Self-efficacy for medication management: a systematic review of instruments

Background: Medication self-efficacy is a potentially important construct in research around optimal use of prescription medications. A number of medication self-efficacy measures are available; however, there is no systematic review of existing instruments and cataloguing of their theoretical underpinnings or psychometric properties, strengths, and weaknesses. The aim of the study was to identify instruments that measure self-efficacy for medication management. The study also aimed to examine the quality, theoretical grounding, and psychometric evaluation of existing measures of self-efficacy for medication management. The study was a systematic review.

Methods: Data were extracted from PubMed, OVID, and MEDLINE using a predefined search strategy. Citations were included if they reported the development and/or psychometric evaluation of an instrument to measure self-efficacy for medication management and were in English. Abstracts were screened for studies potentially meeting eligibility criteria. Full articles of these studies were then reviewed in depth. The review was carried out independently by two members of the research team.

Results: The search identified 158 citations of which 12 were included after screening. Full review identified 3 articles fitting inclusion criteria for the review. Generally, development was theoretically grounded and included patients and experts in the field. Psychometric testing showed evidence of internal consistency (2/3 instruments) and test–retest reliability (1/3 instruments). All instruments showed some validity; however, assessment of all forms of validity for each instrument was lacking.

Conclusion: Although our analysis would recommend the use of the Self-Efficacy for Appropriate Medication Use Scale because of the current evidence of validity and reliability, more psychometric evaluation is required, particularly in terms of responsiveness to change as self-efficacy is a malleable patient-level factor. Three measures of self-efficacy for medication management were identified. Overall, some evidence of reliability and/or validity was demonstrated for all instruments; however, other forms of validity were not tested (ie, responsiveness to change). Use of a well-validated measure of self-efficacy medication management is essential in order to understand relationships between medication self-efficacy and other patient-reported outcomes such as patient-centeredness, patient enablement, and burden of treatment, an important area of research that is currently lacking.

Keywords: self-efficacy, polypharmacy, medication, treatment burden

Introduction
Medication non-adherence is prevalent. Medication non-adherence may indicate a mismatch between patient priorities and medical priorities, or if medication regimes are complex, it may reflect a mismatch between burden of treatment and individual capacity. There is a general consensus that medication non-adherence is related to...
number of poor health and clinical outcomes6–9 and can lead to avoidable health care costs in the health care system.10

There are a number of factors associated with medication non-adherence.11 In a recent systematic review,11 five broad barriers were identified related to medication adherence: patient factors (eg, mental and health state; knowledge and beliefs; demographic variables such as age, sex, and education), medication factors (eg, packaging, drug storage, cost, labeling instructions, regimen complexity), physician factors (eg, poor communication, lack of involvement, trust, dissatisfaction), system-based (eg, lack of medication review, follow-up), and other factors (eg, lack of caregiver, caregiver burden, lack of immediate improvement in health state).

Several behavior change theories have been applied to the understanding of medication adherence and attempt at changing medication-related behavior to improve adherence. The concept of self-efficacy within Social Cognitive Theory (SCT)12 offers a practical framework from which to work. Specifically, self-efficacy is one’s belief that they can successfully perform a specific behavior to achieve a desired outcome. Self-efficacy is considered a person factor within SCT, one that is malleable through the following four sources: mastery experiences, vicarious experiences, verbal persuasion/encouragement, and interpretation of emotional and physiological states.2,13 Although self-efficacy can be a powerful force in behavior change, theorists highlight inherent limitations of its influence as well as the extent to which SCT can explain human behavior.14 For example, it is assumed that self-efficacy beliefs determine action, with the caveat that the person must have the appropriate skills and adequate incentives for performance;15 however, even with the resources and high level of self-efficacy, some people may still not perform the behavior. Despite such limitations, there is a general consensus that self-efficacy is one of the most potent predictors of behavior change and disease self-management.13,16–21

However, the concept of self-efficacy cannot be taken out of context. By its very definition, self-efficacy is context-specific. Many factors, as already noted, impact one’s level of self-efficacy, particularly with respect to medication management. These factors may add to one’s level of treatment burden. The concept of minimally disruptive medicine4,5 is one that marries well with patient self-efficacy for medication management. Minimally disruptive medicine is an approach to care that is based on theory and patient-centeredness, and is context-sensitive with the aim of imposing the smallest possible treatment burden on patients’ lives. One’s level of self-efficacy for medication management may be a proxy assessment of their capacity for treatment. Thus, this highly malleable person factor (self-efficacy) can be both a pathway to improve adherence to prioritized medications and an indicator of match of treatment burden to capacity.

While it is important to have valid and reliable scales to measure medication adherence,22 having valid and reliable instruments to measure one’s self-efficacy for medication management is also important; such instruments may enhance our understanding of, and intervention for, improved medication management. A number of self-report surveys have been developed, evaluated, and used; however, it has been suggested that the majority of these measures are disease specific and do not apply across a range of chronic disease.23 This is an important shortcoming as multimorbidity and comorbidity are a more common reality in clinical practice. It is also not uncommon that multimorbidity and comorbidity are accompanied by factors associated with medication non-adherence, thus making the understanding of medication self-efficacy even more important.

Objectives
The aim of the study was to identify existing instruments that measure self-efficacy for medication management. Secondly, for those instruments identified, the study aimed to examine the theoretical grounding and psychometric evaluation of existing measures of self-efficacy for medication management.

It should be noted that we decided not to explicitly exclude or include instruments that may have been developed and validated in populations characterized by a disease in our search strategy. Rather, we wanted to capture all potential measures of self-efficacy for medication management and adherence. To help us contextualize and understand the findings, where appropriate, search results are differentiated between disease-specific versus non-specific measures of self-efficacy for medication management. However, the focus of this literature review was on non-disease specific measures of self-efficacy medication management.

Methods
The PRISMA guidelines for reporting items for systematic reviews and meta-analyses were followed. A search was completed in PubMed, OVID, and EMBASE to identify abstracts published in English up to May 2017. PubMed was the first database used in the search, with OVID and EMBASE being used as additional sources to identify unique abstracts not identified through PubMed. A Google search was also used to identify articles in the gray literature as well as manually searching the reference lists of included and
excluded articles to identify other possible relevant articles. The search terms and strategy are outlined in Table 1. The strategy was determined by the research team with advice from a librarian.

Abstracts were retrieved and screened. Full articles for which the abstract was determined to be potentially relevant to the search objectives were retrieved and reviewed. Those meeting inclusion criteria (English, development, and/or psychometric evaluation of an instrument to measure self-efficacy for medication management) were retained in the analysis (Table 2 for characteristics of included studies). Exclusion criteria included those abstracts using a measure of self-efficacy for medication management without explicitly assessing the instrument’s psychometric properties. No inclusion or exclusion criteria were included to define study population of the articles. The process of reviewing abstracts and full articles was completed independently by the first two authors. When reviewing full articles, an initial checklist of evaluative criteria to assess the instrument’s developmental measurement properties was used (Table 3).24 Specifically, the checklist was used to evaluate the conceptual model, content validity, reliability, construct validity, scoring and interpretation, and respondent burden and presentation. Each item was scored as 0 or 1, based on whether or not the criterion in question was present in the article. We also used the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist as a means to identify any assessment gaps in surveys in order to identify needs for further research on measurement properties.25

Results

Article selection

Across searches in databases and the gray literature, as well as manual searches of reference lists, the search resulted in 158 citations, of which 12 were identified as potentially relevant abstracts with respect to the search objective and underwent full review. Upon assessment of the full-text article for eligibility, 9 were excluded. Reasons for exclusions are outlined in Figure 1. The remaining 3 articles were included in our synthesis23,26,27 and include review of the Medication Understanding and Use Self-Efficacy Scale (MUSE),26 the Self-Efficacy for Appropriate Medication Use Scale (SEAMS),23 and the Long-Term Medication Behavior Self-Efficacy Scale (LTMBSES).27 It should be noted that validity of the LTMBSES was explored in a separate article28 than its development.27 As such, the article on tool development was also included and interpreted in combination with the article on tool validity. Figure 1 shows the results and flow of the search, identification, and screening process. There were no discrepancies between the two reviewers with respect to inclusion/exclusion of articles for full review.

Characteristics of selected articles

The study sample sizes ranged from 267 to 1,021 participants. The studies validating the MUSE and SEAMS recruited patients from a primary care center, while data were pooled from six studies to validate the LTMBSES. The study validating the MUSE had a sample with no homogenous

Table 1 Search terms

<table>
<thead>
<tr>
<th>Search terms</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(medication adherence[MeSH Terms]) OR medication adherence[Title/Abstract]</td>
<td>Final search = (((medication adherence[MeSH Terms]) OR medication adherence[Title/Abstract]) AND ((Self efficacy[MeSH Terms]) OR self efficacy[Title/Abstract]) AND (surveys and questionnaires[MeSH Terms]) )</td>
</tr>
<tr>
<td>(Self efficacy[MeSH Terms]) OR self efficacy[Title/Abstract]</td>
<td></td>
</tr>
<tr>
<td>(Surveys and questionnaires[MeSH Terms])</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Name of survey</th>
<th>Language</th>
<th>Population</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cameron et al26</td>
<td>MUSE scale</td>
<td>English</td>
<td>N = 267</td>
<td>Outpatients attending primary care in three USA cities (Chicago, New York City, and Shreveport)</td>
</tr>
<tr>
<td></td>
<td>adults; age = 47.9 years (±13.4); 20% less than high school education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risser et al23</td>
<td>SEAMS</td>
<td>English</td>
<td>N = 436</td>
<td>Primary care clinic in one USA city (Atlanta)</td>
</tr>
<tr>
<td></td>
<td>individuals with coronary heart disease; age = 63.80 (±10.4); 91.12% African American; 53% completed high school</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Geest et al27</td>
<td>LTMBSES</td>
<td>English</td>
<td>N = 1,021</td>
<td>Reported in individual studies, with subjects from the USA, Belgium, and the Netherlands</td>
</tr>
<tr>
<td></td>
<td>Dutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>French</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Checklist of evaluative criteria to assess the instrument’s developmental measurement properties

<table>
<thead>
<tr>
<th>Checklist item</th>
<th>Score</th>
<th>Cameron et al28</th>
<th>Risser et al29</th>
<th>De Geest et al30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual model</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Validity</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1. The construct to be measured has been defined</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2. The intended respondent population has been described</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3. The conceptual model addresses whether a single construct/scale or multiple subscales are expected</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4. There is evidence that members of the intended respondent population were involved in the measure’s development</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5. There is evidence that content experts were involved in the measure’s development</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6. There is description of the methodology by which items/questions were determined</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Construct</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7. There is reported quantitative justification that single scale or multiple subscales exist in the measure</td>
<td>1</td>
<td>1</td>
<td>1c</td>
<td></td>
</tr>
<tr>
<td>8. There are findings supporting expected associations with existing measures or with other relevant data</td>
<td>1</td>
<td>1</td>
<td>1c</td>
<td></td>
</tr>
<tr>
<td>9. There are findings supporting expected differences in scores between relevant known groups</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10. The measure is intended to measure change over time? If YES, there is evidence of both test–retest reliability AND responsiveness to change. Otherwise, award 1 point if there is an explicit statement that the measure is NOT intended to measure change over time</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11. There is evidence that the measure’s reliability was tested</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>12. Reported indices of reliability are adequate and/or justified</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Scoring and interpretation</td>
<td></td>
<td>0</td>
<td>1c</td>
<td></td>
</tr>
<tr>
<td>13. There is documentation for how to score the measure</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14. A plan for managing and/or interpreting missing responses has been described</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>15. Information is provided about how to interpret the measure scores, normative data, and/or a definition of severity</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1c</td>
</tr>
<tr>
<td>Qualities related to the respondent</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16. The time to complete the measure is reported and reasonable. If it is not reported, the number of questions appropriate for the intended application is provided</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>17. Is there a description of the literacy of the measure?</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18. The entire measure is available for public viewing</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total score (/18)</td>
<td>10</td>
<td>14</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table is based on a table by Patel et al.25 Scores represent whether or not the information is provided in the citation/source document (0 = criteria met, 1 = criteria not met, + = Scores for this survey combined the review of two articles.27,28 c = Addressed by Denhaerynck et al.29

qualities around disease states, while the SEAMS had a sample of patients with coronary heart disease and the LTMBSES pooled samples of transplant, hyperlipidemia, and AIDS/HIV patients. The geographic reach of validation varied. MUSE recruited participants from three USA cities, SEAMS from one USA city, and LTMBSES from USA, Belgium, and the Netherlands, using the scale in three different languages. Summary of characteristics of each article is given in Table 2.

Scale development

The MUSE was created through selection and modification of items from an existing instrument, the Communication and Attitudinal Self-Efficacy – Cancer scale.29 Items relating to patients’ self-efficacy in the action of taking their medications were added, resulting in an 18-item scale. Items were written at a 6th grade reading level or lower, as per Lexile analysis.30 The 18 items were pilot tested, first with exploratory factor analysis, which eliminated 9 items due to low factor loading. Initial principal components factor analysis was then done, eliminating one more item due to low factor loading, leaving 8 remaining items. Two subscales emerged – the first with 4 items associated with taking medication and the second with 4 items associated with learning about medication.

Similar to the MUSE, item generation for the SEAMS was done from existing instruments. Specifically, items were generated from questions of multiple instruments, giving priority to items that were shown to perform well in published psychometric analyses. From there, the authors generated new items to fill gaps. A multidisciplinary team with relevant expertise modified these items, rephrasing items from the literature search to simplify the wording and make questions more generalizable, as opposed to pertaining to a single medical condition. Patient interviews were conducted to evaluate the thought process of participants in interpreting
and formulating answers to scale items. Wording suggestions from patients were incorporated into the development of the 21-item scale. The 21 items were evaluated by factor analysis and were eliminated if they did not load well on the medication self-efficacy construct. Exploratory principal component factor analysis further reduced the number of items. This evaluation yielded a final instrument with a 13-item scale.

Unlike the MUSE and SEAMS, items of the LTMBSES were generated from interviews using a phenomenological approach. Specifically, 14 chronically ill patients were asked to share their opinions about the knowledge and skills needed to take medication correctly and about situations where it was difficult to take medications correctly. Using themes that emerged from these interviews, along with themes from Bandura’s framework, a self-efficacy scale was generated. The clarity of the items was evaluated by nurses and a physician seasoned in working with patients with chronic disease. This resulted in a 33-item scale. Psychometric analysis was done in three different patient groups, and questions on side effects were tailored to the specific medications of these groups.

Evaluation of psychometric properties
The overall scores for the quality of the scales were 14, 10, and 12 out of a possible 18 points for the SEAMS, MUSE, and LTMBSES, respectively (Table 3). Two quality checklist items were missed by all three validation studies: identifying if it is expected that the conceptual model addresses a single construct, or multiple subscales, and whether the scale is to measure change over time (and as assessment of both test-retest reliability and responsiveness to change). One article reported the number of missing values, but description as to how to handle incomplete surveys was not included.

Validity
Construct validity (structural) of the MUSE was measured through exploratory factor analysis. Cross-loaded items or those with a loading below 0.60 were eliminated, resulting in a 9-item scale. In an initial principal components factor analysis, two factors were identified. One more item was removed because of a factor loading less than 0.30 on both identified factors. This resulted in an 8-item scale. Predictive
validity was also assessed by testing participants’ ability to take a medication correctly using mock pill bottles and determining the association with the MUSE score. For this test, 10 mock pill bottles with common prescription label instructions were given to participants; participants were rated on their ability to teach the research assistant how and when to take the medications. The hypothesized relationship was supported such that higher scores on the MUSE were significantly positively correlated in moderate to large magnitude with taking and learning about prescription instructions, with 55% of the variance of taking and learning about medication explained by MUSE scores.

Construct validity (structural) of the SEAMS was assessed through an exploratory factor analysis. This analysis yielded four factors; however, the authors were unable to find an underlying dimension of self-efficacy for each of the four factors, and therefore they were not divided into identifiable domains. Given this, the authors examined the inter-item correlations to determine poor performance items, resulting in a 13-item scale. Factor analysis of the 13-item scale showed two separate factors: self-efficacy for taking medications under difficult circumstances (factor 1) and self-efficacy for continuing to take medications when circumstances relating to medication taking were uncertain (factor 2). Generally, the factor structure was the same regardless of literacy level, with the exception of item 10 loading on factor 1 versus factor 2 in low literacy patients rather than high literacy patients. Criterion-related validity of the 13-item SEAMS was also analyzed by comparing scores on the 4-item Morisky adherence scale. As expected, the two scales were significantly strongly positively correlated (Spearman’s \( \rho = 0.51, \rho = 0.0001 \)). Criterion-related validity was also assessed by examining differences in SEAMS scores between patients with controlled and uncontrolled blood pressure as well as examining differences in high versus low self-efficacy patients (based on SEAMS scores) and blood glucose levels. Findings showed that SEAMS scores between the two blood pressure groups were not significantly different, nor were blood glucose levels between high and low self-efficacy groups; thus, the hypotheses were not supported.

Construct validity (structural) of the LTMBSES was measured through principal component analysis, which showed that all items were loaded onto the single factor (>0.43). Further, criterion validity was assessed and showed that scores on the LTMBSES were related to medication adherence (\( p < 0.0001 \)) and medication taking compliance (ie, average number of pills taken; \( p < 0.0001 \)) through Generalized Estimating Equations model. However, area under the curve using receiver operating characteristics curve showed that the LTMBSES had poor predictive capability.

**Reliability**

Internal consistency of the MUSE was measured with Cronbach’s alpha, which was computed separately for the two subscales. Factor 1, relating to taking medication, had a Cronbach’s alpha of 0.77 and factor 2, relating to learning about medication, had a Cronbach’s alpha of 0.68.

Internal consistency reliability of the SEAMS was measured using Cronbach’s alpha, item-total correlation coefficient for scale items, and the inter-item correlation matrix. Test–retest reliability was also used to assess the SEAMS, and was determined through a baseline and 3-month follow-up administration of the scale among 96 patients. Cronbach’s alpha for the overall scale was 0.90 and would not have improved with deletion of any items. Cronbach’s alpha was 0.89 in those patients identified as having low literacy (\( n = 197 \)). The item-total correlation coefficients of all 21 items ranged from 0.36 to 0.67. The mean inter-item correlation of items was 0.32, and among all scale items it ranged from 0.08 to 0.71. Test–retest reliability revealed a Spearman’s \( \rho \) of 0.62, \( p = 0.0001 \).

No reliability testing was completed on the LTMBSES. However, it should be noted that reporting of internal consistency is possible in the original articles from which the pooled information was analyzed.

**Discussion**

It is acknowledged among health care leaders that medication non-adherence is both indicative of problems in patient-centered care, and is problematic in itself as it can result in a number of negative consequences for individual patients and adds to the burden within the health care system. One malleable factor that has been proposed as an important predictor of a variety of health behaviors, including medication adherence, is medication self-efficacy. Further, because self-efficacy is context-specific by its very definition, it may also indicate one’s current level of capacity when considering treatment burden (the context). Thus, it is important to have valid and reliable instruments to measure medication self-efficacy, particularly to assess the effectiveness of interventions designed to change one’s medication self-efficacy and assess the (mis)match between capacity and treatment burden. Having valid and reliable ways to assess medication self-efficacy is particularly important given our lack of current understanding about medication self-efficacy and its role in capacity of treatment. Our proposal of a relationship
between medication self-efficacy and capacity for treatment is a novel one.

The purpose of the systematic review was to identify instruments that assess self-efficacy for medication management and describe their theoretical underpinnings as well as their psychometric properties. Our systematic review identified three studies, which developed and evaluated three different measures of self-efficacy for medication management.23,26,27

In all three articles, Bandura’s concept of self-efficacy12 was either mentioned as a founding idea in the introduction or specifically used as a guide in instrument development. Both MUSE and SEAMS used existing scales to generate items, while the LTMBSES used a phenomenological approach for item generation. Both the SEAMS and LTMBSES involved patients as well as experts in the field for editing wording of the items. Assessment of the structure of all the instruments involved an exploratory principal component factor analysis.

Generally, our review supports the notion that development across all instruments was theoretically grounded. Also, the way in which initial factor structure was assessed is adequate. However, both the SEAMS and LTMBSES used feedback from experts and patients to further refine items, an approach not used in the development of the MUSE.

Construct validity was assessed for all three scales by generally evaluating the structure of the scale and examining the relationships between scores of the scale and other variables hypothesized to be related. Findings showed that hypotheses were supported in some instances, but not in others. For example, MUSE scores were significantly positively correlated with comprehension of prescription instructions, SEAMS scores were significantly positively related to scores on the 4-item Morisky adherence scale, and LTMBSES scores were significantly positively related to adherence and medication taking compliance; however, blood pressure and glucose levels were not different based on SEAMS scores. It appears that measures of medication self-efficacy may be related to other types of medication-related constructs (ie, adherence, compliance) but there is no evidence of a relationship to distal factors such as intermediate indicators related to health conditions, which may be physiological in nature (ie, blood pressure in hypertension and glucose levels in diabetes). None of the studies assessed health outcomes directly. These findings are not surprising given self-efficacy’s predictive relationship to health behaviors, that is, self-efficacy may be a more potent proximal predictor of intention to perform a given behavior rather than performance of the actual behavior.31–34

In other words, self-efficacy more strongly relates to proximal outcomes, such as patient-reported health outcomes, rather than more distal factors, such as disease states. Further, it is highlighted that there is a complex web of variables, including self-efficacy, leading to behavior. It is evident from our review that examining how medication self-efficacy is related to other important patient-reported outcomes such as patient-centeredness, patient enablement, and perceived treatment burden, among other factors, will only provide a richer understanding about medication use. Changing the level and nature of such patient-reported outcomes may influence medication self-efficacy and subsequent health outcomes. To date it is unclear how patient-reported outcomes and medication self-efficacy both directly and indirectly influence medication management. It is clear that medication self-efficacy is only one of many critical factors that affect taking (or not taking) medications as prescribed.

In terms of reliability assessment, both the MUSE and SEAMS had evidence of adequate internal consistency as determined by Cronbach’s alpha (generally approximately \( \alpha = 0.70 \)). Only the study evaluating the SEAMS included an assessment of test–retest reliability, which was deemed reasonable. It should be noted that no assessment of reliability was completed for the LTMBSES, which is a weakness of its psychometric support.

Although our overall assessment would lead us to conclude that the SEAMS is the most appropriate measure of self-efficacy for medication management at this time given that it had the highest quality rating and was the only measure with test–retest reliability support, evidence to support the psychometric properties of the three scales identified is fair at best. The difference in quality rating scores between the SEAMS and MUSE was a result of the lack of clarity around the presence of a question stem and item scoring of the MUSE, as well as the involvement of experts and patients in development. Reliability and literacy level description were key differences between the SEAMS and LTMBSES quality scores.

It is important to note some limitations of this review. Our search was limited to articles that developed and/or validated a measure of self-efficacy for medication management. While our search likely identified initial development and validation of existing measures, it likely excluded studies that may contain validation information about the measures, yet was not the primary focus of the study. For example, studies that may have used the MUSE, SEAMS, or LTMBSES to examine the impact of an intervention on medication adherence may show how such measures change across time. Further, many
studies not focused on development or validation of measures (ie, studies examining effectiveness of interventions, describing characteristics of special populations/disease states) will report on internal consistency of all their measures for the study sample because this is good reporting practice; however, our search would exclude such studies as the focus was not primarily on developing or assessing psychometric properties of measures of self-efficacy for medication management. Given these limitations, caution should be taken regarding conclusions about the measures.

It is clear that some support for the scales to measure self-efficacy for medication management exists; however, much more validation work is necessary. Other forms of validation such as convergent and discriminant validity were not assessed and may add support to the use of one scale over another. Using the COSMIN and Francis et al’s24 checklists as a guide to provide suggestions for further validation testing, psychometric work in terms of interpretability, generalizability, and cultural sensitivity should be undertaken. Interpretability may involve assessing how missing values or incomplete surveys should be dealt with, ceiling and floor effects, and determining the minimal important change or difference. Assessing generalizability will be important in terms of demonstrating the scales’ use across various disease states and clinical settings, particularly in the current clinical context where multimorbidity and polypharmacy are very common. Although initial support for cultural sensitivity was shown for the LTMBSES as it was translated into three languages, more work is needed in this area of psychometric evaluation. Finally, we suggest that assessing responsiveness to change for each scale is a pressing need, as self-efficacy is considered a malleable construct and, theoretically, should change as sources of self-efficacy or outcomes are targeted by an intervention. This is likely an important next step in validation work in this area.

Acknowledgments

The authors would like to acknowledge Jennifer Lawson of the McMaster University Sentinel and Information Collaboration, McMaster Department of Family Medicine, for her help refining our search strategy. The authors would also like to acknowledge the Team Approach to Polypharmacy Evaluation and Reduction team for their support of this review.

Disclosure

The authors report no conflicts of interest in this work.

References

5. May C, Montori VM, Mair FS. We need minimally disruptive medicine. BMJ. 2009;339:b2803.


