This is conceptualized by the mediation model in Figure 1:

1. **Relationship a:** between diabetes distress (independent variable) and all potential mediators (dependent variables)
2. **Relationship b:** between all potential mediators (dependent variable) and HbA₁c

The potential mediators that were found to be significantly associated with the change in diabetes distress and HbA₁c at 6 months were selected for formal mediation testing by using seemingly unrelated linear regression techniques [30]. We evaluated each individual mediator separately as well as the shared effect of the combined mediators on the mediation pathway through relationships a and b (the indirect pathway) [30]. We calculated bias-corrected 95% confidence intervals from a bootstrapping method with 5000 replications [30].

Finally, sociodemographic factors (age, race, highest attained education, income, employment) and baseline clinical and psychosocial attributes (health literacy, HbA₁c, duration of diabetes, insulin use, diabetes-specific social support, depressive symptoms) were assessed as potential moderators of the relationship between change in diabetes distress and change in HbA₁c at 6 months. Multivariable linear regressions include an interaction term between the change in diabetes distress at 6 months and each of the potential moderators as well as those variables themselves. The change in HbA₁c at 6 months was the independent variable in these models and covariates included age, race, health literacy, duration of diabetes, and baseline HbA₁c except where the variable was tested as a moderator. This moderator model is conceptualized in Figure 1 (ie, differential effects on relationship d). For each potential moderator, the significance of the interaction term was assessed.
for different subgroups, and the difference in coefficients between the subgroups was evaluated for significance.

**Results**

**Description of the Sample**

A total of 290 veterans with T2D were enrolled in the two intervention arms of the RCT. Baseline characteristics of the full cohort are shown in Table 1. Being a veteran population, 98% of the participants were male with an average age of 63 (SD 10.2) years, and 63% were African American. The average HbA1c was 9.1% (SD 1.7%) with a mean of 15 years of diabetes duration, and 60% of the participants were insulin-dependent. At 6 months, diabetes distress improved by 4.8 points (95% CI 2.2 to 7.5; \( P < .001 \)) and mean HbA1c levels improved by 0.7% (95% CI –0.9 to –0.5; \( P < .001 \)) in all participants (Multimedia Appendix 2). Scores for goal setting, perceived competence, intrinsic motivation, and decisional conflict improved by 14.3, 6.9, 6.8, and 6.8 points, respectively (all \( P < .001 \)) at 6 months (Multimedia Appendix 2).
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>63 (10.2)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>7 (2)</td>
</tr>
<tr>
<td>Male</td>
<td>283 (98)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>181 (62)</td>
</tr>
<tr>
<td>White</td>
<td>106 (37)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (0.7)</td>
</tr>
<tr>
<td><strong>Work status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>74 (26)</td>
</tr>
<tr>
<td>Not employed</td>
<td>49 (17)</td>
</tr>
<tr>
<td>Retired</td>
<td>141 (49)</td>
</tr>
<tr>
<td>Disabled</td>
<td>23 (8)</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>12 (4)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>78 (27)</td>
</tr>
<tr>
<td>Some tech or vocational</td>
<td>23 (8)</td>
</tr>
<tr>
<td>Some college or more</td>
<td>177 (61)</td>
</tr>
<tr>
<td><strong>Income ($), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>1-15,000</td>
<td>61 (21)</td>
</tr>
<tr>
<td>16,000-30,000</td>
<td>81 (28)</td>
</tr>
<tr>
<td>31,000-55,000</td>
<td>59 (20)</td>
</tr>
<tr>
<td>56,000 and above</td>
<td>46 (16)</td>
</tr>
<tr>
<td>Prefer not to discuss</td>
<td>42 (15)</td>
</tr>
<tr>
<td>Baseline HBA(_{1c}), mean (SD)</td>
<td>9.1 (1.7)</td>
</tr>
<tr>
<td>Number of years with diabetes, mean (SD)</td>
<td>15.2 (10.0)</td>
</tr>
<tr>
<td>Insulin use, n (%)</td>
<td>171 (60)</td>
</tr>
<tr>
<td>Number of oral antihyperglycemic meds, mean (SD)</td>
<td>1.1 (0.8)</td>
</tr>
<tr>
<td>Health literacy, mean (SD)</td>
<td>7.0 (1.9)</td>
</tr>
<tr>
<td>Diabetes-specific social support(^b), mean (SD)</td>
<td>54.4 (14.3)</td>
</tr>
<tr>
<td>Depression(^c), mean (SD)</td>
<td>76.9 (27.0)</td>
</tr>
</tbody>
</table>

\(^a\)HBA\(_{1c}\): hemoglobin A\(_{1c}\).

\(^b\)Based on the Diabetes-Specific Social Support Needs assessment \([31]\), scaled score ranging from 0 to 100, with more positive outcomes reflected by higher numbers.

\(^c\)Based on the Patient Health Questionnaire–2 scaled score ranging from 0 to 100, with more positive outcomes reflected by higher numbers.

**Results of the Main Relationship**

A significant association between the improvement in diabetes distress and decreased HbA\(_{1c}\) was found in the unadjusted model (\(\beta\) coefficient \(-0.017; 95\% \text{ CI} -0.028 \text{ to } -0.006; \text{ } P=.003\) (relationship d). This association remained significant in the adjusted model, controlling for age, race, health literacy, duration of diabetes, and baseline HbA\(_{1c}\) (\(\beta\) coefficient \(-0.015; 95\% \text{ CI } -0.025 \text{ to } -0.006; \text{ } P=.001\)).

**Results of the Mediator Analysis**

Improvement in goal setting at 6 months was associated with improvements in diabetes distress (\(\beta\) coefficient 0.225, \(P=.02\)) and reduction in the HbA\(_{1c}\) (\(\beta\) coefficient \(-0.09; \text{ } P=.004\)) at 6 months. Similarly, improvement in perceived competence at 6 months was associated with both improvements in diabetes distress (\(\beta\) coefficient 0.182, \(P=.002\)) and the improvement in HbA\(_{1c}\) (\(\beta\) coefficient \(-0.011; \text{ } P=.03\)) at 6 months. Neither
intrinsic motivation or decisional conflict were associated with the change in diabetes distress or change in HbA\textsubscript{1c} at 6 months so were removed from further mediation analyses. These results are highlighted in Table 2.

**Table 2.** Adjusted estimates of the effect of diabetes distress on all potential mediators (relationship a) and the effect of all mediators on hemoglobin A\textsubscript{1c} (relationship b).

<table>
<thead>
<tr>
<th>Potential mediator (outcome in relationship a; predictor in relationship b)</th>
<th>β coefficient</th>
<th>95% CI</th>
<th>P value</th>
<th>β coefficient</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal setting</td>
<td>.225</td>
<td>.036 to .414</td>
<td>.02</td>
<td>−.009</td>
<td>−.015 to .002</td>
<td>.004</td>
</tr>
<tr>
<td>Perceived competence</td>
<td>.183</td>
<td>.065 to .300</td>
<td>.002</td>
<td>−.011</td>
<td>−.021 to −.001</td>
<td>.03</td>
</tr>
<tr>
<td>Intrinsic motivation</td>
<td>.007</td>
<td>−.127 to 1.141</td>
<td>.91</td>
<td>−.008</td>
<td>−.017 to .001</td>
<td>.07</td>
</tr>
<tr>
<td>Decisional conflict</td>
<td>.101</td>
<td>−.053 to .255</td>
<td>.20</td>
<td>−.007</td>
<td>−.015 to .003</td>
<td>.06</td>
</tr>
</tbody>
</table>

aDiabetes distress, hemoglobin A\textsubscript{1c}, and all potential mediators assessed as the mean change from baseline to 6 months.

bModels included diabetes distress as the independent variable and potential mediators as dependent variables; covariates include age, race, health literacy, duration of diabetes, and baseline A\textsubscript{1c} variables.

cModels included potential mediators as the independent variable and hemoglobin A\textsubscript{1c} as the dependent variable; covariates include age, race, health literacy, duration of diabetes, and baseline A\textsubscript{1c} variables.

Table 3 presents the extent to which the association between improvement in HbA\textsubscript{1c} and the improvement in diabetes distress was mediated by goal setting or perceived competence (through the pathway that encompasses relationships a and b in Figure 1). We found that both goal setting and perceived competence are modest mediators with a combined 20% shared total effect (combined indirect effect −0.003, 95% CI −0.0072 to −0.0005).

**Table 3.** Mediating effects of goal setting and perceived competence in the relationship between diabetes distress and hemoglobin A\textsubscript{1c} (mediator analysis).

<table>
<thead>
<tr>
<th>Potential mediator\textsuperscript{a}</th>
<th>Indirect effect\textsuperscript{b} (95% CI)</th>
<th>Share of total effect (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal setting</td>
<td>−0.002 (−0.0052 to −0.0001)</td>
<td>13.3</td>
</tr>
<tr>
<td>Perceived competence</td>
<td>−0.001 (−0.0045 to −0.0002)</td>
<td>6.7</td>
</tr>
<tr>
<td>Combination of goal setting and perceive competence</td>
<td>−0.003 (−0.0072 to −0.0005)</td>
<td>20</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Goal setting and perceived competence assessed as the mean change from baseline to 6 months.

\textsuperscript{b}Covariates include age, race, health literacy, duration of diabetes, and baseline hemoglobin A\textsubscript{1c}.

**Results of the Moderator Analysis**

As shown in Table 4, the within-group estimates for the relationship between the change in diabetes distress and the change in HbA\textsubscript{1c} at 6 months was significant for participants who are younger than age 65 years, have more than a high school education, are employed, have an income greater than $30,000 per year, have lower health literacy, have more depressive symptoms, who have more social support, who have had diabetes for fewer years, and those with a baseline HbA\textsubscript{1c} <8.5%. The between group estimates suggest there is a significant difference in the relationship between the change in diabetes distress and the change in HbA\textsubscript{1c} at 6 months by race and the status of insulin use: stronger for whites compared with African Americans (P=.002) and for those who were using insulin compared with those not (P=.02).
Table 4. Adjusted estimates on the effect of improved diabetes distress on improved glycemic control, by groups with different baseline characteristics (moderator analysis).

<table>
<thead>
<tr>
<th>Potential moderator</th>
<th>N</th>
<th>Baseline mean diabetes distress (Predictor)</th>
<th>Baseline mean HBA1c&lt;sup&gt;a&lt;/sup&gt; (Outcome)</th>
<th>Adjusted estimates</th>
<th>β coefficient for change at 6 months (within subgroup)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>P value</th>
<th>Difference in β coefficients (between subgroups)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age in years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65</td>
<td>154</td>
<td>71.7</td>
<td>9.3</td>
<td>–0.019</td>
<td>0.002</td>
<td>0.007</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>&gt;65</td>
<td>136</td>
<td>74.9</td>
<td>8.8</td>
<td>–0.012</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>181</td>
<td>74.0</td>
<td>9.1</td>
<td>–0.006</td>
<td>.28</td>
<td>0.029</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>106</td>
<td>72.2</td>
<td>9.0</td>
<td>–0.035</td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;HS&lt;sup&gt;c&lt;/sup&gt;</td>
<td>12</td>
<td>77.8</td>
<td>8.8</td>
<td>0.024</td>
<td>.52</td>
<td>0.040</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>&gt;HS</td>
<td>278</td>
<td>73.0</td>
<td>9.1</td>
<td>–0.016</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Employment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None&lt;sup&gt;d&lt;/sup&gt;</td>
<td>213</td>
<td>74.6</td>
<td>9.1</td>
<td>–0.011</td>
<td>.19</td>
<td>0.008</td>
<td>.58</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>74</td>
<td>69.6</td>
<td>8.9</td>
<td>–0.018</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Income ($)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30,000</td>
<td>142</td>
<td>73.1</td>
<td>9.1</td>
<td>–0.012</td>
<td>.07</td>
<td>0.011</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>&gt;30,000</td>
<td>105</td>
<td>73.8</td>
<td>9.0</td>
<td>–0.023</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health literacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>152</td>
<td>70.4</td>
<td>9.1</td>
<td>–0.026</td>
<td>&lt;.001</td>
<td>0.018</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>138</td>
<td>76.3</td>
<td>9.1</td>
<td>–0.008</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline depression&lt;sup&gt;e&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>132</td>
<td>81.9</td>
<td>8.8</td>
<td>–0.013</td>
<td>.10</td>
<td>0.003</td>
<td>.64</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>158</td>
<td>66.0</td>
<td>9.3</td>
<td>–0.015</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline social support&lt;sup&gt;f&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>111</td>
<td>76.9</td>
<td>9.2</td>
<td>–0.012</td>
<td>.15</td>
<td>–0.004</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>130</td>
<td>72.2</td>
<td>9.0</td>
<td>–0.016</td>
<td>.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Duration of diabetes in years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10</td>
<td>111</td>
<td>71.4</td>
<td>9.3</td>
<td>–0.026</td>
<td>.006</td>
<td>0.016</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>179</td>
<td>74.3</td>
<td>8.9</td>
<td>–0.008</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline HBA1c (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;8.5</td>
<td>109</td>
<td>78.1</td>
<td>7.7</td>
<td>–0.021</td>
<td>.004</td>
<td>0.011</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>&gt;8.5</td>
<td>134</td>
<td>70.8</td>
<td>10.2</td>
<td>–0.010</td>
<td>.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Insulin use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>119</td>
<td>73.7</td>
<td>8.8</td>
<td>–0.006</td>
<td>.40</td>
<td>0.024</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>171</td>
<td>72.9</td>
<td>9.3</td>
<td>–0.029</td>
<td>.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>HBA1c: hemoglobin A1c.

<sup>b</sup>Adjusted for age, race, health literacy, duration of diabetes and baseline hemoglobin A1c except where these variables were tested as moderators.

<sup>c</sup>HS: high school.

<sup>d</sup>Includes not employed, retired and disabled.
improvement of the HbA1c due to the reduced levels of diabetes distress. Race was found to a moderator, suggesting that Caucasian veterans responded more to the peer support intervention than African American patients. Prior studies suggest that peer supporters who are culturally appropriate (including concordant age, race, and gender) may be more effective peer supporters for African Americans with diabetes [34,35]. Given that the burden of T2D falls heavily on minority populations, including African American and Latino populations [36], further studies are needed to understand the characteristics of effective interventions that target these high-risk populations, such as cultural concordance among peer supporters. Additionally, insulin use was found to be a moderator, suggesting that peer support interventions targeting high distress levels in insulin-requiring T2D patients lead to better glycemic control. This is important because approximately one-quarter of T2D patients in the United States currently require insulin, and this proportion is on the rise [37].

Strengths and Limitations
This study has several strengths. The first is that, to our knowledge, this is the first study looking at mediators and moderators between glycemic control and diabetes distress in an intervention that improves both. We incorporated robust statistical methods to assess the mediation pathway, finding that goal setting and perceived competence are important for future interventions targeting both glycemic and psychosocial outcomes for T2D. This is also one of the first studies to more specifically examine a broad array of socioeconomic and diabetes-specific characteristics that might moderate the relationship between diabetes distress and glycemic control. This is important because this can facilitate screening and targeted interventions using information readily captured by electronic medical records.

We also recognize that our study has several important limitations. First, this study was conducted in primarily African American male veterans with T2D, which limits the generalizability of our findings. It is therefore possible that, in other populations, goal setting and perceived competence have less significance in the mechanistic pathway between elevated levels of diabetes distress and worse glycemic control. Additionally, our use of brief validated scales to measure multiple complicated psychological constructs is a potential limitation, as these short-form scales did not permit in-depth investigation into different facets of these constructs. For example, we used the Diabetes Distress Scale 2 to measure diabetes distress, rather than the full 17-item Diabetes Distress Scale. Although the 2-item Diabetes Distress Scale has been found to correlate well with the larger Diabetes Distress Scale questionnaire, it does not provide subtypes of distress as it only measures emotional distress and this may have impacted our moderator analyses [27]. Prior studies indicate Black patients have higher levels of provider-related distress [38], which was not specifically measured in our study. It is possible that there are differences in the subtypes of diabetes distress (emotional burden, provider-related, interpersonal, and regimen-related)
among different populations (such as race/ethnicity) that account for the differential response in White versus Black participants in our study. The study population was also nearly exclusively male and does not therefore generalize to women with T2D, who often have higher levels of diabetes distress [39]. Future studies should include evaluation of interventions of women with T2D with high diabetes distress levels and use of more comprehensive scales to measure diabetes distress in order to more accurately generalize to all T2D populations. Finally, we hypothesized a priori that there would be 4 potential mediators and found that only goal setting and perceived competence were mediators. However, combined, these mediators only accounted for 20% of the mediation effect, suggesting that there are other important mediators in the mechanistic pathway between diabetes distress and glycemic control that we did not measure. Future studies are therefore needed to clarify these additional mediating mechanisms.

Conclusions

In conclusion, we found that in a peer support intervention for T2D in primarily African American male veterans both goal setting and perceived competence are important mediators in the mechanistic pathway between diabetes distress and glycemic control. Additionally, we found that this peer support intervention that improved diabetes distress was most effective in reducing HbA1c levels in White and insulin-requiring veterans with T2D. These findings are important for informing future interventions that target both psychosocial and glycemic outcomes and efforts to tailor interventions to best meet the needs of patients with different characteristics.

Acknowledgments

This research was supported by grants from the Veterans Affairs Health Services Research and Development Service (12-412) and the National Institute of Diabetes and Digestive and Kidney Diseases (P30DK092926 MCDTR).

Authors’ Contributions

KMS, HC, GP, and MH designed the study. HC and MH collected the data. KMS, HC, and CR analyzed the data. KMS wrote the first draft of the manuscript. KMS, HC, CR, GP, and MH edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Diabetes distress, goal setting, perceived competence, intrinsic motivation, and decisional conflict scales.

Multimedia Appendix 2
Summary of the change in diabetes distress, change in HbA1c, and hypothesized mediators between baseline and 6 months.

References


Abbreviations

- HBA1c: hemoglobin A1c
- RCT: randomized controlled trial
- T2D: type 2 diabetes
- VA: Veterans Affairs

Please cite as:

Mizokami-Stout K, Choi H, Richardson CR, Piatt G, Heisler M
Diabetes Distress and Glycemic Control in Type 2 Diabetes: Mediator and Moderator Analysis of a Peer Support Intervention
JMIR Diabetes 2021;6(1):e21400
URL: https://diabetes.jmir.org/2021/1/e21400
doi:10.2196/21400
PMID: 33427667

©Kara Mizokami-Stout, Hwajung Choi, Caroline R Richardson, Gretchen Piatt, Michele Heisler. Originally published in JMIR Diabetes (http://diabetes.jmir.org), 11.01.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), 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.