



Published in final edited form as:

Med Care. 2014 June ; 52(6): 511–518. doi:10.1097/MLR.000000000000131.

Re-Thinking the Frequency of Between-Visit Monitoring for Patients with Diabetes

John D. Piette, MSc, PhD^{1,2,3}, James Aikens, PhD⁴, Ann Marie Rosland, MD, MSc^{1,3}, and Jeremy B. Sussman, MD^{1,3}

John D. Piette: jpiette@umich.edu; James Aikens: aikensj@umich.edu; Ann Marie Rosland: arosland@umich.edu; Jeremy B. Sussman: jeremysu@umich.edu

¹Center for Clinical Management Research, VA Ann Arbor Healthcare System, Ann Arbor, Michigan, USA

²Department of Health Behavior and Health Education, University of Michigan School of Public Health, Ann Arbor, Michigan, USA

³Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA

⁴Department of Family Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA

Abstract

Background—Health systems increasingly look to mobile health (mHealth) tools to monitor patients cost-effectively between visits. The frequency of assessment services such as interactive voice response (IVR) calls is typically arbitrary, and no approaches have been proposed to tailor assessment schedules based on evidence regarding which measures actually provide new information about patients' status.

Methods—We analyzed longitudinal data from over 5000 weekly IVR monitoring calls to 298 diabetes patients using logistic models to determine the predictability of IVR-reported physiologic results, perceived health indicators, and self-care behaviors. We also determined the implications for assessment burden and problem detection of omitting assessment items that had no more than a 5% predicted probability of a problem report.

Results—Assuming weekly IVR assessments, episodes of hyperglycemia were difficult to predict (area under the curve [AUC] = 69.7; 95% CI: 50.2, 89.2) based on patients' prior assessment responses. Hypoglycemic symptoms and fair/poor perceived health were more predictable, and self-care behaviors such as problems with medication adherence (AUC=92.1; 95% CI: 89.6, 94.6) and foot care (AUC=98.4; 95% CI: 97.0, 99.8) were highly predictable. Even

Send correspondence to: John D. Piette, Center for Clinical Management Research HSR&D Center of Excellence, VA Ann Arbor Healthcare System, P.O. Box 130170, Ann Arbor, Michigan 48113-0170, United States; jpiette@umich.edu. Fax: 734-936-8944. Phone: 734-998-5922.

Address for Drs. Piette, Rosland, and Sussman, University of Michigan, North Campus Research Complex, 2800 Plymouth Road, Building 16, 3rd floor, Ann Arbor, MI 48109-2800

Address for Dr. Aikens, Department of Family Medicine, University of Michigan Health System, 1018 Fuller SPC 5708, Ann Arbor, MI 48109-2800

Telephone for all authors: 734-936-5216

Fax for all authors: 734-936-8944

if patients were only asked about foot inspection behavior when they had >5% chance of a problem report, 94% of foot inspection assessments could be omitted while still identifying 91% of reported problems.

Conclusions—mHealth monitoring systems could be made more efficient by taking patients' reporting history into account. Avoiding redundant information requests could make services more patient-centered and might increase engagement. Time saved by decreasing redundancy could be better spent educating patients or assessing other clinical problems.

Keywords

telemedicine; care management; chronic disease; diabetes

Introduction

Patients with diabetes often experience fluctuations in their health status, and frequent monitoring with feedback is the hallmark of proactive disease management systems.^{1–3} In addition to self-monitored blood glucose (SMBG) levels and symptoms of poor glycemic control, patients' overall perceived health status is a significant predictor of long-term disease risk and an important indicator for diabetes care management.⁴ An ongoing dialogue about self management behaviors such as medication adherence and foot inspection is considered central to effective diabetes self-management support.⁵

Although communication about patients' health status and self-care between face-to-face encounters is vital, frequent contacts with patients by care managers and health educators can be costly. As a consequence, health systems increasingly seek cost-effective alternatives to automate the process of patient contact via mobile health (mHealth) tools such as text messaging, smart phone “apps,” and interactive voice response (IVR) telephone calls.⁶ IVR calls with tree-structured calling algorithms are particularly attractive because they make possible a rich, patient-specific exchange of information while avoiding barriers associated with other approaches, such as limitations in patients' ability to text or the expense of a smartphone and data plan.⁷ Clinical data collected via IVR is reliable and valid,^{8–11} and IVR self-management support services can improve outcomes including hospitalization rates, weight management, blood pressure control, and psychiatric symptoms.^{6,12–15}

The frequency of assessments using mHealth tools is typically arbitrary – similar to the frequency of regular diabetes follow-up visits, “live” care management calls, and many other health services used to monitor patients' status.^{16–19} Lengthy assessment intervals could miss important health problems, while overly frequent assessments may cause patient dissatisfaction and program dropout. For patients with chronic diseases such as diabetes, excessively frequent assessments for one condition also may generate an unmanageable amount of clinical feedback while potentially missing opportunities to monitor comorbidities that represent greater threats to the patient's health.²⁰ To our knowledge, no studies have suggested an empirically based method for personalizing the frequency and content of mHealth monitoring for chronic conditions. In a recent study,²¹ we found that weekly IVR calls assessing depression-related outcomes provided valid and reliable clinical

information, but may produce a significant amount of data that could be inferred based on the patient's prior assessments and baseline characteristics.

In this study, we examined data on more than 5000 weekly IVR assessment reports for 298 patients with type 2 diabetes who participated in the CarePartner IVR care management program. The primary goal of the study was to determine the extent to which patients' IVR-reported assessment data could be predicted using information from their IVR reporting history and information collected at program entry. Because monitoring calls included questions about participants' glycemic control, perceived health status, and self-management behaviors, the study provides an opportunity to determine whether some of these domains should be monitored more frequently while others may be stable enough to monitor more occasionally. More generally, this study provides new insights regarding the relative value of time spent in telephone chronic disease management programs conducting repeated patient assessments, as opposed to devoting that time to educating patients about diabetes self-management, supporting behavior change, or assessing other clinical problems.

Methods

Patient Eligibility and Recruitment

Patients with type 2 diabetes were enrolled from 16 Department of Veterans Affairs (VA) primary care clinics. To be eligible, patients had to have: an ICD-9 coded diagnosis of type 2 diabetes; 1+ VA primary care visit in the prior 12 months; and a current prescription for antihyperglycemic medication. Patients with diagnoses indicating cognitive impairment, serious mental illness, schizophrenia, or bipolar disorder were excluded. Potential participants were initially contacted by mail and then screened for eligibility by telephone. After providing informed consent, patients were mailed additional program information, including materials describing effective communication with informal caregivers and clinicians. Patients were enrolled in two waves, with the first wave receiving IVR calls weekly for three months and the second receiving weekly calls for six months. The study was approved by the Ann Arbor VA and University of Michigan human subjects committees.

Program Features

The goals of the IVR intervention were to: (1) monitor patients' symptoms and self-management problems and provide tailored information about diabetes self-care based on those reports; (2) provide clinicians with actionable feedback via fax about patients' IVR-reported problems; and (3) provide informal caregivers with feedback about patients' status via structured emails along with guidance on supporting diabetes self-management. Each week that an assessment was scheduled, the system made up to three attempts to contact patients on up to three different patient-selected day/time combinations. Patients responded to requests for information using their touch-tone cellular or landline telephone. As reported previously,⁷ patients completed IVR assessments on more than 80% of the occasions for which one was attempted.

Data Collection

Baseline telephone surveys included information about participants' sociodemographic characteristics, including their age, gender, race, and educational attainment; and self-reported diagnoses of chronic diseases. To measure patients' baseline severity of illness, we created a count of the number of comorbid conditions including: hypertension, heart disease (prior myocardial infarction, atherosclerosis, angina, or heart failure), cancer, stroke, chronic pain (arthritis or chronic low back pain), pulmonary disease (COPD or asthma), and osteoporosis. Surveys also included: the Problem Areas in Diabetes (PAID) measure of diabetes-related distress,²² the Medical Outcome Study 12-Item Short Form (SF-12),²³ and the Morisky Medication Adherence index.

Assessment questions and tailored educational messages delivered during IVR calls were developed by experts in mobile health message design, diabetes education, endocrinology, and diabetes management in primary care. The content of IVR assessments was identical over time and across patients, and designed to be consistent with the suggested focus for diabetes education programs certified by the American Diabetes Association.²⁴ Assessments focused on three areas: patients' physiologic health (high and low SMBG levels), perceived health (hyperglycemic symptoms, hypoglycemic symptoms, and perceived general health), and health behaviors (days in bed due to health problems, diabetes medication adherence, foot inspection, and SMBG). For patients with hypertension who monitored their blood pressure at home, additional items were included to assess high and low levels of systolic blood pressure and antihypertensive medication adherence. Hypertension-related analytic results are included as an online supplement (see online Tables 1–2 and Figures 1–2). Exact wording of each assessment item and other information provided during the IVR calls are available from the authors on request.

We created the following binary indicators of patients' health and self-care problems: two or more instances in the past week of morning SMBG values > 300 mg/dL (among patients reporting SMBG); two or more instances in the past week of morning SMBG values *> 90* (among patients reporting SMBG); any "symptoms of high blood sugar in the past week, sometimes called hyperglycemia, such as increased thirst, blurred vision, or frequent urination"; any "symptoms of low blood sugar in the past week, sometimes called hypoglycemia, such as sweating, trembling, or feeling so weak that you were going to fall down"; "fair" or "poor" overall health on the day of the assessment; ever staying in bed all or most of the day because of health problems during the past week; not "always" taking diabetes medication as prescribed in the past week; not "checking [your] feet for sores, blisters and cuts every day"; and not "checking [your] own blood sugar in the morning before eating or having another person check [your] blood sugar."

Analytic Approach

We used t-tests and chi-squared tests to identify significant differences in baseline scores for the PAID, MCS, PCS, and Morisky Adherence scale between groups who ever did versus did not report each health and self-care problem during one or more of their IVR assessments. We used data at the level of the assessment occasion to predict IVR-reported health and self-care problems within and across assessments. First, we used cross-

tabulations to examine the bivariate association between each problem and other problems reported during the same IVR assessment. Next, we used contingency tables to determine the extent to which patients' IVR-reported problems were repeated across multiple, successive contacts. Specifically, we identified all completed assessments that were linkable to five prior completed weekly assessments. Using these data, we examined the proportion of patients reporting each health or self-care problem, given that the problem was reported in zero to five of those prior assessments. Statistically significant bivariate associations were identified using F-tests that adjusted for the clustering of repeated assessments within patient.

We used multivariable logistic models to predict assessment responses using: the five prior reports of that same health or self-care problem, other health and self-care problems reported during prior assessments, and information collected at baseline. For each outcome, we created a series of multi-level, hierarchical models with one observation for each "index" assessment linked to prior assessment results used in the prediction. Variance estimates of the predictors were adjusted for the within-patient correlation using robust variance estimation via the "vce(robust)" option in Stata which calculates the Huber-White sandwich estimates of the variances.^{25,26} Robust variances treat the higher-level effects (i.e., patient effects) as a nuisance factor and give accurate assessments of the variation in parameter estimates across patients even when the model is misspecified.²⁷⁻²⁹ Initial models considered three levels, with index assessments nested within patients and patients nested within recruitment sites. However, we observed no correlation attributable to site and the final models were calculated using two levels.

When fitting each model, we used two strategies to prevent over-fitting to the current dataset. First, we used 10-fold cross validation, in which the model was fit 10 times based on random 90% training samples, and then used to predict the outcomes in mutually-exclusive 10% test samples. Second, for each of the ten replications, we used stepwise regression (with a p-value of .20 for removal) to identify the most significant subset of predictors. To investigate the relationship between assessment frequency and the predictability of patients' reports, we used an identical process, but with index assessments linked to up to five prior assessments that had at least a 14-day and a 21-day gap between each completed IVR call.

The ability of each model to predict a given response within the index assessment was evaluated using Receiver Operator Characteristic (ROC) curves and measures of the area under the curve (AUC), along with each AUC's 95% confidence interval.³⁰ Also, we used the predicted probabilities from the fitted models to determine the implications of skipping IVR assessments for each outcome when a model predicted that the probability of a problem report was 5%. Specifically, using the 5% cutoff, we calculated: the proportion of index assessments that would be omitted, the proportion of reported problems that would still be captured after omitting those assessments (i.e., the sensitivity of the 5% cutoff for identifying subsequent problems), and the proportion of negative assessment reports that would be omitted (i.e., the specificity of the 5% cutoff for identifying patients who were unlikely to subsequently report a problem). All analyses were conducted using Stata version 11.2.

Results

Patient Characteristics

A total of 72% of patients screened by telephone consented to participate in the program. The sample of 298 patients was 97% male, largely Caucasian (60.8%), and with 47% reporting no more than a high school education. Eight-two percent of the sample was at least 60 years old, with 30% at least 70 years of age. Patients reported a mean of 2.7 (SD: 2.4) comorbid chronic conditions including hypertension (81.9%), chronic pain (68.0%), cardiovascular disease (45.5%), pulmonary disease (26.6%), cancer (17.1%), and osteoporosis (4.3%). The sample reported relatively poor physical function scores (baseline physical composite summary score from the SF-12 mean, SD: 32.3, 12.3), but relatively good psychological function (baseline mental composite summary mean, SD: 50.1, 11.7; PAID mean, SD: 13.0, 12.8). A total of 30.6% of patients had baseline scores of ≤ 2 on the Morisky medication adherence scale, indicating significant problems adhering to their antihyperglycemic regimens.³¹

Patients enrolled in the three-month program completed calls for an average of 11.1 (SD: 1.1) weeks; patient in the six-month program completed calls for an average of 26.8 (8.1) weeks. Patients in the analytic sample completed assessments on 5,059 patient-weeks in which one was attempted or 83.4 percent (SD: 19.2; median: 91.3 percent). Completion rates were lower among patients with: less physical impairment at enrollment as measured by the SF-12 ($p=.04$), elevated diabetes-related distress ($p=.007$), or reports of medication non-adherence at baseline. Patients had 3,831 index assessments that could be linked to five prior assessments assuming a weekly assessment frequency, 249 patients had 2,737 index assessments that could be linked with five prior assessments that had at least a 14-day gap between each one, and 133 patients had 1,274 index assessments linkable with five prior assessments having a 21+ day gap between each one.

Patients reported a variety of health or self-care problems at least once during their IVR assessments (Table 1). At least half of participants reported at least once that they had experienced: (a) ≥ 2 instances in the prior week of SMBG values > 90 mg/dL (56% of participants), (b) hyperglycemic symptoms (63% of participants), (c) hypoglycemic symptoms (63%), and (d) not always taking their antihyperglycemic medication exactly as prescribed (62%). Patients reporting hypoglycemic symptoms had lower baseline PAID scores, as well as poorer MCS, PCS, and Morisky adherence scores (all $p < .05$). Similarly, scores on each baseline measure were worse among patients who reported symptoms of hyperglycemia, fair/poor health, and days in bed due to health problems during their IVR calls (all $p < .05$).

Bivariate Associations between IVR-Reported Health and Self-Care Problems

Reported health and self-care problems were significantly associated within assessments. After adjustment for the clustering of assessments within patients, patients reporting SMBG values > 300 mg/dL were: more than four times more likely than other patients to simultaneously report symptoms of hyperglycemia (65% versus 14%), more than four times more likely to report days in bed due to poor health (29% versus 7%), and more than twice

as likely as other patients to report fair or poor health (49% versus 19%; all $p < .001$). Patients reporting not taking their diabetes medication as prescribed were more likely than other patients to simultaneously report: days in bed due to poor health (13% versus 7%), symptoms of hyperglycemia (28% versus 18%), and not checking their feet (14% versus 8%; all $p < .05$). Twenty-six percent of patients reporting fair or poor health simultaneously reported one or more days in bed in the prior week, compared to 4% of patients reporting that their health was good, very good, or excellent ($p < .001$).

Patients' responses were highly stable across assessments (Figure 1). For example, using the sample of index assessments with five or more prior consecutive weekly assessments, we identified 348 assessments for which patients reported medication adherence problems in all five of their recent IVR calls. In 94% of those assessments, patients once again reported nonadherence. As another example, 87% of assessments had no reports of less than daily foot inspection in any of the prior five IVR calls; and during those index assessments, patients reported less than daily foot inspection only 0.5% of the time.

Predictability of Assessment Results Using Multivariate Models

As shown in Table 2 and Figure 2, instances in which the patient reported at least two morning SMBG values exceeding 300 mg/dL during an assessment were poorly-predicted by prior weekly IVR assessments and baseline information (AUC: 69.7; 95% CI: 50.2, 89.2). In contrast, patient reports of symptoms and perceived general health were predicted with good accuracy; and health behaviors, particularly medication adherence (AUC: 92.1; 95% CI: 89.6, 94.6) and foot inspection (AUC: 98.4; 95% CI: 97.0, 99.8) were very well predicted. As expected, more frequent IVR reports tended to have greater predictive power than reports collected only every three weeks; however in many instances there was no significant difference in the AUC when comparing more versus less frequent IVR assessment calls. The pattern of regression coefficients across logistic models (data not shown) suggested that prior information about a given health or behavioral problem was consistently the most significant predictor of patients' subsequent responses.

Analyses suggested that many repeat assessment items could be omitted without significant loss of information (Table 2). For example, 52% of hypoglycemia assessments could have been omitted if that item were dropped for patients whose prior data indicated less than a 5% probability of a problem report, and yet 85% of the instances of reported hypoglycemia symptoms would still be identified. Also as Table 2 shows, 94% of the foot inspection assessments could have been omitted if the item was used only when patients' prior data indicated >5% chance of reporting a problem, and the monitoring system still would detect 91% of all reports of less-than-daily foot inspection.

Discussion

This study provides important insight into the extent to which serial mHealth assessments provide useful new information about diabetes patients' status, as opposed to replicating information that could have been predicted based on prior IVR reports and baseline characteristics. In general, we found that reports of morning hyperglycemia were rare and therefore difficult to predict. In contrast, reports of frequent hypoglycemia and fair/poor

general health were common and predictable, and therefore may not need to be monitored as frequently. Self-care behaviors such as diabetes medication adherence and foot inspection also were highly predictable after the patient reported problems in these areas a few times.

Redundant assessments represent major lost communication opportunities, and the time could be better spent assessing other comorbid health problems that have either a higher probability of occurrence or a greater risk to the patient's overall health. Alternatively, contact time saved by avoiding redundancies could be devoted to delivering more targeted self-management education. In the long-run, patients likely will find it burdensome if they experience excessive follow-up of health and behavioral issues that are unchanging, and this may lead to program dissatisfaction and dropout.

Even for highly predictable outcomes, a small number of health events will invariably be missed if patients are asked less frequently about that potential event. In any given system, designers must make explicit decisions about the trade-offs between the potential "payoff" of repeated assessments of a given health parameter, relative to the cost in terms of potentially increased patient burden, dissatisfaction and dropout. It may be that other measures of predictability may be preferable in some instances, depending on patients' clinical characteristics and whether the feedback to clinicians is actionable given health system constraints. For example, even if overall predictability of patients' responses is high, clinicians may weigh false positive versus false negatives differently for health indicators such as hypoglycemia which could pose a significant risk. The current analyses presents a framework for making those choices explicit, rather than using rote decision rules developed arbitrarily by system designers. This approach could easily incorporate novel measures of performance that take clinical usefulness into account.³²

It remains possible that repeatedly asking patients about self-management activities such as foot inspection or SMBG may have the benefit of reinforcing these important behaviors.³ Future mHealth research should investigate whether such repetition is a worthwhile approach to eliciting and sustaining health behaviors compared to more tailored, in-depth messages that address patients' motivational and other barriers to self-care.

This study has several limitations that may affect the predictability of patients' IVR-reported health and self-care problems. Patients were older, mostly Caucasian males receiving care in the VA healthcare system. Patients who are uninsured or recently diagnosed with diabetes may have a less stable disease course and more difficulty accessing diabetes management services and medications – both factors that could impact the predictability of their repeated IVR assessment responses. Also, it is important to note that the predictive accuracy of these models may vary across important subgroups of patients. For example, patients with prior hospitalizations and missed clinic appointments are more likely to skip IVR assessments;⁷ and disease severity as measured by A1c levels, insulin use, or other factors may affect the time lag between completed assessments or the consistency of responses between them. Unfortunately data on patients' glycemic control and medication use were not available for analysis.

As mHealth monitoring and self-management support systems become increasingly integrated with electronic medical records, new approaches will be possible that identify the most appropriate content of each interaction utilizing techniques such as “machine learning,” in which a system modifies its own functional parameters based upon the data that it collects.³³ Techniques such as those used in the current study could identify when assessment responses become highly predictable, so that subsequent assessments could either shift to patient education or the assessment of comorbid health problems. Data mining techniques also could be used to identify the most prognostically useful patient-reports with respect to subsequent urgent events or complications.

As mHealth services become more prevalent, we need to re-think how we design mobile health monitoring, so that systems collect the most important information about patients’ status while avoiding missed opportunities to use patient contact time more effectively for promoting behavior change or attending to important comorbidities. The next generation of mHealth services should use machine learning and other forms of artificial intelligence to automatically adapt to each patient’s unique and evolving needs.

Acknowledgments

John Piette is a VA Senior Research Career Scientist. Ann Marie Rosland is VA Career Development Awardees. Other financial support came from grants number P30DK092926 and 1-R18-DK-088294-01-A1 from the National Institute of Diabetes and Digestive and Kidney Diseases.

References

1. Funnell MM, Brown TL, Childs BP, et al. National standards for diabetes self-management education. *Diabetes Care*. 2012; 25 (Suppl 1):S101–S8. [PubMed: 22187467]
2. DiClemente CC, Marinilli AS, Singh M, Bellino LE. The role of feedback in the process of health behavior change. *American Journal of Health Behavior*. 2001; 25:217–27. [PubMed: 11322620]
3. McClure JB. Are biomarkers useful treatment aids for promoting health behavior change? *American Journal of Preventive Medicine*. 2002; 22:200–7. [PubMed: 11897465]
4. Grant MD, Piotrowski ZH, Chappell R. Self-reported health and survival in the longitudinal study of aging, 1984–1986. *Journal of Clinical Epidemiology*. 1995; 48:375–87. [PubMed: 7897459]
5. Aikens JE, Piette JD. Longitudinal association between medication adherence and glycaemic control in type 2 diabetes. *Diabetic Medicine*. 2013; 30:338–44. [PubMed: 23075262]
6. Free C, Phillips G, Galli L, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLOS Medicine*. 2013; 10:e1001362. [PubMed: 23349621]
7. Piette JD, Rosland AM, Marinec NS, Striplin D, Bernstein SJ, Silveira MJ. Engagement with automated patient monitoring and self-management support calls: experience with a thousand chronically-ill patients. *Medical care*. 2013; 51:216–23. [PubMed: 23222527]
8. Searles J, Perrine M, Mundt J, Helzer J. Self-report of drinking using touch-tone telephone: extending the limits of reliable daily contact. *Journal of studies on alcohol*. 1995; 56:375–82. [PubMed: 7674671]
9. Piette JD, McPhee SJ, Weinberger M, Mah CA, Kraemer FB. Use of automated telephone disease management calls in an ethnically diverse sample of low-income patients with diabetes. *Diabetes Care*. 1999; 22:1302–9. [PubMed: 10480775]
10. Turvey CL, Willyard D, Hickman DH, Klein DM, Kukoyi O. Telehealth screen for depression in a chronic illness care management program. *Telemedicine and e-Health*. 2007; 13:51–6. [PubMed: 17309355]

11. Millard RW, Carver JR. Cross-sectional comparison of live and interactive voice recognition administration of the SF-12 health status survey. *The American journal of managed care*. 1999; 5:153–9. [PubMed: 10346511]
12. Piette JD, Datwani H, Gaudioso S, et al. Hypertension management using mobile technology and home blood pressure monitoring: results of a randomized trial in two low/middle income countries. *Telemedicine and e-Health*. 2012; 18:613–20. [PubMed: 23061642]
13. Graham J, Tomcavage J, Salek D, Sciandra J, Davis DE, Stewart WF. Postdischarge monitoring using interactive voice response system reduces 30-day readmission rates in a case-managed Medicare population. *Medical care*. 2012; 50:50–7s. [PubMed: 21822152]
14. Estabrooks PA, Shoup JA, Gattshall M, Dandamudi P, Shetterly S, Xu S. Automated telephone counseling for parents of overweight children: a randomized controlled trial. *Am J Prev Med*. 2009; 36:35–42. [PubMed: 19095163]
15. Mahoney DF, Tarlow BJ, Jones RN. Effects of an automated telephone support system on caregiver burden and anxiety: findings from the REACH for TLC intervention study. *Gerontologist*. 2003; 43:556–67. [PubMed: 12937334]
16. Lichtenstein MJ, Steele MA, Hoehn TP, Bulpitt CJ, Coles EC. Visit frequency for essential hypertension: observed associations. *The Journal of Family Practice*. 1989; 28:667–72. [PubMed: 2723595]
17. Schectman G, Barnas G, Laud P, Cantwell L, Horton M, Zarling EJ. Prolonging the return visit interval in primary care. *The American Journal of Medicine*. 2005; 118:393–9. [PubMed: 15808137]
18. DeSalvo KB, Block JP, Muntner P, Merrill W. Predictors of variation in office visit interval assignment. *International Journal for Quality in Health Care*. 2003; 15:399–405. [PubMed: 14527983]
19. DeSalvo KB, Bowdish BE, Alper AS, Grossman DM, Merrill WW. Physician practice variation in assignment of return interval. *Archives of Internal Medicine*. 2000; 160:205–8. [PubMed: 10647759]
20. Piette JD, Kerr EA. The impact of comorbid chronic conditions on diabetes care. *Diabetes Care*. 2006; 29:725–31. [PubMed: 16505540]
21. Piette JD, Sussman JB, Pfeiffer PN, Silveira MJ, Singh SB, Lavieri M. Maximizing the value of mobile health monitoring by avoiding redundant patient reports: prediction of depression-related symptoms and adherence problems in automated health assessment services. *J Med Internet Res*. 2013; 13:e118. [PubMed: 23832021]
22. Polonsky WH, Fisher L, Earles J, et al. Assessing psychosocial distress in diabetes: development of the diabetes distress scale. *Diabetes Care*. 2005; 28:626–31. [PubMed: 15735199]
23. Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical care*. 1996; 34:220–33. [PubMed: 8628042]
24. Haas LMM, Beck J, Cox CE, et al. National standards for diabetes self-management education and support. *Diabetes Care*. 2013; 36:S100–S8. [PubMed: 23264420]
25. Huber, PJ. The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; Berkeley, CA: University of California Press; 1967. p. 221-33.
26. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*. 1980; 48:817–30.
27. Rogers WH. Regression standard errors in clustered samples. *Stata Technical Bulletin*. 1993; 13:19–23. Also see: http://www.stata.com/support/faqs/statistics/stb13_rogers.pdf.
28. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics*. 2000; 56:645–6. [PubMed: 10877330]
29. Froot KA. Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data. *Journal of Financial and Quantitative Analysis*. 1989; 24:333–55.
30. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*. 1997; 30:1145–59. 10.1016/S0031-3203(96)00142-2

31. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. *The Journal of Clinical Hypertension*. 2008; 10:348–54. [PubMed: 18453793]
32. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology*. 2010; 21:128–38. [PubMed: 20010215]
33. Michalski, RS.; Carbonell, JG.; Mitchell, TM. *Machine Learning: An Artificial Intelligence Approach*. Tioga Publishing Company; 1983.

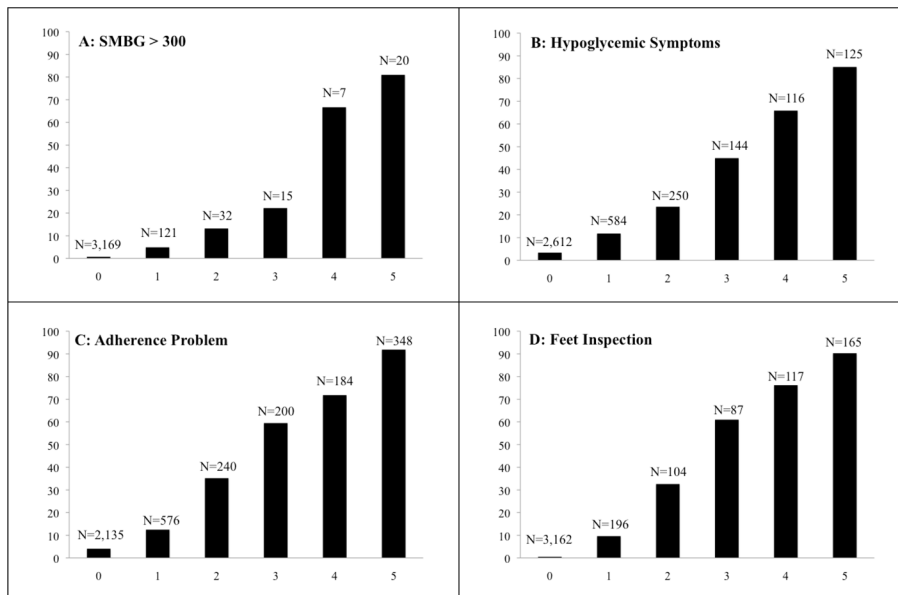


Figure 1. Probability of IVR-Reported Problems by the Frequency of Problem Reports during the Preceding Five Assessments

Notes: N's are the number of assessments with each frequency of prior positive reports. For example, as shown in Panel A, patients completed an assessment 3,169 times in which they were asked whether they had an SMBG > 300 in the prior week, even though they had never reported levels that high in the past five calls. Among those 3,169 assessments, patients reported SMBG > 300 only 0.6% of the time.

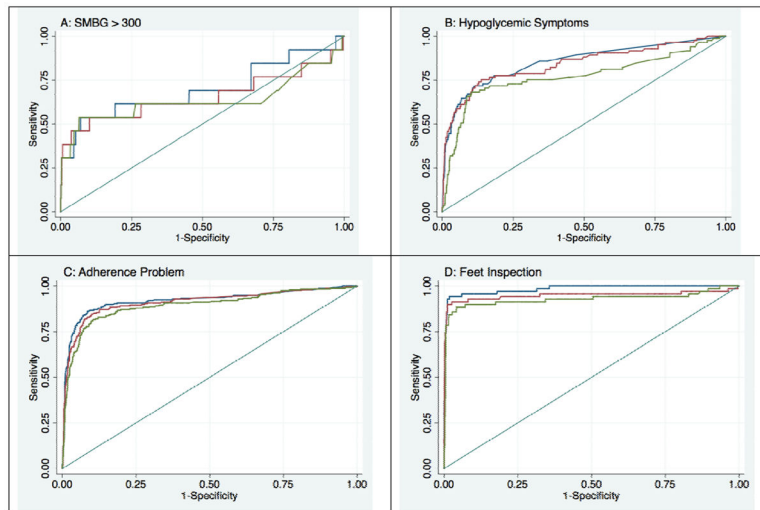


Figure 2. Receiver Operator Characteristic Curves for Models Predicting Patients' IVR Responses Based on Prior Reports and Baseline Data

Notes: Blue lines are based on models predicting outcomes using data collected from patients at baseline, as well as during the most recent five IVR assessments captured assuming weekly assessment attempts. Red lines are based on models assuming bi-weekly assessment attempts, and green lines are based on models assuming assessments collected no more than every three weeks.

Table 1

Unadjusted Associations Between Selected Baseline Risk Factors and Patients' IVR Reports of Health and Self-Care Problems

N = 298 unique patients	Baseline Measures				
	Ever Reported	PAID ^a Mean (SD)	MCS ^b Mean (SD)	PCS ^c Mean (SD)	Morisky ^d (% classified nonadherent)
SMBG^e > 300 mg/dL					
Yes (13.0%) ^f	16.9 (2.2)*	49.7 (12.2)	26.7 (33.4)	44.7*	
No (87.0%)	12.4 (12.6)	50.0 (11.7)	33.4 (12.2)	27.9	
SMBG < 90 mg/dL					
Yes (56.3%)	14.2 (12.9)	50.0 (11.8)	30.7 (12.7)	33.7	
No (43.7%)	11.4 (12.5)	49.9 (11.7)	35.0 (11.4)	24.4	
Hyperglycemic Symptoms					
Yes (62.7%)	15.8 (13.8)*	48.4 (11.7)*	30.5 (11.9)*	36.7*	
No (37.3%)	8.4 (9.1)	52.6 (11.1)	35.8 (12.2)	19.4	
Hypoglycemic Symptoms					
Yes (62.8%)	15.7 (13.6)*	48.6 (11.0)*	31.0 (11.9)*	33.7*	
No (37.2%)	8.4 (9.7)	52.5 (11.0)	35.5 (12.4)	23.4	
Fair/Poor Health					
Yes (47.5%)	16.4 (14.1)*	45.8 (11.9)*	25.8 (11.0)*	36.6*	
No (52.5%)	9.9 (10.7)	53.6 (10.1)	38.5 (10.2)	24.2	
1+ Days in Bed					
Yes (44.5%)	16.5 (15.2)*	46.9 (12.3)*	28.4 (11.3)*	38.5*	
No (55.5%)	10.2 (9.7)	52.6 (10.4)	36.1 (12.0)	23.0	
Adherence Problem					
Yes (61.8%)	14.8 (13.9)*	49.3 (12.0)	33.0 (11.8)	41.9*	
No (38.2%)	10.1 (10.2)	51.2 (11.0)	32.1 (13.0)	10.7	
Not Checking Feet					
Yes (29.6%)	15.2 (13.7)*	47.5 (11.5)*	32.6 (11.1)	38.1*	
No (70.4%)	12.1 (12.3)	51.1 (11.6)	32.6 (12.8)	26.6	
No SMBG					
Yes (34.9%)	15.9 (14.2)*	47.3 (12.3)*	32.1 (12.0)	39.8*	
No (65.1%)	11.4 (11.7)	51.4 (11.1)	32.9 (12.4)	24.9	

Notes: Asterisks indicate statistically significant differences ($P < .05$) in that baseline measure between groups who ever versus never reported that problem during their IVR assessments.

^aPAID: Problem Areas in Diabetes scale, a measure of diabetes-related distress. Higher numbers indicate greater distress.

^bMCS: The Mental Composite Summary Score from the SF-12, a measure of mental health-related functioning. Higher numbers indicate better functioning.

^cPCS: The Physical Composite Summary Score from the SF-12, a measure of physical health-related functioning. Higher numbers indicate greater functioning.

^d Morisky: The Morisky medication adherence scale. Percentages represent the number of respondents with a score of 2+ indicating significant medication adherence problems.

^e SMBG: self-monitored blood glucose.

^f Percentages represent the number of patients who ever versus never reported that problem during one of their assessments. IVR: Interactive Voice Response.

Table 2

Area Under the ROC Curve for Logistic Models Predicting Health and Self-Care Problems, and the Potential Impact on Detection and Respondent Burden of Dropping Assessments for Which the Respondent has 5% Probability of A Problem Report According to Previously-Reported Information

	Avoiding Items When P = 5%				
	AUC	95% CI	Assessments Avoided (%)	% of Problems Captured	% Neg Calls Avoided
<u>GLUCOSE LEVELS</u>					
SMBG > 300 mg/dL					
Weekly reporting	69.7	50.2, 89.2	96.7	46.3	97.3
Bi-weekly reporting	65.6	43.7, 87.5	97.1	38.9	97.5
Tri-Weekly reporting	63.6	40.7, 86.5	95.5	46.2	96.3
SMBG < 90 mg/dL					
Weekly reporting	70.1	63.4, 78.0	50.9	79.8	54.3
Bi-weekly reporting	61.0	52.9, 69.0	63.6	74.5	68.7
Tri-Weekly reporting	70.0	63.2, 76.4	68.4	55.3	70.1
<u>PERCEIVED HEALTH</u>					
Hyperglycemic Symptoms					
Weekly reporting	85.4	80.5, 90.4	69.2	87.5	77.0
Bi-weekly reporting	84.3	78.9, 89.7	71.7	84.5	78.2
Tri-Weekly reporting	77.2	70.6, 84.0	73.6	78.9	81.2
Hypoglycemic Symptoms					
Weekly reporting	80.7	76.0, 85.3	52.1	85.3	58.1
Bi-weekly reporting	77.9	73.2, 82.6	54.6	81.3	58.7
Tri-Weekly reporting	77.2	77.5, 82.0	55.5	78.6	59.9
Fair/Poor Health					
Weekly reporting	96.2	94.6, 97.7	72.9	90.3	87.7
Bi-weekly reporting	93.5	91.2, 95.8	68.8	91.0	82.7
Tri-Weekly reporting	93.3	91.1, 95.4	63.8	92.1	76.4
<u>HEALTH BEHAVIORS</u>					
1+ Days in Bed					
Weekly reporting	86.2	81.0, 91.5	79.3	75.2	83.7
Bi-weekly reporting	83.8	77.6, 90.0	82.4	70.6	86.4

	Avoiding Items When P = 5%				
	AUC	95% CI	Assessments Avoided (%)	% of Problems Captured	% Neg Calls Avoided
Tri-Weekly reporting	81.0	74.4, 87.7	83.6	67.6	87.5
Adherence Problem					
Weekly reporting	92.1	89.6, 94.6	64.9	90.1	78.2
Bi-weekly reporting	91.2	88.7, 93.8	58.5	88.5	70.0
Tri-Weekly reporting	89.4	86.7, 92.2	54.2	90.4	65.2
Not Checking Feet					
Weekly reporting	98.4	97.0, 99.8	93.7	90.5	94.9
Bi-weekly reporting	94.9	90.4, 99.4	90.2	88.1	95.5
Tri-Weekly reporting	92.9	87.8, 97.9	90.7	87.0	95.5
Not SMBG					
Weekly reporting	82.2	75.7, 88.6	89.3	73.0	93.4
Bi-weekly reporting	77.3	69.4, 85.3	90.3	75.3	94.3
Tri-Weekly reporting	83.4	76.6, 90.4	90.8	68.1	93.1

Notes: See Methods for a definition of each problem. AUC=area under the curve for the receiver operator characteristic curves measuring the predictive significance of each model.