Comparison of a new 3-item self-reported measure of adherence to medication with pharmacy claims data in patients with cardiometabolic disease

Julie C. Lauffenburger, PharmD, PhD,a,b Constance P. Fontanet, MPH,a,b Thomas Isaac, MD, MBA, MPH,c Chandrasekar Gopalakrishnan, MD, MPH,b Thomas D. Sequist, MD, MPH,d Joshua J. Gagne, PharmD, ScD,b Cynthia A. Jackevicius, PharmD, MSc,e Michael A. Fischer, MD, MS,b Daniel H. Solomon, MD, MPH, b,f and Niteesh K. Choudhry, MD, PhD a,b Boston, Newton, USA; and Toronto, Canada

Background Less than half of patients with cardiometabolic disease consistently take prescribed medications. While health insurers and some delivery organizations use claims to measure adherence, most clinicians do not have access during routine interactions. Self-reported scales exist, but their practical utility is often limited by length or cost. By contrast, the accuracy of a new 3-item self-reported measure has been demonstrated in individuals with HIV. We evaluated its concordance with claims-based adherence measures in cardiometabolic disease.

Methods We used data from a recently-completed pragmatic trial of patients with cardiometabolic conditions. After 12 months of follow-up, intervention subjects were mailed a survey with the 3-item measure that queries about medication use in the prior 30 days. Responses were linearly transformed and averaged. Adherence was also measured in claims in month 12 and months 1-12 of the trial using proportion of days covered (PDC) metrics. We compared validation metrics for non-adherence for self-report (average \( \overline{b} = 0.80 \)) compared with claims (PDC \( \overline{b} = 0.80 \)).

Results Of 459 patients returning the survey (response rate: 43.5%), 50.1% were non-adherent in claims in month 12 while 20.9% were non-adherent based on the survey. Specificity of the 3-item metric for non-adherence was high (month 12: 0.83). Sensitivity was relatively poor (month 12: 0.25). Month 12 positive and negative predictive values were 0.59 and 0.52, respectively.

Conclusions A 3-item self-reported measure has high specificity but poor sensitivity for non-adherence versus claims in cardiometabolic disease. Despite this, the tool could help target those needing adherence support, particularly in the absence of claims data. (Am Heart J 2020;228:xxx-xxx.)
accessing claims data, limiting their use for real-time decision-making.\textsuperscript{1,14} Existing self-reported adherence measures range from one-item questions inquiring about frequency of missed doses to longer multi-item assessments evaluating beliefs associated with adherence.\textsuperscript{13,15,20} The practicability of these measures is often limited by their cost (eg, licensing fees) or length (eg, patient burden).

One promising measure is the recent 3-item self-report scale by Wilson et al., which queries patients about how many days they missed medications and overall consistency over the last 30 days.\textsuperscript{21,22} Unlike other measures, this 3-item scale is short, easy to administer, and does not involve licensing fees. The 30-day recall also matches the most common interval for prescription dispensing by retail pharmacies (eg, 30 days) and would therefore hypothetically align well with claims.\textsuperscript{25} The accuracy of a 3-item self-reported adherence measure has been demonstrated in individuals with HIV compared to data collected from electronic pill bottles. However, despite this measure’s growing acceptance\textsuperscript{24}, this metric has not been compared directly with pharmacy claims data, nor studied in cardiometabolic disease.

Therefore, using data from a recently-completed pragmatic trial,\textsuperscript{25,26} we compared adherence collected via this 3-item self-reported metric\textsuperscript{22} versus pharmacy claims data as the referent in patients with cardiometabolic disease. We also explored the extent to which self-report adherence measures correspond with different claims-based adherence levels and thresholds and examined patient factors that might affect agreement between the measures.

## Methods

### Study design and patient population

We used data from the Study of a Tele-pharmacy Intervention for Chronic diseases to (2) Improve Treatment adherence (STIC2IT) pragmatic trial.\textsuperscript{25,26} This 2-arm trial was conducted at a large integrated delivery network in Massachusetts between August 2015 and August 2017. Subjects were included in the trial if they demonstrated suboptimal hyperlipidemia, hypertension, or diabetes disease control and were non-adherent to prescribed medications for these conditions.\textsuperscript{25,26} Patients were excluded from the trial if they had <6 months of plan enrollment or were <18 or >85 years of age before randomization. In total, 2038 patients were randomized to the intervention arm. As previously described\textsuperscript{25,26}, the central intervention component was a clinical pharmacist consultation. Patients were excluded from the present analysis if they lost continuous enrollment in the 12-month follow-up. The overall study design is in Appendix Figure 1.

### Measuring self-reported adherence through the 3-item scale

Patients in the intervention arm who received ≥1 clinical pharmacist consultation were administered a mailed survey 12 months after randomization to assess adherence. The survey consisted of a cover letter written on behalf of the patient’s primary care provider and a second page with the 3-item adherence measure. The cover letter also provided details on how to complete the survey online through Qualtrics, a Web-based survey platform. The surveys were otherwise anonymized and contained scrambled patient identifiers to link with claims data. Individuals not returning the survey within 30 days were sent a second copy with a $5 gift card. In the rare event that patients returned both rounds, we included only the first round in the analyses.

The 3 self-report items queried patients about the number of days on which medication was taken, frequency of use, and overall ratings of adherence, respectively, as follows:

1. Days Taken: “In the last 30 days, on how many days did you miss at least one dose of any of your medications?” (number of days: 0-30). As previously done, this item was converted to “days taken” by subtracting the number of days missed from 30.\textsuperscript{21}
2. Frequency: “In the last 30 days, how often did you take your medications in the way you were supposed to?” (never/rarely/sometimes/usually/always/always)
3. Overall Rating: “In the last 30 days, how good a job did you do at taking your medications in the way you were supposed to?” (very poor/poor/fair/good/very good/excellent)

As done in previously-published studies, we linearly transformed the 3 self-reported items to a 0-100 scale, with zero being the lowest adherence, and 100 being the best.\textsuperscript{21} This means, for example, that the Frequency and Overall Rating items that had 6 levels were transformed to 0, 20, 40, 60, 80, and 100, respectively and that the Days Taken item was transformed to units of 3.33 for each (ie, 100/30), so Days Taken = 29 was transformed to 96.67%. An average was then calculated as a mean of the 3 individual items on this 0-100 scale.\textsuperscript{21} If there was missing data (<4%) for any individual item, we took the mean of the non-missing items.

### Pharmacy claims-based adherence

We created a drug supply diary from prescription claims data from the date of randomization until the end of the 12-month follow-up for each medication used for any of the 3 study conditions.\textsuperscript{27} Medications in the same therapeutic class were considered interchangeable (eg, 2 \(\beta\)-blockers). Supply for overlapping fills could accumulate up to 180 days excess supply. From these supply diaries, for each eligible medication, we then measured the proportion of days that patients had medications available, or the proportion of days covered (PDC) within each month of
follow-up and the entire 12-month period. Using an “average of averages” approach, we averaged this PDC across all medications used to treat a single condition (eg, hypertension) and then calculated an overall average adherence across each patient’s eligible conditions.28,29

We used this “average of averages” calculation to generate 2 different continuous measures of claims-based adherence, including (1) PDC in the 12 months after randomization (“PDC in months 1-12”) and (2) PDC in the 12th month after randomization because this period overlapped with the measurement by self-report (“PDC in month 12”). The 3-item survey and claims-based metrics both measure the implementation phase of adherence.30

Covariates

We used claims and electronic health record (EHR) data to measure potential predictors of adherence during the 365 days before randomization. If patients had less than 1 year of insurance eligibility, covariates were measured upon enrollment. Using EHR data, we measured age, sex, race/ethnicity (classified as “white”/“non-white”), marital/partner status, Body Mass Index (BMI), and current tobacco use. Using International Classification of Disease 9th edition (ICD-9) or 10th edition (ICD-10) codes in medical claims data, we measured comorbidities, a combined comorbidity score (measure of overall health status),31,32 number of outpatient physician office visits, hospitalizations, and emergency room (ER) visits. Using pharmacy claims data, we measured the number of unique medications each patient filled.27 By definition, no factors measured in claims were missing because insurance coverage was a trial criterion.25,33 While the overall proportion missing BMI was low (<5%), we imputed using mean values for those without one.35

Statistical analyses

We first described characteristics of survey responders and non-responders and calculated absolute standardized differences. Then, we examined distributions of adherence measured in claims by levels of self-reported adherence on the 3 self-reported items using box plots. As in prior work, we aggregated adjacent categories when there were a small number of responses in a given category (eg, Very Poor/Poor).21

To enable dichotomous comparisons, in claims data, we classified patients with a mean PDC <0.80 as “non-adherent”, the most commonly-used threshold for “non-adherence”.34 In the survey data, we classified patients who self-reported as having taken medication on <24 days per month as “non-adherent” (ie, the analogue to PDC = 0.80, 24/30 = 0.80).21 Similarly, we classified patients who responded with “almost always/always” for frequency and “very good/excellent” for overall ratings as “adherent”. For the average 3-item score, we classified patients <80% as “non-adherent”. We used these cutpoints to calculate validation metrics using non-adherence using claims measures as the standard. Sensitivity provides the ability of self-report to identify “true positives” (ie, proportion non-adherent in the survey measure of those non-adherent in claims). Specificity provides the ability of self-report to identify “true negatives” (ie, proportion adherent in the survey measure of those adherent in claims). We also calculated positive and negative likelihood ratios and positive and negative predictive values (PPV, NPV). We also examined patient characteristics by their concordance and for these analyses stratified age into 3 categories (<55, 55-64, ≥65 years), BMI as <30 or ≥30, combined comorbidity score ≥1 or <1, office visits as <2 or ≥2, and unique medications as <5, 5-10, ≥10.32 We also conducted exploratory subgroup analyses by chronic condition qualifying the patient for trial entry.

Finally, we conducted sensitivity analyses, including restricting to those who returned the first survey and excluding those with a PDC = 0 in the measurement window to reduce the possibility of inaccurate claims for that patient.35 To explore the optimal definition of “non-adherence”, we varied the thresholds for “non-adherence” for the average 3-item score and claims. Analyses were conducted using SAS 9.4 (Cary, NC). The Institutional Review Board of Brigham and Women’s Hospital approved this study. The trial is registered on ClinicalTrials.gov (NCT02512276).

This research was supported by an NIH grant (R01HL117918). The authors are solely responsible for the design and conduct of this study, all study analyses and drafting and editing of the paper.

Results

In the trial, 1069 intervention patients received the initial pharmacist consultation and were mailed the survey (52.4% of intervention group). Of these 1069, 465 patients (43.5%) returned the surveys; >50% within 1 week. Of these, 6 (1.3%) surveys did not contain usable study identifiers to link their responses to claims data and were excluded. As a result, 459 patient surveys were included in analyses.

The mean age of those who responded was 64 years, and 57% were male (Appendix Table I). Compared with non-responders, patients who returned the survey were slightly older, more likely to be white, have a partner, have a diagnosis of hyperlipidemia or osteoporosis, and visit the ER in the prior 12 months. PDC adherence was also slightly higher among responders.

In month 12, mean claims-based PDC was 56.8 (SD: 43.4), median 76.7 (IQR: 0.0-96.7). Over months 1-12, mean PDC was 57.7 (SD: 32.9), median 65.5 (IQR: 33.2-85.5). For the 3-item score, the mean was 87.4 (SD: 13.8), median 91.1 (IQR: 83.3-100.0). The distribution of responses to each survey question is shown in Appendix Table II. The
percentages of patients classified as “non-adherent” by self-report and pharmacy claims are in Figure 1. In month 12, fewer patients were classified as “non-adherent” by patient self-report (20.9%) than claims (50.1%).

The raw self-reported items and PDC measures for each item level are compared in Figure 2. Mean PDC generally increased as self-reported adherence increased. The notable exception is that mean PDC in months 1-12 was slightly higher in those reporting taking 29 days of medication versus those taking all 30 days. Patients who did not report adherence were much less likely to be adherent in claims.

Validation metrics for identifying non-adherence are shown in Table I. Sensitivity was relatively low for identifying non-adherence (eg, 0.25 for average of the 3-item score for month 12). By contrast, specificity was high for identifying non-adherence (ie, $\geq 0.83$ for all comparisons). PPV and NPV were modest (eg, 0.59 and 0.25, respectively for month 12). These metrics were similar when comparing self-report with adherence measured in claims across both intervals (month 12 and months 1-12), although PPV was higher for months 1-12 (0.80). For the self-reported items individually, the “Days Taken” metric appeared to perform best versus claims.

Baseline characteristics of patients by whether their survey and months 1-12 claims-based PDC agree are shown in Appendix Table III. In general, patients who were classified as optimally adherent by both measures tended to be male and older. Patients who were non-adherent by both measures tended to be younger, non-white, single/unpartnered, and have had fewer office visits. However, these inferences were limited by small sample size.

Exploratory subgroup analyses by eligible study condition are shown in Appendix Table IV. There were no substantial differences across subgroups, but conclusions were limited by small sample size. Restricting to patients who returned the surveys in the first round slightly improved but did not materially change concordance (Appendix Table V).

Variations of the 3-item self-reported score and claims thresholds for non-adherence suggest that accuracy may improve with alternative cutpoints (Appendix Table VI). Specificity was extremely high for all claims-based thresholds at survey thresholds at or below $<70\%$ (Figure 3).

Discussion

In this study of patients with cardiometabolic disease, we observed that self-reported adherence was consistently higher than claims-based adherence. In addition, patients reporting non-adherence were much less likely to be adherent in claims. Using pharmacy claims as the referent, the 3-item self-report measure had high specificity but very low sensitivity for non-adherence; PPV and NPV suggest that there may be false positives or negatives, depending on prevalence of non-adherence. Survey metrics were similar for short (1-month) and longer (12-month) claims measurement windows.

To our knowledge, this is the first study to directly compare this self-report metric with pharmacy claims in patients with cardiometabolic disease. This item has become quickly embraced by many in the HIV community, but has not been studied to our knowledge in cardiometabolic diseases or in claims.24 Previously-developed metrics have been limited by their length, difficulty of administration, or...
licensing fees. Concerns have been raised about other metrics that query about the extent of non-adherence while also asking about reasons for non-adherence. Other metrics ask about use during shorter time windows, yet prior work has suggested that a 1-month recall window may best approximate adherence—ie, the recall window in this 3-item metric. These results have several important insights relevant to clinical care. First, patients who self-identify with any degree of non-adherence also appeared to have suboptimal adherence in claims. This finding is not necessarily surprising given the overall tendency of patients to over-report adherence. The creators of the 3-item scale do not advocate for specific levels at which patients should be considered sub-optimally adherent, so any admission of non-perfect adherence could therefore prove potentially actionable. Sensitivity analyses of non-adherence cutpoints support this. The thresholds themselves therefore may not perfectly mirror claims data calculations, as for example, patients using multiple medications may report differently than how adherence was calculated in claims. The 3-item scale also appears to perform similarly for patients' average adherence in the prior year as in the prior month.

While some metrics are modest, this brief 3-item self-report scale is inexpensive to administer and could be a useful tool in clinical care or population health management. Its strong specificity for identifying non-adherence may lend itself to some desirable properties, especially when pharmacy claims data may be unavailable or delayed. In clinical care, this self-report scale could be best used as a brief screening tool to identify patients with suboptimal adherence who would be candidates for more in-depth discuss about how to address their barriers. In patients with suspected non-adherence, due to the poor sensitivity, it may also be useful to assume non-adherence and then use the survey to confirm non-adherence. Of course, if employed as is, resulting false positive and false negative rates may still misclassify patients. Implementation strategies would need to recognize that because of the low sensitivity, many potential candidates would not be identified. Given that non-adherence to cardiometabolic medications is highly prevalent in real-world practice, identifying even a subset of patients for intervention represents an improvement.

Our results largely corroborate the diagnostic parameters observed in prior examinations of other self-reported metrics. For example, our finding that self-reported adherence was higher than adherence measured in

---

**Figure 2**

Adherence measured in pharmacy claims by levels of the 3-item patient self-reported adherence score. **A**, Self-reported 3-item adherence: Days Taken item. **B**, Self-reported 3-item adherence: Frequency item. **C**, Self-reported 3-item adherence: Overall Rating item. **D**, Self-reported 3-item adherence: Average.
pharmacy claims is consistent with evaluations of other tools. It is impossible to distinguish whether the higher estimation by self-report in general is reflective of social desirability bias or perhaps an underestimation of adherence by claims. Regardless, prior research has shown that while adherence measured via self-report and claims may not correlate well with each other, both are separately correlated with blood pressure control, suggesting that the measures provide complementary information about patients’ medication-taking.

The modest associations observed in this work and others could be explained in several ways. First, the metrics could functionally be measuring different things; for example, pharmacy claims-based adherence measures long-term adherence, whereas self-report measures short-term behaviors. Second, the results could be a function of how claims-based adherence is averaged, yet adherence changes over time and therefore standard deviations are fairly wide. Prior work has shown that calibrating to the specific comparison data source can improve accuracy of self-reported scales by reducing overall overestimation by self-report. Of note, the metrics studied here were for "non-adherence"—if applied to identifying "optimal" adherence (eg, PDC ≥ 0.80), they would be reversed. Finally, the scale does not perform as well in cardiometabolic disease as HIV at the threshold used in prior studies but may perform better at alternative cutpoints.

Several limitations should be acknowledged. First, patients were part of a larger quality improvement trial conducted in one region, eastern Massachusetts, and patients had to have a pharmacist consultation, prior non-adherence and poor control to receive a survey, which may have had their own inaccuracies, such as potential for missing claims when filled outside insurance; for example, excluding patients with PDC = 0 slightly improved metrics. Prior work has shown that calibrating to the specific comparison data source can improve accuracy of self-reported scales by reducing overall overestimation by self-report. Of note, the metrics studied here were for "non-adherence"—if applied to identifying "optimal" adherence (eg, PDC ≥ 0.80), they would be reversed. Finally, the scale does not perform as well in cardiometabolic disease as HIV at the threshold used in prior studies but may perform better at alternative cutpoints.

Several limitations should be acknowledged. First, patients were part of a larger quality improvement trial conducted in one region, eastern Massachusetts, and patients had to have a pharmacist consultation, prior non-adherence and poor control to receive a survey, Table I. Accuracy of self-reported non-adherence compared with pharmacy claims.

<table>
<thead>
<tr>
<th>Claims PDC</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive likelihood ratio (95% CI)</th>
<th>Negative likelihood ratio (95% CI)</th>
<th>Positive predictive value (95% CI)</th>
<th>Negative predictive value (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days Taken item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-adh</td>
<td>31</td>
<td>12</td>
<td>(0.09-0.18)</td>
<td>(0.91-0.97)</td>
<td>(1.36-4.88)</td>
<td>(0.86-0.97)</td>
</tr>
<tr>
<td>Adherent</td>
<td>199</td>
<td>217</td>
<td>0.95</td>
<td>2.57</td>
<td>0.91</td>
<td>0.72</td>
</tr>
<tr>
<td>Months 1-12, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-adh</td>
<td>36</td>
<td>7</td>
<td>(0.08-0.15)</td>
<td>(0.92-0.99)</td>
<td>(1.06-5.10)</td>
<td>(0.88-0.98)</td>
</tr>
<tr>
<td>Adherent</td>
<td>280</td>
<td>136</td>
<td>0.95</td>
<td>2.33</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-adh</td>
<td>31</td>
<td>19</td>
<td>(0.09-0.18)</td>
<td>(0.88-0.95)</td>
<td>(0.95-2.79)</td>
<td>(0.88-1.01)</td>
</tr>
<tr>
<td>Adherent</td>
<td>199</td>
<td>210</td>
<td>0.92</td>
<td>1.62</td>
<td>0.94</td>
<td>0.62</td>
</tr>
<tr>
<td>Overall rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-adh</td>
<td>73</td>
<td>22</td>
<td>(0.18-0.28)</td>
<td>(0.79-0.91)</td>
<td>(0.97-2.32)</td>
<td>(0.83-1.00)</td>
</tr>
<tr>
<td>Adherent</td>
<td>243</td>
<td>121</td>
<td>0.85</td>
<td>1.50</td>
<td>0.91</td>
<td>0.77</td>
</tr>
<tr>
<td>Average 3-item score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-adh</td>
<td>57</td>
<td>39</td>
<td>(0.19-0.31)</td>
<td>(0.77-0.88)</td>
<td>(1.01-2.09)</td>
<td>(0.82-1.00)</td>
</tr>
<tr>
<td>Adherent</td>
<td>173</td>
<td>190</td>
<td>0.83</td>
<td>1.46</td>
<td>0.91</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Abbreviations: PDC, Proportion of days covered; Non-adh, Non-adherent; CI, Confidence Interval.
Note: Three-item score linearly transformed to 0-100 scale and then averaged. “Non-adherence” was defined as follows: Claims PDC < 80%, Days Taken < 24, Frequency: Never/Rarely/Sometimes/Usually, Overall Rating: Very poor/Poor/Fair/Good, and Average 3-item score < 80%.
potentially reducing generalizability. However, trial patients with these characteristics may be a key population of interest for this scale), and patients who participated in the consultation were generally similar to those in the parent trial. A 44% response rate is modest; those responding may be more subject to social desirability bias, which may affect generalizability. Finally, sample sizes by study conditions were small.

Conclusion

This new 3-item self-reported adherence metric had high specificity but poor sensitivity versus pharmacy claims data in patients with cardiometabolic disease. Patients who self-report any degree of non-adherence in this metric are also highly likely to be non-adherent based on claims data.

Acknowledgements

We wish to thank the clinical pharmacists at Atrius for their support with the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ahj.2020.06.012.

References

Lauffenburger et al


