Evaluation of Multimedia Medication Reconciliation Software: A Randomized Controlled, Single-Blind Trial to Measure Diagnostic Accuracy for Discrepancy Detection

Blake J. Lesselroth1,2  Kathleen Adams1  Victoria L. Church1  Stephanie Tallett1  Yelizaveta Russ3
Jack Wiedrick4  Christopher Forsberg5  David A. Dorr2

1NorthWest Innovation Center, Veterans’ Affairs Portland Healthcare System, Portland, Oregon, United States
2Department of Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, Portland, Oregon, United States
3Division of Primary Care, Veterans’ Affairs Portland Healthcare System, Portland, Oregon, United States
4Oregon Clinical and Translational Research Institute, Oregon Health and Science University, Portland, Oregon, United States
5Center of Innovation, Veterans’ Affairs Portland Healthcare System, Portland, Oregon, United States

Address for correspondence  Blake J. Lesselroth, MD, MBI, NorthWest Innovation Center, Veterans’ Affairs Portland Healthcare System, Portland, Oregon, United States (e-mail: blake.lesselroth@va.gov).


Abstract

Background  The Veterans Affairs Portland Healthcare System developed a medication history collection software that displays prescription names and medication images.

Objective  This article measures the frequency of medication discrepancy reporting using the medication history collection software and compares with the frequency of reporting using a paper-based process. This article also determines the accuracy of each method by comparing both strategies to a best possible medication history.

Study Design  Randomized, controlled, single-blind trial.


Participants  Of 212 patients with primary care appointments, 209 patients fulfilled the study requirements.

Intervention  Patients randomized to a software-directed medication history or a paper-based medication history. Randomization and allocation to treatment groups were performed using a computer-based random number generator. Assignments were placed in a sealed envelope and opened after participant consent. The research coordinator did not know or have access to the treatment assignment until the time of presentation.

Main Outcome Measures  The primary analysis compared the discrepancy detection rates between groups with respect to the health record and a best possible medication history.

Keywords
► medication reconciliation
► clinical information systems
► quality improvement
► medication errors
► medication adherence
► clinical informatics
► consumer health informatics
► medication review

received  November 18, 2017
accepted after revision  March 20, 2018

ISSN 1869-0327.
Results  Of 3,500 medications reviewed, we detected 1,435 discrepancies. Forty-six percent of those discrepancies were potentially high risk for causing an adverse drug event. There was no difference in detection rates between treatment arms. Software sensitivity was 83% and specificity was 91%; paper sensitivity was 81% and specificity was 94%. No participants were lost to follow-up.

Conclusion  The medication history collection software is an efficient and scalable method for gathering a medication history and detecting high-risk discrepancies. Although it included medication images, the technology did not improve accuracy over a paper list when compared with a best possible medication history.

Trial Registration  ClinicalTrials.gov Identifier: NCT02135731.

Background and Significance

Any incongruity in prescription information between two or more different information sources may result in a medication discrepancy.1 The risk of an individual discrepancy is relatively small.2 However, in aggregate, medication discrepancies during care transitions and clinic visits are a common cause of preventable adverse drug events (ADEs) and patient harm.3–11 Unintentional discrepancies in allergy and medication information contribute to nearly half a million hospitalizations and cost the United States’ (U.S.) health care system upwards of $1 billion annually.12–15

According to the Institute of Healthcare Improvement, medication reconciliation (MR) is the “process of identifying the most accurate list of all medications a patient is taking... and using this list to provide correct medications for patients anywhere within the healthcare system.”16 Studies show that standardized MR programs reliably identify discrepancies and reduce medical error.5,17–21 While quality advocacy organizations and regulatory agencies recommend implementing scalable systems-based solutions, health care organizations typically struggle with an array of individual- and system-based implementation barriers.3,16,22–30

Clinicians compiling medication histories under time pressure often adopt manual workarounds to overcome poor electronic health record (EHR) usability and lack of data interoperability.29–33 They cite challenges with tight schedules, competing care priorities, and limited patient reliability.30 This is unsurprising given that a medication history can require 20 minutes and a high-quality reconciliation can take up to 80 minutes.33 Therefore, many clinicians dismiss MR as a set of “administrative accounting tasks” and skip reviewing medications with patients altogether.34 By contrast, quality MR programs foster patient-centered discussion, pull data from multiple sources, and promote interdisciplinary communication.23,35 Clearly, there is a need for user-centered tools that assist with history collection and that improve the efficiency, consistency, and accuracy of MR.17,26,27,36–41

Problem Statement

The Veterans Affairs (VA) initiated an MR campaign to establish standards, promote change, and measure implementation effectiveness.42 The campaign required clinicians to collect a medication history at every encounter and compare with facility documentation. Many clinics failed to meet these guidelines. Clinicians often found it difficult to collect or update medication information, opting instead to copy and paste existing EHR medication lists into their notes. In a previous survey of local primary care providers, respondents identified three barriers affecting MR: poor EHR design, inflexible workflows, and insufficient time.29 We therefore developed a medication history collection software to engage patients, improve data integration, and streamline the process.38

We modeled our Automated Patient History Intake Device (APHID) software after interactive self-service functionality commonly featured on retail Web sites and commercial air travel check-in kiosks.38,43 Patients used an APHID-enabled kiosk located in the clinic waiting room to check-in for their appointments and review the names and pictures of their prescriptions. Clinicians then imported the results into the EHR for reference during the medication interview.

Although our initial piloting efforts demonstrated the viability of this approach, we had not yet measured its accuracy compared with traditional methods of history collection such as interviews and paper questionnaires.29,38,43 Few studies have demonstrated the effectiveness of self-service software or multimedia tools for gathering reliable medication information.44–49 In small nonrandomized pilot studies, Hornick et al and Kimmel et al found that medication images improved patient recall for certain classes of medications.50,51 Before these strategies can be recommended, high-quality studies are needed to inform design and establish the accuracy of a self-reported medication history.

Objective

This study compared the diagnostic accuracy of the APHID collection process to a paper-based collection process. To determine which of these methods was more effective, we compared both strategies to the EHR list and a reference standard. For our reference standard, we assembled a best possible medication history (BPMH). (Table 1): a systematic clinician-conducted history using several information
sources. We hypothesized that participants using APHID would report more medication discrepancies than those using a paper list. We also hypothesized that the participant-reviewed APHID list would be more accurate than the participant-reviewed paper list when compared with the BPMH.

**Materials and Methods**

**Theory**

We based our study design on two well-established theories: (1) the Systems Engineering to Improve Patient Safety (SEIPS) framework; and (2) the Pictorial Superiority Effect (PSE). Carayon et al’s SEIPS framework argues that system constructs (i.e., people, technologies, workflow, and culture) dictate health system processes and clinical outcomes. In our adaptation of the framework, MR is a cyclical macroprocess composed of linked subprocesses (e.g., history collection, data adjudication, discrepancy resolution). The quality of MR can only be as effective as the initial history. We therefore designed our intervention to address system constructs supporting history collection.

The PSE contends that humans encode, store, and retrieve images from memory more easily than text or auditory information. Studies indicate patient education materials using pictures and pictograms can influence health literacy and improve comprehension. Since it is rare for patients to bring their medications to clinic, we hypothesized that providing patients with medication images would reduce the number of errors caused by “look-alike” and “sound-alike” medications.

**Description of the Technology**

Our technology consists of three main components: (1) a self-service kiosk; (2) the APHID medication history collection software; and (3) an interface to the facility EHR. For a complete description of the technology, please refer to our technology development and deployment manuscript.

The APHID software has access to all prescription information in the EHR and can use stored metadata to match each medication with a digital image. Our organization manages the medication supply chain for most prescriptions; medications written by VA prescribers are typically dispensed from VA facilities or regional mail-out distribution centers. The EHR stores medication dispense dates, prescription refill histories, medication inventory numbers, and U.S. National Drug Code (NDC) numbers. The NDC numbers are unique 10-digit, 3-segment numbers assigned by the U.S. Food and Drug Administration to all drugs distributed in the United States. APHID uses a combination of NDC numbers and dispensing data to match an image with each prescription. To assemble a medication list for patient review, APHID retrieves prescription data from all VA facilities and pairs each medication name and instructions for use with a single digital photograph of the prescription. If a patient has received several prescriptions for the same drug, APHID uses the image associated with the last dispense date.

It is customary for VA clinicians to document prescriptions from non-VA practitioners and nonprescription medications reported by the patient. Approximately 10 to 30% of prescriptions are procured outside the VA. APHID will display this information when available in the EHR. However, the VA does not exchange health data with community drug dispensaries. Thus, APHID cannot match images to prescriptions procured outside the VA.

**Subjects and Settings**

We recruited U.S. Veteran patients from three primary care clinics associated with the VA Portland Healthcare System (VAPORHCS). VAPORHCS is a 300-bed tertiary care hospital with eight associated ambulatory care centers located in...
northern Oregon and southern Washington states. It is part of the Veterans Health Affairs, a nationwide care network of over 150 hospitals. Most Veterans are male (94%), older than the average civilian patient, have a greater number of medical comorbidities, and use more medications.

**Study Design**

We conducted a prospective parallel-randomized, controlled, single-blind study of the APHID software. Using the BPMH as a reference standard, we compared the discrepancy counts reported using a paper list with those reported by APHID (Fig. 3). We designed our study as per the CONSORT guidelines for reporting clinical trials and STARD guidelines for reporting studies of diagnostic tests. For a detailed description of the methods and instruments, please refer to our previously published protocol.

From June 2009 through December 2011, a research coordinator contacted all patients scheduled in participating clinics and screened them for inclusion in the study. Patients were eligible for inclusion if they were over the age of 18 years, taking three or more medications, and had completed at least one appointment in the past. Exclusion criteria included: the inability to read or speak English; the presence of cognitive impairment or mental illness; visual impairment; and physical impairment that might preclude use of a mouse or keyboard. The research coordinator asked participating patients to bring in all their medications to the appointment.

On the day of the study appointment, the research coordinator confirmed the study participant’s eligibility and then consented, randomized, and assigned the participant to one of the two treatments (Fig. 3, step 2). A member of the research team determined the treatment assignment using a computer-based random number generator and placed the output in a sealed envelope. The research coordinator did not know or have access to the treatment assignment until the time of presentation when the envelope was opened.

The research coordinator asked participants to review, correct, and amend their medication lists (Fig. 3, step 3). Participants in the intervention arm used APHID at a standard computer workstation, whereas participants in the control arm used paper on a clipboard. Participants using APHID completed an onscreen questionnaire asking them about (1) their prescriptions dispensed by the study facility; (2) other VA prescriptions; (3) nonprescription medications; and (4) 6 months of expired prescriptions. Each prescription was displayed, one at a time, with an image where available. Participants were asked to indicate adherence using one of
Participants then added any new or missing medications. Participants in the control arm reviewed a paper list that included (1) prescriptions dispensed by the study facility; (2) other VA prescriptions; (3) nonprescription medications; and (4) 6 months of expired prescriptions. This control included more information than what was available in the facility EHR and was more stringent than the “usual care” practice of importing an automated list into clinic notes. We selected this control to isolate the effects of medication pictures and self-service software. It was crucial that the paper and APHID lists contained the same medications for review and the same number of opportunities to report a discrepancy.

The research coordinator asked the control group participants to place a “yes” on the sheet next to medications they were taking and “no,” “differently,” or “unsure” next to medications that they were not taking, taking differently, or had a question about. We coded any response other than “yes” as a discrepancy for active medications. We coded any response other than “no” as a discrepancy for expired medications. Participants listed additional medications at the bottom of the form.

Clinician researchers (an internal medicine physician, a clinical pharmacist, and an advanced practice nurse) were trained to collect a medication history using an interview script. The researchers then independently interviewed standardized patients to eliminate variation in interview technique. The performance characteristics were identical between researchers after two practice cycles.

In the study, a blinded clinician researcher met with each participant and completed a medication history using the interview script (Fig. 3; step 5). This is the most common method used in MR studies to establish a reference standard. The researchers were also instructed to review the EHR and inspect the medication containers. Then the researcher recorded a discrepancy on a spreadsheet if the participant was not taking a medication associated with a current prescription, taking a medication associated with an expired prescription, taking a medication differently than instructed, or taking a new medication.

Our reference standard (i.e., the BPMH) included an additional step and an additional information source (Fig. 3, step 6). The BPMH included the clinician-gathered medication history, a review of EHR pharmacy records, a prescription vial or direct pill inspection, and a double-check using the participant-reviewed blinded list. The research coordinator furnished the clinician researcher with the blinded list. The clinician then completed the “double-check” with the patient, adjudicating mismatches identified between the clinician history and the blinded list. The researcher then furnished the BPMH to the primary care provider.

Our institutional review board (IRB) required the adjudication step (Fig. 3, step 6) for safety purposes; it was crucial to disclose all information sources to the primary care team. Furthermore, we used both histories—the clinician-gathered medication history and the BPMH—during statistical analysis to screen for differences or trends between the first and second clinician review.
Researchers classified all discrepancies by descriptive type using a typology adapted from Pippins et al. Researchers also classified discrepancies by root cause using an instrument adapted from Orrico and Smith et al. A blinded clinician panel then assigned risk scores using a classification scheme adapted from Pippins et al and Wong et al (see Fig. A1 for risk assessment protocol). We used sample sets of discrepancies collected during piloting to train clinician raters on the instrument and calibrate responses.

**Analysis**
To determine discrepancy rates, we compared the participant-reviewed lists against the EHR list and then calculated the proportion of medications with a discrepancy (Fig. 3, step 1–step 2). Similarly, we compared the participant-reviewed lists against the BPMH and calculated the proportion of medications with a discrepancy (Fig. 3, step 2–step 6). The primary outcome measure was the difference in discrepancy rates between each arm with respect to the EHR list. Our secondary outcome measures were the difference in discrepancy rates and high-risk discrepancy rates between each arm compared with the BPMH. Assuming a discrepancy base rate of 25% per participant list, our power calculations indicated that we needed a sample size of 210 participants to detect a difference of 15% in detection rates between treatment arms. We used a crossed random-effects model (three phases of evaluation crossed with participant lists) to account for discrepancy-clustering effects at the participant level. Discrepancy rates were calculated for two sets of comparisons: (1) treatment arm versus EHR; and (2) treatment arm versus BPMH. We completed a poststratified analysis to compare differences in discrepancy counts between treatment arms correcting for the number of medications between treatment arms (i.e., opportunities to detect an error). We also computed intercorrelation coefficients (ICCs) of absolute agreement by treatment arm for individual discrepancies at the level of the medication list item across evaluation phases (clinician-gathered medication history and BPMH) to assess whether one of the treatment arms showed significantly more agreement than the other with respect to a given medication list item. No gross difference in agreement was seen between the treatment arms.

We assessed the accuracy of treatment methods relative to the BPMH by comparing the discrepancy status (either “yes” or “no”) according to the BPMH within each treatment arm of each medication list item that was identified by both the BPMH and the treatment. We then tallied the counts of (“yes,” “yes”), (“yes,” “no”), (“no,” “yes”), and (“no,” “no”), where the first response in each pair was from the BPMH and the second response was from a treatment method. This yielded two sets of four counts each; arranged in two 2 × 2 cross-tabulation tables and individually analyzed on standard diagnostic agreement metrics. We assumed that the BPMH represented the true status and the treatment method represented the test status. The metrics we report include sensitivity (fraction of test positives among true positives), specificity (fraction of test negatives among true negatives), positive predictive value (fraction of true positives among test positives), negative predictive value (fraction of true negatives among test negatives), positive likelihood ratio (sensitivity divided by false positive fraction among true negatives), and negative likelihood ratio (false negative fraction among true positives, divided by specificity), where the likelihood ratios are weighted by the prevalence odds. Confidence intervals were calculated for each metric on each cross-tabulation using exact binomial distributions.

**Results**
We assessed 614 patients for eligibility; 220 participants were enrolled and randomized for the study (Fig. 4). The study pilot involved eight participants and three participants were withdrawn, leaving 209 participants included in the final analysis. No participants were lost to follow-up. There were no incidents of accidental unblinding treatment status for the clinician interviewers.

Participant sample characteristics and descriptive statistics are reported in Table 2. Most enrolled participants were male, with a mean age of 66.5 years, and an estimated 5.6 chronic medical illnesses. Each participant had an average of 11.5 active prescriptions. When accounting for recently expired and nonprescription medications, we reviewed an average of 16.7 medications per participant.

Descriptive statistics for all discrepancies detected are outlined in Table 3. Using all information sources, our team identified 3,500 medications and 1,435 discrepancies. We detected 530 discrepancies using the paper list and 594 discrepancies using the APHID list. Of the 1,435 discrepancies identified by the BPMH, 657 (46%) were high or very high risk. We traced 47% of the discrepancies to a system-based root cause (i.e., clinical documentation errors) and the remainder to a participant-based root cause (e.g., nonadherence). VA medications accounted for 49% of the discrepancies by prescription status, expired VA medications accounted for 15%, and medications not in the EHR accounted for 37%. We did not identify any differences between treatment arms.

There were no statistically significant differences in the rate of discrepancies reported (i.e., primary outcome) for each treatment arm when compared with the study facility EHR (Table 4). An average of 35% (0.35 ± 0.20) of medications on each list included a discrepancy (p = 0.89); 15% of all medications in the control arm and 17% of all medications in the intervention arm included a high-risk discrepancy (this corresponds to 43% of all discrepancies in the control group and 49% of all discrepancies in the intervention group).

When comparing treatment arms to the BPMH (i.e., secondary outcome), we did not identify any differences in discrepancy rates (Table 4). The BPMH identified an additional 253 medications not recorded in the EHR or by either treatment arm. The BPMH included ~13% (0.13 ± 0.13) more discrepancies than either treatment; 31 to 38% of those discrepancies (4–5% of all medications reviewed) were high risk. As a test for bias (i.e., an interaction between test and reference standard), we compared detection rates for the clinician-gathered medication history with the BPMH.
calculated a concordance rate of 98.6% (kappa = 0.97), arguing against any interaction between the test and the reference standard. When comparing the EHR lists to the BPMH, 207 of the 209 charts (99%) included one or more discrepancies.

Table 5 lists the cross-tabulations for each treatment arm compared with the BPMH. Cross-tabulations only included medications present in the treatment list and the BPMH. New medications only identified on the BPMH were undefined for the treatment arms and could not be counted.
as true positives or false positives. The paper-based process had a sensitivity of 81% and APHID had a sensitivity of 83%. The paper-based process had a specificity of 94% and APHID had a specificity of 91%. Overall, we did not detect a statistically significant difference in the accuracy of either method for detecting discrepancies or high-risk discrepancies when compared with the BPMH.

**Discussion**

**Principal Findings**

Discrepancy detection rates were similar between the paper-based and APHID processes for all dimensions of analysis. The addition of medication images did not affect history accuracy. Discrepancy rates, sensitivity statistics, and negative predictive values were virtually identical when compared with the EHR, even after adjusting for participant characteristics, discrepancy risk category, or root cause.

APHID offers an efficient and patient-centered method for collecting a medication history and documenting discrepancies. It compiled a list of medications from across the VA enterprise and helped identify nearly 90% of all discrepancies in our sample. APHID’s detection rates compared favorably with best-practice interviews (i.e., ≥1 discrepancy in 99% of all EHR lists). In all, over a third of the prescriptions included a discrepancy, nearly half of which were high-risk.

---

**Table 1** Glossary of terms used in this article

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adverse drug event (ADE)</td>
<td>Allergic reactions, adverse effects, or unintentional overdoses⁶</td>
</tr>
<tr>
<td>Automated Patient History Intake Device (APHID)</td>
<td>Veterans’ Affairs (VA) software that supports collection and documentation of a medication history³⁸</td>
</tr>
<tr>
<td>APHID list</td>
<td>A medication list generated by the APHID software that includes current prescriptions written at the study facility, 6 months of expired prescriptions written at the study facility, non-VA prescriptions and nonprescription medications documented at the study facility, and prescriptions written at other VA facilities³⁸</td>
</tr>
<tr>
<td>Best possible medication history (BPMH)</td>
<td>Clinician-gathered medication history adjudicated with the participant-reviewed blinded list; reference standard³⁴</td>
</tr>
<tr>
<td>Clinician-gathered medication history</td>
<td>A medication history collected by the clinician researcher while referencing the EHR and inspecting medications brought in by the patient</td>
</tr>
<tr>
<td>Electronic health record (EHR) list</td>
<td>List of prescription information in the VA electronic health record</td>
</tr>
<tr>
<td>Expired medications</td>
<td>Electronic prescriptions that have passed a predefined expiration date set by the EHR</td>
</tr>
<tr>
<td>Kiosk</td>
<td>Self-service hardware equipped with a touch-screen and installed with APHID software³⁸,⁴⁵</td>
</tr>
<tr>
<td>Medication discrepancy</td>
<td>Any unintended incongruity between prescription lists from two sources⁵</td>
</tr>
<tr>
<td>Medication errors</td>
<td>Failure in prescribing or treatment process that can result in patient harm¹²⁵</td>
</tr>
<tr>
<td>Medication history</td>
<td>A process of identifying the list of all medications a patient is taking by interviewing the patient/family and reviewing available documentation⁹</td>
</tr>
<tr>
<td>Medication information</td>
<td>Information about prescribing/administration/consumption of medications</td>
</tr>
<tr>
<td>Medication reconciliation</td>
<td>A process of identifying the most accurate list of all medications a patient is taking and using this list to provide correct information across the continuum of care¹⁶</td>
</tr>
<tr>
<td>Non-VA medications</td>
<td>Medications procured outside the VA by the patient and documented in the EHR by a clinician</td>
</tr>
<tr>
<td>Paper list</td>
<td>A medication list that includes current prescriptions written at the study facility, 6 months of expired prescriptions written at the study facility, non-VA prescriptions and nonprescription medications documented at the study facility, and prescriptions written at other VA facilities</td>
</tr>
<tr>
<td>Participant adherence</td>
<td>Extent to which medications are taken as directed</td>
</tr>
<tr>
<td>Participant-reviewed APHID list</td>
<td>APHID list of medications with participant adherence documented</td>
</tr>
<tr>
<td>Participant-reviewed blinded list</td>
<td>Form created by the research coordinator listing medications and participant-furnished adherence history</td>
</tr>
<tr>
<td>Participant-reviewed paper list</td>
<td>Paper list of medications with participant adherence documented¹</td>
</tr>
<tr>
<td>Root cause (participant-based)</td>
<td>Discrepancy caused by a factor under the patient’s control¹</td>
</tr>
<tr>
<td>Root cause (system-based)</td>
<td>Discrepancy caused by a clinician or health system factor</td>
</tr>
<tr>
<td>VA medications (local)</td>
<td>Electronic prescriptions written and documented in the study facility EHR</td>
</tr>
<tr>
<td>VA medications (remote)</td>
<td>Electronic prescriptions written at VA facilities other than the study facility</td>
</tr>
</tbody>
</table>
Approximately 15% of prescriptions with discrepancies were expired, indicating that asking about expired prescriptions identified errors of omission that might otherwise have gone undetected. This illustrates the value of using a patient-centered and standardized process to collect a medication history.50,67,68,73,83 Although the self-service history was more complete that the EHR list, the BPMH identified an additional high-risk discrepancy in nearly 5% of medications reviewed. This demonstrates the importance of clinician engagement. Self-service history collection techniques work in conjunction with—not in lieu of—a clinician-mediated medication history.41,46,68,73,84 Several possibilities may explain why our study detected more discrepancies than most other MR studies published in the last 5 years (34–88%);25,27,37,46,47,72,85,86 First, Veterans tend to be older, more medically complex, and take more medications than the general population. Each of these factors has been shown to correlate with the incidence of medication errors.74,76,87–90 Second, sequential history collection steps may improve patient recall. Studies suggest single histories may be less reliable, prone to drift, and overestimate compliance by up to 20%.76,77,87,91–94 Third, studies show that data gathered from self-administered questionnaires are less affected by social desirability bias than interviews95; patients may be more likely to report nonadherence when independently correcting a medication list. Compared with pill counts, interviews have a reported sensitivity of 55 to 80% and a specificity of 70 to 87%.77,93 Therefore, our self-reporting techniques may have been more effective than a typical interview. Fourth, our BPMH used a combination of supply chain metadata, local prescription lists, and patient furnished data.25,27,96 Combining data sources increases the ability to document “ground truth”—an important consideration for reconciliation systems.34,74,76,83,90 This study underscores the importance of sociotechnical fit when implementing a reconciliation program. Clinician engagement, standardized processes, and patient-centered strategies may have greater influence upon the accuracy, quality, and overall success of MR than any specific technology.85 Most MR methods—including paper—may be equifinal if informaticians optimize workflow and implementation climate (i.e., culture, leadership, and education). Nonetheless, facilities contending with staff shortages, high patient volumes, or large geographic areas may consider using patient-centered technologies such as kiosks, secure messaging, or online portals to improve efficiency and scalability without sacrificing effectiveness or accuracy.27,29,37

### Study Strengths and Limitations

To the best of our knowledge, there are very few published studies comparing MR strategies and only four other randomized controlled trials (RCTs) conducted in the ambulatory setting, none of which adhere to the CONSORT/STARD guidelines.19,21,47,74,85,86,96,98–100 Systematic reviews of ambulatory MR programs suggest more research is needed to prove the clinical impact of MR.21,27,74,96,98,101 Our study makes an important contribution to the literature by providing accuracy statistics for multiple collection strategies.

This study has several limitations that may affect the validity of our findings. First, the high number of reported medications, older age, and unbalanced gender distribution of our sample likely affected the results. Our findings may not be generalizable to other settings and the prevalence of medication discrepancies may be lower in other populations.43,102 Second, limitations in health information exchange and EHR interoperability prevented us from matching images to nonprescription and non-VA medications (~27% of medications). This may have caused a type II error. Finally, to provide equal opportunity for discrepancy detection in each treatment arm, we compiled medication

### Table 2 Descriptive statistics of participant sample

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control</th>
<th>Intervention</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants, count (%)</td>
<td>102 (48.8)</td>
<td>107 (51.2)</td>
<td>209 (100.0)</td>
</tr>
<tr>
<td>Male gender, count (%)</td>
<td>99 (97.1)</td>
<td>101 (94.4)</td>
<td>200 (95.7)</td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>67.6 (12.3)</td>
<td>65.5 (12.2)</td>
<td>66.5 (12.3)</td>
</tr>
<tr>
<td>Chronic conditions, mean (SD)</td>
<td>5.5 (2.0)</td>
<td>5.7 (2.1)</td>
<td>5.6 (2.1)</td>
</tr>
<tr>
<td>Medications, mean (SD)</td>
<td>16.8 (7.1)</td>
<td>16.7 (8.5)</td>
<td>16.7 (7.8)</td>
</tr>
<tr>
<td>Current prescriptions in EHR</td>
<td>11.5 (5.7)</td>
<td>11.5 (6.0)</td>
<td>11.5 (5.9)</td>
</tr>
<tr>
<td>Expired prescriptions in EHR</td>
<td>2.5 (2.4)</td>
<td>2.9 (3.1)</td>
<td>2.7 (2.8)</td>
</tr>
<tr>
<td>Newly reported medications not in EHR</td>
<td>2.9 (2.6)</td>
<td>2.2 (2.7)</td>
<td>2.5 (2.7)</td>
</tr>
<tr>
<td>Education, count (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 12th grade</td>
<td>6 (5.9)</td>
<td>10 (9.3)</td>
<td>16 (7.7)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>26 (25.5)</td>
<td>22 (20.6)</td>
<td>48 (23.0)</td>
</tr>
<tr>
<td>Some college (no degree)</td>
<td>26 (25.5)</td>
<td>23 (21.5)</td>
<td>49 (23.4)</td>
</tr>
<tr>
<td>College degree</td>
<td>44 (43.1)</td>
<td>52 (48.6)</td>
<td>96 (45.9)</td>
</tr>
</tbody>
</table>

Abbreviations: EHR, electronic health record; SD, standard deviation.
information from other VA facilities and generated longer medication lists than those accessible in our local EHR. This may have diluted our ability to identify a treatment effect.

**Implications for Future Research and Policy Development**

Several hypotheses may account for the apparent absence of a PSE (i.e., the phenomenon where pictures are more likely to be remembered than words). First, we could not match images to non-VA and nonprescription medications. Second, the prescriptions and images were shown on screen, one medication at a time. The similarity in physical appearance of drugs manufactured and distributed in the United States may hamper correct identification. Showing the complete list on one screen might have helped participants disambiguate similar appearing drugs. Third, there is mixed evidence to suggest that the strength of the PSE decays with age. Finally, the use of computer technology may have caused an interference effect. It is critical that future consumer informatics research using pictures and visual displays isolate the design attributes that reliably promote attention, comprehension, and recall.

MR is a complex adaptive system that demands equally adaptive technologies to support practitioner workflow. Commensurate effort should be applied to interface usability, data interoperability, and clinician decision support. Federated health systems and managed care enterprises can promote these efforts by spearheading use of semantically rich medication terminologies such as Rx Norm. They can also help by enforcing EHR interoperability standards and funding regional health organizations. We believe our findings also emphasize the need to invest further in consumer informatics tools that engage patients and address U.S. EHR “meaningful use” standards. Stage 2 of the Health Information Technology for Economic and Clinical Health (HITECH) Act includes MR as one of the 25 criteria for functionality and Stage 3 expands scope to incorporate patient-entered data. These expectations, while laudable, may further stress systems,
inadvertently incentivizing organizations to implement less effective solutions such as interruptive reminders or boilerplate templates. Personal health records, mobile devices, secure messaging, and other consumer-driven technologies may provide time-sensitive alternatives to engage patients and collect information that might otherwise be skipped during a busy clinic visit. The effectiveness of these hi-tech solutions can be augmented by modest interventions focusing upon purposeful interface design, patient education best-practices, and provider interviewing strategies.

Finally, patient-centered assistive technologies like smart pillboxes, wearable devices, and Internet-enhanced living environments can improve the fidelity of our data streams.

**Conclusion**

Our study suggests that gathering patient-generated data using EHR-based technologies or pen-and-paper processes can be equally effective in supporting MR. We believe the technology and workflow described herein offer a practical, safe, and scalable method to foster collaboration between patients and care teams. We have offered a strategy that combines EHR technology, business process reengineering, and patient-generated data to augment traditional history collection and improve patient engagement. Future research should study how to promote MR technology adoption, improve patient self-reporting, and optimize use in specialty care settings.

**Clinical Relevance Statement**

Consumer informatics technology, such as self-service kiosks, offers a workflow-compatible solution to collect an accurate medication history and satisfy the Stage 2 Meaningful Use criteria. This randomized controlled trial shows that the patient-facing medication reconciliation software, when thoughtfully implemented using a systems engineering approach, can incorporate patient-furnished data and substantially improve discrepancy detection as compared with usual care. A variety of data collection strategies may be equifinal and further usability research is needed to

---

Table 4  Disagreement between the EHR list, treatment lists, and the reference standard

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Paper (102 participants)</th>
<th>APHID (107 participants)</th>
<th>Raw difference</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment versus EHR # medications</td>
<td>1,574</td>
<td>1,673</td>
<td>99</td>
<td></td>
</tr>
</tbody>
</table>

Proportion of list with discrepancies

| Total discrepancies, mean (SD) | 0.35 (0.20) | 0.35 (0.19) | 0.00 | 0.89 |
| High and very high risk, mean (SD) | 0.15 (0.13) | 0.17 (0.15) | 0.02 | 0.34 |
| System-based, mean (SD) | 0.14 (0.14) | 0.15 (0.13) | 0.01 | 0.64 |
| Participant-based, mean (SD) | 0.21 (0.21) | 0.20 (0.16) | 0.01 | 0.64 |
| Number of discordant lists – count (%) | 99 (97) | 104 (97) | 5 | 1.00 |

Treatment versus BPMH # medications

| Total discrepancies, mean (SD) | 0.35 (0.20) | 0.35 (0.19) | 0.00 | 0.89 |
| High and very high risk, mean (SD) | 0.15 (0.13) | 0.17 (0.15) | 0.02 | 0.34 |
| System-based discrepancies, mean (SD) | 0.14 (0.14) | 0.15 (0.13) | 0.01 | 0.64 |
| Participant-based discrepancies, mean (SD) | 0.21 (0.21) | 0.20 (0.16) | 0.01 | 0.64 |
| Number of discordant lists, count (%) | 99 (97) | 104 (97) | 5 | 1.00 |

Proportion of list with discrepancies

| Total discrepancies, mean (SD) | 0.13 (0.12) | 0.13 (0.13) | 0.00 | 0.90 |
| High and very high risk, mean (SD) | 0.04 (0.07) | 0.05 (0.07) | 0.01 | 0.67 |
| System-based discrepancies, mean (SD) | 0.04 (0.06) | 0.04 (0.06) | 0.00 | 0.97 |
| Participant-based discrepancies, mean (SD) | 0.06 (0.09) | 0.05 (0.07) | 0.01 | 0.31 |
| Number of discordant lists, count (%) | 78 (76) | 76 (71) | 2 | 0.43 |

EHR versus BPMH # medications

| Total discrepancies, mean (SD) | 0.43 (0.20) | 0.39 (0.17) | 0.04 | 0.12 |
| High and very high risk | 0.19 (0.14) | 0.18 (0.14) | 0.00 | 0.67 |
| System-based | 0.20 (0.17) | 0.17 (0.14) | 0.03 | 0.24 |
| Participant-based | 0.23 (0.20) | 0.22 (0.17) | 0.01 | 0.55 |
| Discordant lists, count (%) | 101 (99) | 106 (99) | 5 | 1.00 |

Abbreviations: APHID, Automated Patient History Intake Device; BPMH, best possible medication history; EHR, electronic health record; SD, standard deviation.

Note: Discrepancies are reported as the proportion of medications per list with an error. Statistics calculated using the participant as the unit of analysis. N = 209 participants; 3,500 medications.

The BPMH identified 253 additional medications not on the EHR or in either treatment.
understand how to effectively use medication images in consumer-facing interfaces.

Multiple Choice Questions

1. What do results of this randomized controlled trial indicate about self-service medication history collection software?
   a. The inclusion of medication images significantly improves patient accuracy when compared with a reference standard
   b. Kiosk technology does not perform as well as more traditional data collection methods such as distribution of paper-based questionnaires
   c. Patient-facing self-service technologies tend to be less accurate than usual care clinician-conducted interviews
   d. The accuracy of standardized methods for medication history collection including self-service kiosks and paper questionnaires are comparable when compared with a reference standard

Correct Answer: The correct answer is option d. This study did not demonstrate a statistically significant difference in the performance characteristics of the software as compared with a paper-based control that did not include images. Rather, discrepancy rates, sensitivity statistics, and negative predictive values were virtually identical when compared with the reference standard. It would not be completely accurate to say the technology was less effective than a clinician-mediated history since the software’s diagnostic performance was at least comparable—if not superior to—that reported for clinician-conducted histories.

2. Which statement best describes the distribution of medication history errors reported in this study?
   a. System-based sources of medication discrepancies are more common than patient nonadherence or recall errors
   b. An estimated 35 to 40% of collected medication histories include one or more discrepancies
   c. Approximately 15% of medication discrepancies are associated with expired prescriptions that the patient is still taking
   d. Between 22 and 25% of all medication discrepancies were classified as high-risk by a panel of blinded clinician raters

Correct Answer: The correct answer is option c. Nearly 49% of discrepancies were associated with current prescriptions; 15% were associated with expired medications, and 37% were newly identified over-the-counter medications. Slightly more than half of the errors (an estimated 53%) had a patient-based root cause such as nonadherence. Note that 41% of all medications reviewed were associated with a discrepancy and 99% of medication histories included one or more discrepancies when compared with the reference standard. Over 45% of the discrepancies detected were rated as high or very high risk.

3. When implementing medication history collection software as part of an organizational medication reconciliation strategy, what factors are likely to improve performance?
   a. The software is primarily effective in health systems where regional health information exchange is available
   b. An estimated 35 to 40% of collected medication histories include one or more discrepancies
   c. Approximately 15% of medication discrepancies are associated with expired prescriptions that the patient is still taking
   d. Between 22 and 25% of all medication discrepancies were classified as high-risk by a panel of blinded clinician raters

Correct Answer: The software is primarily effective in health systems where regional health information exchange is available.

Table 5 Discrepancy reporting accuracy for each treatment

<table>
<thead>
<tr>
<th></th>
<th>Control (paper)</th>
<th>Intervention (APHID)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPMH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrepancy (+)</td>
<td>474</td>
<td>500</td>
</tr>
<tr>
<td>No discrepancy (-)</td>
<td>108</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>936</td>
<td>979</td>
</tr>
<tr>
<td></td>
<td>530</td>
<td>594</td>
</tr>
<tr>
<td>Total = 1,574 medications</td>
<td>Total = 1,673 medications</td>
<td></td>
</tr>
<tr>
<td>Validity measure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>81 (78, 85)</td>
<td>83 (80, 86)</td>
</tr>
<tr>
<td>Specificity</td>
<td>94 (93, 96)</td>
<td>91 (89, 93)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>89 (87, 92)</td>
<td>84 (81, 87)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>90 (88, 91)</td>
<td>91 (89, 92)</td>
</tr>
<tr>
<td>Positive likelihood ratioa</td>
<td>8.46 (6.60, 10.86)</td>
<td>5.32 (4.40, 6.42)</td>
</tr>
<tr>
<td>Negative likelihood ratioa</td>
<td>0.12 (0.10, 0.14)</td>
<td>0.10 (0.08, 0.12)</td>
</tr>
</tbody>
</table>

Abbreviations: APHID, Automated Patient History Intake Device; BPMH, best possible medication history.
Note: An additional 253 medications were identified by the BPMH that were not identified in either treatment arm. Item-wise comparison cannot be calculated for medications not defined in both methods; they are not represented in the accuracy calculation. Percents rounded to nearest percentile. N = 3,247 medications.
*aLikelihood ratios weighted by prevalence.
b. The software should include patient-facing affordances that improve patient reliability including medication images, plain language, and simple response controls

c. The implementation team should pay attention to the sociotechnical fit of the product including workflow compatibility and implementation climate

d. The software should be used primarily in clinical settings where resource constraints preclude collection of a more time-intensive clinician-mediated history

Correct Answer: The correct answer is option c. This study seems to suggest that a variety of data collection methods and technologies can produce a high-quality history. Clinician engagement, standardized process, and patient-centered models may have a greater influence upon the accuracy and success of MR than any specific approach. While the availability of regional data exchange could certainly improve the quality of the history, this study shows that even without a data exchange network, the information gathered may considerably improve upon usual care. Unfortunately, the study did not show that the patient-facing interface or the inclusion of images conferred any additional benefit beyond a simple medication list. The accuracy of each treatment arm was at least comparable to the published statistics associated with a clinician-mediated history. Since results of this study suggest that combining several data sources substantially improves accuracy, a standardized patient-driven history collection method should be combined with a clinician-mediated history.

Authors’ Contributions

All the authors contributed equally to the study.

Protection of Human and Animal Subjects

This study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. Our local Institutional Review Board (IRB) approved this study. All participants provided written consent. We registered with ClinicalTrials.Gov (Identifier: NCT02135731) prior to data analysis. This study was funded in part by the VA National Center for Patient Safety.

Conflict of Interest

The authors report no conflicts of interest in the research; there are no plans to commercialize the software.

Funding

This study was funded by VA National Center for Patient Safety.

Acknowledgments

The authors wish to acknowledge the generous support from the VA Portland Healthcare System and the VA Chief Business Office. The authors also thank Richard Pham, PharmD, for assistance with the development and piloting our discrepancy classification instrument.

References

25 Dullabh PM, Sondheimer NK, Kathi E, Evans MA. How patients can improve the accuracy of their medical records. eGEMs (Generating Evidence & Methods to improve patient outcomes); 2014; 2(3):19. Available at: https://egems.ama-assn.org/article/abstract/10.13606/2327-9214.1080
36 Samarth A, Grant E. Using Health Information Technology to Perform Medication Reconciliation: Findings from the AHRQ Health IT Portfolio. In: Quality AHRQs, ed. Rockville, MD: AHRQ National Resource Center for Health Information Technology; 2010:5
Evaluation of Multimedia Medication Reconciliation Software

Lesselroth et al.

64 Lambert BL, Lin S-J, Chang K-Y, Gandhi SK. Similarity as a risk factor in drug-name confusion errors: the look-alike (orthographic) and sound-alike (phonetic) model. Med Care 1999;37(12):1214–1225
69 Jaeschke R, Guyatt GH, Sackett DL. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994;271(09):703–707
84 Kushneruk A, Santos SL, Pourakis G, Nebeker JR, Boockvar KS, Eds. Cognitive analysis of a medication reconciliation tool: applying laboratory and naturalistic approaches to system evaluation. ITCH 2011; February 2011; Victoria, Canada
86 Walsh EK, Cussen K. “Take ten minutes”: a dedicated ten minute medication review reduces polypharmacy in the elderly. Ir Med J 2010;103(08):236–238
93 Stephenson BJ, Rowe BH, Haynes RB, Macharia WM, Leon G. The rational clinical examination. Is this patient taking the treatment as prescribed? JAMA 1993;269(21):2779–2781

Applied Clinical Informatics Vol 9 No. 2/2018
Evaluation of Multimedia Medication Reconciliation Software

Lesselroth et al.

104 Grindrod K, Burns CM, eds. Medication reconciliation: more than just a check. 2015 International Symposium on Human Factors and Ergonomics in Health Care: Improving the Outcomes; April 26–29, 2015; Baltimore, MD: Human Factors and Ergonomics Society
114 Nazi KM. Veterans’ voices: use of the American Customer Satisfaction Index (ACSI) Survey to identify My HealtheVet personal health record users’ characteristics, needs, and preferences. J Am Med Inform Assoc 2010;17(02):203–211
### Medication Discrepancy Risk Scoring Tool

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Criteria</th>
<th>Risk Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Very High</strong></td>
<td>An error that can cause signs/symptoms that if not treated puts the patient at risk of death.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>High potential to cause cardiopulmonary arrest in the dose ordered.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medication listed as “high alert” by PVAMC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Chemotherapeutic agents (including methotrexate for non-oncologic use)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Botulinum Toxin type A injection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Concentrated (20 mg/ml) oral morphine &amp; oxycodone solution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Intravenous heparin, thrombolytics, and oral warfarin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Investigational Drugs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- I.V. Insulin &amp; SQ U-500 Insulin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Parenteral opiates, including epidurals</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Anti-arrhythmics, not including digoxin, metoprolol and diltiazem</td>
<td></td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>Can cause signs/symptoms associated with a serious level of risk, not life threatening.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cardiac, respiratory, antidepressant, anticonvulsant, hypoglycemic agent.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contraindicated secondary to a drug–drug interaction or drug–disease interaction.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manages a condition listed on the patient’s problem list.</td>
<td></td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td>Can cause patient symptoms, poses little or no threat to life function</td>
<td>1</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>Does not meet any of the criteria above.</td>
<td>0</td>
</tr>
</tbody>
</table>

**Step 1:** Does medication discrepancy meet any criteria for the Very High risk level?  
If yes, assign Risk Code 3. If not, proceed to Step 2.

**Step 2:** Does medication discrepancy meet any criteria for the High risk level?  
If yes, assign Risk Code 2. If not, proceed to Step 3.

**Step 3:** Does medication discrepancy meet any criteria for the Medium risk level?  
If yes, assign Risk Code 1. If not, proceed to Step 4.

**Step 4:** Assign Risk code 0.

---

*Appendix Fig. A1 Medication discrepancy risk scoring tool.*