Early Adopters of Patient-Generated Health Data Upload in an Electronic Patient Portal

Jessica S. Ancker¹ Elizabeth Mauer¹ Robin B. Kalish² Joshua R. Vest³ J. Travis Gossey¹,4,5

¹Department of Healthcare Policy and Research, Weill Cornell Medicine, New York, New York, United States
²Department of Obstetrics and Gynecology, Weill Cornell Medicine, New York, New York, United States
³Department of Health Policy and Management, Indiana University Richard M. Fairbanks School of Health, Indianapolis, Indiana, United States
⁴Department of Medicine, Weill Cornell Medicine, New York, New York, United States
⁵Weill Cornell Physician Organization Information Services, Weill Cornell Medicine, New York, New York, United States

Address for correspondence Jessica S. Ancker, MPH, PhD, Department of Healthcare Policy and Research, Weill Cornell Medicine, 425 East 61st Street, New York, NY 10065, United States (e-mail: jsa7002@med.cornell.edu).


Abstract

Background and Objective Patient-generated health data (PGHD) may help providers monitor patient status between clinical visits. Our objective was to describe our medical center’s early experience with an electronic flowsheet allowing patients to upload self-monitored blood glucose to their provider’s electronic health record (EHR).

Methods An academic multispecialty practice enabled the portal-linked PGHD tool in 2012. We conducted a retrospective observational study of adult ambulatory patients using this tool between 2012 and 2016, comparing clinical and demographic characteristics of data uploaders with those of a group of patients with diabetes diagnoses and patient portal accounts seen by the same health care providers.

Results Over four years, 16 providers chose to use the tool, and 53 adult patients used it to upload three or more blood glucose values within any 9-month period. Of these patients, 23 were pregnant women and 30 were nonpregnant adults with diabetes. Uploaders had more encounters and portal log-ins than comparison patients but did not differ in socioeconomic status. Among the chronic disease patients, uploaders’ mean hemoglobin A1c and body mass index (BMI) both dropped significantly in the months after upload.

Conclusion Despite the potential value of PGHD in health care, the rate of adoption of a tool allowing patients to upload PGHD to their provider’s EHR has been slow. Among chronic disease patients, PGHD upload was associated with improvements in blood glucose control and BMI, but it is possible that the changes were because of increased motivation or intensive changes in medical management.

Keywords► blood glucose self-monitoring► self-management► patient portals► diabetes► obstetrics–gynecology

Background and Significance

Patient-generated health data (PGHD), that is, data recorded or generated by patients during their daily lives, could give clinicians a much more comprehensive understanding of patient health status than is available from clinical visits alone. It seems likely that a variety of types of PGHD could enrich clinical care, including exercise, activity, or diet tracking, through wearables and mobile devices¹ or through self-tracked data from personal medical devices such as blood pressure and blood glucose monitors²,³ Engaging patient in tracking and sharing PGHD could be a way to promote health knowledge, disease self-management skills,
and motivation, all of which are known to improve health outcomes. \textsuperscript{4–6}

Some important barriers stand in the way of integrating PGHD into clinical practice. The use of computers, smartphones, and other health information technology remains lowest among the patients in the greatest need, including people who are elderly, less educated, or less affluent. \textsuperscript{7–10} Although some communities embrace PGHD and data tracking, other people with chronic medical conditions may be deterred from self-tracking by feelings of frustration or anxiety in response to their PGHD data, which can serve as an unwelcome reminder of their poor health. \textsuperscript{11} On the provider side, physicians who have expressed interest in potentially using PGHD (for example, in diabetes) \textsuperscript{12} still require additional support before these data are likely to be useful. Providers are likely to need practice protocols to guide clinical responses to the data, visualizations, data reduction, or decision support to help make sense of the information while preventing overload, electronic health record (EHR) integration, and integration into clinical workflow. \textsuperscript{3,12,13} Nationality studies, feasibility studies, and demonstration projects to integrate PGHD with EHRs are ongoing. \textsuperscript{14,15}

Prior to these national studies, in 2012, our academic multispecialty practice sought to facilitate the use of PGHD by enabling an electronic data tracking tool called a flowsheet. The functionality was made available in the EHR where it could be adopted by providers as needed, but there was no specific implementation plan, encouragement, or incentivization to adopt this function. In this retrospective study, we sought to describe adoption rates and characteristics of early adopters of the PGHD functionality under these naturalistic conditions, along with preliminary data about associations with patient outcomes.

**Objectives**

In this retrospective study, our objectives were to describe adoption rates and characteristics of early adopters of the PGHD functionality under these naturalistic conditions and to collect preliminary data on associations with patient outcomes.

**Methods**

**Setting and Technology**

The Weill Cornell Physician Organization is the multispecialty faculty practice of Weill Cornell Medical College in New York City, representing both physician ambulatory offices and ambulatory hospital-based clinics. All physicians use the EpicCare electronic medical record and its integrated portal, which is branded as Weill Cornell Connect. Weill Cornell physicians see approximately 300,000 unique patients per year. Of the patients, 42% have electronic patient portal accounts. Providers are not salaried by Weill Cornell and participate in both public and private fee-for-service and managed-care plans. (Weill Cornell is affiliated to NewYork-Presbyterian Hospital, the largest hospital system in New York City, but inpatient electronic systems and databases are separate from the outpatient ones and are not described in this paper.)

The flowsheet functionality was enabled in 2012 and was available to every provider. A provider selects patients for whom it appears appropriate and introduces the flowsheet to them; patients cannot initiate the use of flowsheets by themselves. After a physician enables the flowsheet for a patient, the patient can use it to upload data securely through the electronic patient portal that is accessible either through a web browser or a smartphone app. Each blood glucose value is labeled with the time, and the patient can also enter insulin dose (if any), time of the insulin administration, and free text notes. Patients may upload up to several values per day or less frequently. The physician can then view the data within the EHR in a spreadsheet format.

When this flowsheet functionality was enabled, announcements went out to all EHR users, but the tool was not attached to any specific implementation plans, research studies, or quality initiatives.

**Methods**

We defined flowsheet adopters as adult patients who (1) had uploaded three or more blood glucose values within any 9-month period, (2) had at least a 2-year history at Weill Cornell before the first data upload, and (3) had logged into their patient portal account at least once before the first data upload. The comparison group comprised patients who had encounters with any of the physicians seen by the flowsheet adopter patients during the same time period, had any recorded diagnosis of diabetes or gestational diabetes, and had a 2-year history as well as at least one portal log-in before the date defined as the index date (see definition below). The 2-year history at Weill Cornell was chosen to ensure that there was comparable run-in data on diagnoses, clinical visits, and portal use for both flowsheet adopter and comparison patients. All variables were specified a priori, with the exception of the three-value upload threshold that was developed after inspecting the data.

The index date for the flowsheet adopter group was the date of first PGHD data upload. For the comparison group, the index date was set as the median time from between the first portal log-in and December 2016. We assessed physiological indicators (hemoglobin A1c, blood pressure, low-density lipoprotein cholesterol, and body mass index [BMI]) before and 9 months after the index date. Nine months was chosen to provide sufficient time for two to three hemoglobin A1c measurements.

The Johns Hopkins Aggregated Diagnostic Group algorithm, a case mix metric appropriate for ambulatory patients, was run on patient diagnosis codes to produce a chronic condition count. \textsuperscript{15} Patient addresses were mapped to census block to estimate patient socioeconomic status using the Centers for Disease Control’s Social Vulnerability Index–Socioeconomic Status (SVI-SES) theme, a composite metric of neighborhood-level income and education levels. \textsuperscript{17} Values range from 0 to 1, with 1 indicating the highest degree of vulnerability.
Statistical Analysis
Categorical data were compared using chi-square tests (or Fisher’s exact when cell sizes were smaller than 5). Continuous and count data were compared using independent sample t-tests when normally distributed or Wilcoxon’s tests when skewed. Analyses were conducted using SAS v. 9.4 (SAS Institute Inc., Cary, North Carolina, United States) and R v. 3.5.0 (R Foundation, Vienna, Austria).

Results
All providers had access to the flowsheets over 4 years, including approximately 250 attending physicians in internal medicine (including endocrinologists) and approximately 40 attending physicians in obstetrics and gynecology (ob-gyn). During these 4 years, 37,578 patients had a recorded diagnosis of diabetes, 8,851 (23.5%) of whom had portal accounts. Also, during this time, 1,395 patients had gestational diabetes, 823 (59%) of whom had portal accounts.

Over 4 years, 16 providers chose to use the blood glucose flowsheet, and 12 of them had patients who met our inclusion criteria. Of these 12, 4 were ob-gyn physicians (representing approximately 10% of the ob-gyn with access to the flowsheets), and 4 were physicians in the Department of Medicine (representing ~2% of the internists with access to flowsheet). One of the internists had a specialty of endocrinology. The remaining providers who used the flowsheets were nurses in the Department of Obstetrics and Gynecology (n = 2) and Medicine (n = 1), one individual with unknown credential in the Department of Obstetrics and Gynecology, and one individual missing both department and credential.

Fifty-three (53) patients uploaded three or more blood glucose values over any 9-month period. Of these patients, 23 were pregnant women (representing ~3% of portal users with gestational diabetes) and 30 were nonpregnant adults with diabetes (representing <1% of portal users with diabetes). A total of 201 nonpregnant comparison patients and 41 pregnant comparison patients were identified.

Pregnant Patients
The 23 pregnant patients who uploaded PGHD submitted a median of 201 (range: 23–573) values to six providers in the Department of Obstetrics and Gynecology. Uploaders were similar to comparison group patients in race, ethnicity, age, and socioeconomic status assessed through both SVI-SES and insurance type (Table 1). However, uploaders had more clinical encounters and portal log-ins before initial data upload, somewhat earlier establishment of patient portal accounts, and worse baseline blood pressure.

For pregnant patients, average BMI dropped significantly more in the 9 months after the index date (i.e., post-partum) among uploaders than nonuploaders. However, average blood pressure increased more among uploaders than nonuploaders.

Chronic Disease Patients
A total of 30 nonpregnant patients with diabetes diagnoses uploaded a median of 139.5 (15–1,253) values to 10 providers in the departments of Internal Medicine, Endocrinology, Cardiology, and Nutrition.

Among these chronic disease patients, uploaders were more likely than comparison group patients to be Asian-American and were younger, but the groups did not have other significant demographic differences including in socioeconomic status (Table 1). Uploaders had more clinical visits and portal log-ins before initial data upload.

During the 9-month period after the first glucose value upload, uploaders had significantly larger reductions in hemoglobin A1c and BMI than did nonuploaders (Table 1).

Uploaders also experienced large variations in hemoglobin A1c level after the index date (Fig. 1). Uploaders appeared to fall in two subsets (Fig. 2). One subset had low, well-controlled hemoglobin A1c values both before and during PGHD upload. A second subset began upload at a time when their hemoglobin A1c was elevated and typically experienced a hemoglobin A1c decrease followed by a plateau.

Discussion
Self-tracked or PGHD appears to offer the possibility of improved clinical monitoring for patients between clinical visits. Many patients who track data would like their physicians to review it to identify health problems or reassure them that all is well. Nevertheless, in an academic multispecialty practice with high rates of patient portal use, adoption of a PGHD data upload function has been slow. Over 4 years, 16 providers tried it, and 53 established patients uploaded three or more values.

The most marked difference between uploaders and comparison patients was that uploaders had more visits and portal log-ins. However, uploaders did not have more chronic conditions than nonuploaders and did not have markedly worse baseline physiological indicators (with the exception of higher systolic blood pressure among the pregnant subgroup). Overall, this suggests that uploaders were not substantively sicker than nonuploaders but rather more engaged in their health care, motivated to change behavior, or loyal to the medical center. They may also have been more facile with information technology, a point supported by the fact that uploaders among the chronic disease patients were younger than nonuploaders. Patient portal use has been shown to be associated with socioeconomic status, with more affluent patients having higher average health literacy and better access to computers, broadband internet, and smartphones. However, in our study, we included only patients with portal accounts and a history of using them. Among these patients, there were no socioeconomic differences between uploaders and nonuploaders, as indicated by insurance type or a composite measure of U.S. Census tract-level income and education.

Among the chronic disease patients, PGHD uploading was associated with marked variation in hemoglobin A1c values (Fig. 1) and significant reductions in hemoglobin A1c and BMI over a 9-month period (Table 1). It appears that uploading was associated with a period of medical transition or change. However, given the baseline differences between the two groups, it would be difficult to conclude that the uploading behavior caused the reduction.
Among pregnant patients, BMI is expected to rise during pregnancy, but we found that after birth, BMI dropped more among uploaders than nonuploaders. By contrast, blood pressure increased more among uploaders than comparison patients. This could be an artifact of the fact that PGHD patients were slightly sicker (more than two comorbidities compared with approximately 1; \( p = 0.1 \)) or that some patients with gestational diabetes may have had comorbid preeclampsia. Hemoglobin A1c is not typically monitored during pregnancy, and therefore there were little data on this.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of patient-generated health data uploaders and comparison with patients who did not upload</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Chronic disease patients</strong></td>
</tr>
<tr>
<td></td>
<td>Uploaders ((n = 30))</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
</tr>
<tr>
<td></td>
<td>Black</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
</tr>
<tr>
<td></td>
<td>All other</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Hispanic</td>
</tr>
<tr>
<td></td>
<td>Not Hispanic</td>
</tr>
<tr>
<td></td>
<td>All other</td>
</tr>
<tr>
<td>Age</td>
<td>39.2 (14.4) 47 (15.2)</td>
</tr>
<tr>
<td>Payer</td>
<td>Commercial</td>
</tr>
<tr>
<td></td>
<td>Medicaid</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>Medicare</td>
</tr>
<tr>
<td>Chronic condition count</td>
<td>5.9 (4.7) 6 (4.5)</td>
</tr>
<tr>
<td>SVI-SES</td>
<td>0.27 (0.30) 0.24 (0.26)</td>
</tr>
<tr>
<td>Encounters year before index</td>
<td>10.2 (4.6) 6.5 (5.3)</td>
</tr>
<tr>
<td>Portal log-ins year before index</td>
<td>65.9 (47.6) 23.5 (28.1)</td>
</tr>
<tr>
<td>Year of first encounter</td>
<td>2,008 (3.3) 2,008 (3.1)</td>
</tr>
<tr>
<td>Year established portal account</td>
<td>2,012 (1.6) 2,012 (1.9)</td>
</tr>
<tr>
<td>Baseline hemoglobin A1c (%)</td>
<td>7.8 (3.1) 6.8 (1)</td>
</tr>
<tr>
<td>Baseline systolic blood pressure</td>
<td>117.2 (13) 118.5 (16)</td>
</tr>
<tr>
<td>Baseline diastolic blood pressure</td>
<td>72.3 (10.5) 71.8 (9.2)</td>
</tr>
<tr>
<td>Baseline LDL</td>
<td>118.5 (38.1) 101.1 (33.2)</td>
</tr>
<tr>
<td>Baseline body mass index</td>
<td>29.2 (6.8) 28.2 (6.2)</td>
</tr>
<tr>
<td>Change in hemoglobin A1c (%)</td>
<td>-0.8 (1.9) 0.1 (0.8)</td>
</tr>
<tr>
<td>Change in systolic blood pressure</td>
<td>5.5 (16.7) 1.6 (16.1)</td>
</tr>
<tr>
<td>Change in diastolic blood pressure</td>
<td>3.8 (10) 0.9 (9.4)</td>
</tr>
<tr>
<td>Change in LDL</td>
<td>2.1 (33.5) 0.1 (23.9)</td>
</tr>
<tr>
<td>Change in body mass index</td>
<td>-1.4 (2.2) 0.2 (1.6)</td>
</tr>
</tbody>
</table>

Abbreviations: LDL, low-density lipoprotein; NA, not available; SVI-SES, Centers for Disease Control’s Social Vulnerability Index–Socioeconomic Status.

*As these indicators are not typically monitored multiple times during pregnancy, fewer than 10 pregnant patients had both baseline and follow-up values for LDL and hemoglobin A1c.

Note: p-values of 0.05 or lower are shown in boldface.
The low adoption rate for PGHD upload is probably attributable to several factors. The first factor is barriers to provider uptake. It is widely accepted that integration into the EHR is a prerequisite for the use of novel tools, and, in other situations, lack of integration is cited as a barrier to adoption.\textsuperscript{3,22} However, even though this PGHD upload tool was integrated into the EHR, providers would still have had to select patients for whom it would be appropriate, determine if the patients are interested and train them to use it, and then would presumably feel responsible for checking uploaded values regularly to provide feedback. In our center, rollout of the new function was announced across the institution but was not accompanied by a financial or other incentive program that would have covered the time costs of online monitoring or by an implementation plan with clinical champions to promote use. In this naturalistic situation, providers who adopted the new function were those who were already intensively monitoring patients. This included physicians and nurses in the Department of Obstetrics and Gynecology, which has protocols for routine screening for gestational diabetes and preeclampsia at 24 to 28 weeks of gestation. For women with high-risk pregnancies, physicians promoted daily self-monitoring even before the electronic PGHD tool was enabled. Patient-recorded values are typically checked regularly by a physician or a nurse, and patients are encouraged to call when values are concerning. By contrast, there is no standardized approach to PGHD data upload for patients with chronic diabetes; therefore, uptake and workflow depend upon provider preference. In this group, a small number of providers used the flowsheets for patients at a time of hemoglobin A1c elevation, which may have been a time when patients were motivated to attempt lifestyle changes or when providers were making and monitoring changes in medical management. Others have noted that even providers who recognize the value of PGHD still express the need for support such as practice protocols.\textsuperscript{3}

A second explanation for low adoption rates is that PGHD upload is likely to require active and continued buy-in from patients. Our inclusion criteria specified that patients must already have a patient portal account. However, it is also likely that PGHD would be adopted only by patients who were willing to measure and input glucose values regularly and to check provider feedback through portal messaging.\textsuperscript{23} Self-monitoring of blood glucose can be challenging both physically and emotionally. In some cases, patients may find self-monitoring helpful in understanding the impact of their lifestyle and increasing their sense of control. However, in others, it may lead to feelings of burden, frustration, anxiety, or self-blame, especially when values are not ”good” or when trends do not clearly link to lifestyle changes.\textsuperscript{11,24} This is likely to reduce the numbers of patients willing to commit to routine PGHD upload. It may also lead to abandonment after trying the tool. In many domains, people abandon personal tracking tools after they have satisfied their curiosity, when their health status or life circumstances change, or when tracking starts to become burdensome.\textsuperscript{25,26} In other clinical domains, sustained engagement with self-monitoring has also been linked to strong relationships with the health care providers.
provider, supportive environments, and regular feedback. In our study, another probable barrier was the need for manual data input. Although, in theory, patients could upload glucose values directly from monitors through Apple Healthkit, the data did not allow us to determine how many patients took advantage of this potential convenience, and it is not clear how easy it was for them to do this. In the future, widespread integration of PGHD into clinical care will likely depend upon (1) targeted implementation strategies including measures to ensure that physicians are fully incentivized to collect and review remotely collected PGHD data rather than prioritizing face-to-face visits, (2) modifications that make technologies easier for patients to use, such as by eliminating manual data input, (3) disease- or population-specific decision support that uses the PGHD, and (4) a stronger evidence base demonstrating ways in which the time and energy needed to generate, review, and provide feedback on PGHD measurably improve health, satisfaction, or other important outcomes. This will be important to ensure that patients are helped to track information of personal value to them rather than data to be used to meet quality measurement or value-based purchasing initiatives.

Limitations
The study is limited by its small sample size and unmeasured confounders such as health care engagement, literacy, patient activation, and motivation, which could account for differences between uploaders and non-uploaders. The sample size meant that it was not feasible to apply causal inference methods such as propensity score matching. Because of these problems, we also avoided conducting regression analysis. This makes it impossible to rule out the possibility that the confounders (engagement with health care and patient age) could be responsible for the observed improvement in patient outcomes. The heterogeneity of the settings and of the patients and clinicians in this particular dataset also makes it challenging to determine whether this technology might be appropriate for certain subsets. Targeted studies with selected patient populations would be needed to develop estimates of efficacy of this technology; the current analysis of this natural experiment is useful only for showing its limited effectiveness in this specific setting. Also, this short-term analysis does not provide an evidence on long-term effects or sustainability.

Conclusion
Despite the potential for PGHD to help medical monitoring of patients between visits, adoption of a PGHD function within an academic multispecialty practice was slow. The function appeared to be adopted by providers who sought to monitor patients in a period of medical instability or transition, such as, women during high-risk pregnancy and chronic disease patients after an elevated hemoglobin A1c measurement. Patients who uploaded PGHD had more visits and portal uploads than patients who did not, but they did not appear markedly sicker than non-uploaders, suggesting that they may have been more engaged with their health care or motivated. Chronic disease patients who uploaded PGHD data experienced improvements in their hemoglobin A1c and BMI values. However, the improvement may have been because of motivation or increased engagement, meaning that PGHD upload may be a support tool for patients who are already motivated rather than one that directly improves outcomes.

Clinical Relevance Statement
Self-tracked or PGHD appears to offer the possibility of improved clinical monitoring for patients between clinical visits and may be associated with improved outcomes. Nevertheless, physicians may be slow to embrace PGHD data upload to the EHR unless the process is streamlined, incentivized or reimbursed, or promoted through targeted implementation efforts.

Multiple Choice Questions
1. Diabetes patients who uploaded PGHD to a patient portal showed improvement in their hemoglobin A1c values. Nevertheless, the effect might be because of the following:
   a. Higher socioeconomic status among these patients.
   b. Fewer clinical visits among these patients.
   c. Higher health care utilization among these patients.
   d. Improved medication safety among the patient.

   **Correct Answer:** The correct answer is option c, higher health care utilization among these patients. Although the patients did have improved hemoglobin A1c, they also were more frequent users of the portal and more frequent users of the health care system than non-uploaders. This could be the explanation for their improved outcomes. By contrast, their neighborhood-level socioeconomic status was tested and did not show any significant difference; therefore, this is unlikely to be the explanation for the findings.

2. Patient-generated health data (PGHD) is a term that encompasses which of the following types of information:
   a. Patients answering questions about the quality of their health care.
   b. Patients tracking health indicators of interest to themselves.
   c. Patients describing their own health status.
   d. Patient satisfaction.

   **Correct Answer:** The best answer is option b, patients tracking health indicators of interest to themselves. The other options (a, c, and d) are instead considered patient-reported outcomes.

Protection of Human and Animal Subjects
This study was approved by the Weill Cornell Institutional Review Board with a waiver of consent for use of de-identified data.

Funding
This project was funded by the Agency for Healthcare Research and Quality (K01 HS021531; PI: Ancker).
Conflict of Interest
None declared.

References
3 Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from Project HealthDesign. JMIR Human Factors 2016;3(02):e26
5 Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013;32(02):207–214
10 Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018;25(08):1080–1088