Review Article

Developing evidenced-based quality assessment checklist for real practice in primary health care using standardized patients: a systematic review

Jianjian Wang[1#], Shuaixiang Zhao[1,2#], Dong Roman Xu[3,4], Janne Estill[5,6], Meng Lv[1], Siya Zhao[1], Ming Zhang[1], Yiyuan Cai[7]^, Jing Liao[8,9], Yun Lu[10], Ruixin Wang[11], Xiaohui Wang[1], Yaolong Chen[1,12,13,14]

[1]School of Public Health, Lanzhou University, Lanzhou, China; [2]Jiaozuo People's Hospital, Jiaozuo, China; [3]ACACIA Labs, Institute for Global Health and Dermatology Hospital, Southern Medical University, Guangzhou, China; [4]Center for WHO Studies and Institute for Health Management, School of Health Management of Southern Medical University, Guangzhou, China; [5]Institute of Global Health, University of Geneva, Geneva, Switzerland; [6]Institute of Mathematical Statistics and Actuarial Science, University of Bern, Bern, Switzerland; [7]Department of Epidemiology and Medical Statistics, School of Public Health, Guizhou Medical University, Guiyang, China; [8]Department of Medical Statistics & Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, China; [9]Global Health Institute (SGHI) , Institute of State Governance, Sun Yat-sen University, Guangzhou, China; [10]School of Public Health, the Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Medical University, Guiyang, China; [11]Department of Health Economics, School of Public Health, Fudan University, Shanghai, China; [12]Lanzhou University, Affiliate of the Cochrane China Network, Lanzhou, China; [13]WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou University, Lanzhou, China; [14]Guideline International Network Asia, Lanzhou University, Lanzhou, China

Contributions: (I) Conception and design: J Wang, S Zhao, X Wang, Y Chen; (II) Administrative support: Y Chen; (III) Provision of study materials or patients: J Wang, S Zhao, X Wang, Y Chen; (IV) Collection and assembly of data: J Wang, S Zhao; (V) Data analysis and interpretation: J Wang, S Zhao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xiaohui Wang; Yaolong Chen. No. 199, Donggang West Road, Chengguan District, Lanzhou 730000, China.
Email: wangxiaohui@lzu.edu.cn; chenyaolong@lzu.edu.cn.

Background: The aim of this review was to explore the quality assessment checklists development methods in previous researches using standardized patients (SPs), as well as to propose an evidence-based checklist development procedure for quality assessment of common conditions in primary health care (PHC) settings.

Methods: We conducted a systematic review of studies that described checklist development method and extracted the methodology in terms of the developer, the basis and processes. Based on that, we formulated the development procedure according to the recommendations of the *WHO Handbook for Guideline Development*.

Results: We identified a total of 13 articles, and proposed the following five key steps: (I) forming a multidisciplinary team; (II) selecting and evaluating relevant references; (III) extracting medical information and forming the basic items; (IV) clinical expert consensus on the items; and (V) pre-testing the item pool and determining final items.

Discussion: SP has been proven to be an effective method to assess performance in practice. There are still some deficiencies in the developing of case-specific checklists using SPs. To ensure the validity and reliability of checklists, the development processes need to be more standardized and procedural.

Keywords: Primary health care (PHC); standardized patient (SP); common condition; systematic review

Submitted Mar 26, 2021. Accepted for publication Jun 11, 2021.
doi: 10.21037/apm-21-712
View this article at: https://dx.doi.org/10.21037/apm-21-712

^ ORCID: 0000-0002-8845-7309.

Introduction

Primary health care (PHC) institutions act as cornerstone in national health systems. They usually play a major role in achieving universal health coverage (UHC) and many other Sustainable Development Goals (SDGs) (1). By investing in the PHC system through comprehensive reforms, China has undoubtedly made a great improvement in expending access to basic clinical care and public health services (2). However, challenges in primary care still remain, particularly in the education and qualifications of workforce (2-5), quality of care (6) and the application of scientific quality assessment tools (7). Taking hypertension treatment as an example, high-value antihypertensive medications are prescribed only rarely in Chinese PHC settings, although being affordable and recommended by guidelines (8).

Various methods have been used to assess whether health services meet acceptable levels of quality worldwide, and current methods have significant limitations as well as strengths (9). Previous researches have shown that care measured through direct observation is generally assumed to be of much higher quality in most contexts because providers are aware they are being assessed (10,11). Indirect methods to check the behaviors include medical chart review, retrospective surveys and case presentations (12). In developed countries, these indirect methods are feasible in practice. In less developed countries, however, this may be limited due to incomplete, inconsistent or even non-existent medical records, particularly at public ambulatory care facilities. Although these methods are often used, few empirical researches have shown their validity for measuring the quality of health services delivering (13). For these reasons, it is necessary to use other means of obtaining data for the evaluation of quality.

Standardized patients (SPs) are individuals who pretend to be patients in a standardized and consistent way in a formal examination setting, or in an unannounced clinical practice to assess the performance of medical workers (14,15). This approach has been used to simulate clinical encounters for decades, despite its high cost (16), and it was commonly used in medical education and examination (17). A recent review of the SPs highlights that the types of studies conducted so far including the assessment of medical workers' knowledge, medical skills and communications practice, etc. (18-27). SPs method has been proven to be an effective and reliable method to assess performance (28-31). And it usually reports the performance of medical workers on a checklist, which can be generic or case-specific (20,21,30). A checklist of items and standardized answers to questions that the providers may ask can support the conduction of high-quality SPs studies (32,33). Recently, the number of studies with SPs has constantly grown. However, gaps in the reporting of checklist development methods remain to a certain extent, including the failure to base on high-quality evidence, and lack of validation by multidisciplinary experts, etc.

As part of the ongoing ACACIA Study that will use SPs to assess the quality of primary care across 7 provinces of China (34,35), this research aims to investigate and summarize the procedure and methodology of the current SPs checklist development. Our study focuses on three main aspects: who developed the checklist, what development procedure was used, and whether the development process was evidence-based or generated through consensus procedures. And proposed a detailed method on how to standardize the development process of checklists. We present the following article in accordance with the PRISMA reporting checklist (available at https://dx.doi.org/10.21037/apm-21-712).

Methods

Search strategy

The following six electronic databases were searched: MEDLINE (via PubMed), Cochrane Library, Epistemonikos, CNKI (China National Knowledge Infrastructure), WanFang Data and CBM (China Biology Medicine disc) up to December 31, 2019. The main terms were "standardized patient" and "simulated patient" (The details of the search strategy can be found in Appendix 1). In addition, we screened the references lists of all included publications manually for further potential studies. Only studies written in English and Chinese were included.

Inclusion and exclusion criteria

We included all studies that assessed the behavior of medical staff during diagnosis and treatment, meanwhile described the checklist development information (at least include one of the three aspects about checklist developers, development basis, and development process). The following types of studies were excluded: (I) redundant studies; (II) research unrelated to SPs; and (III) research using SPs for teaching and evaluation of medical students' clinical practice ability.

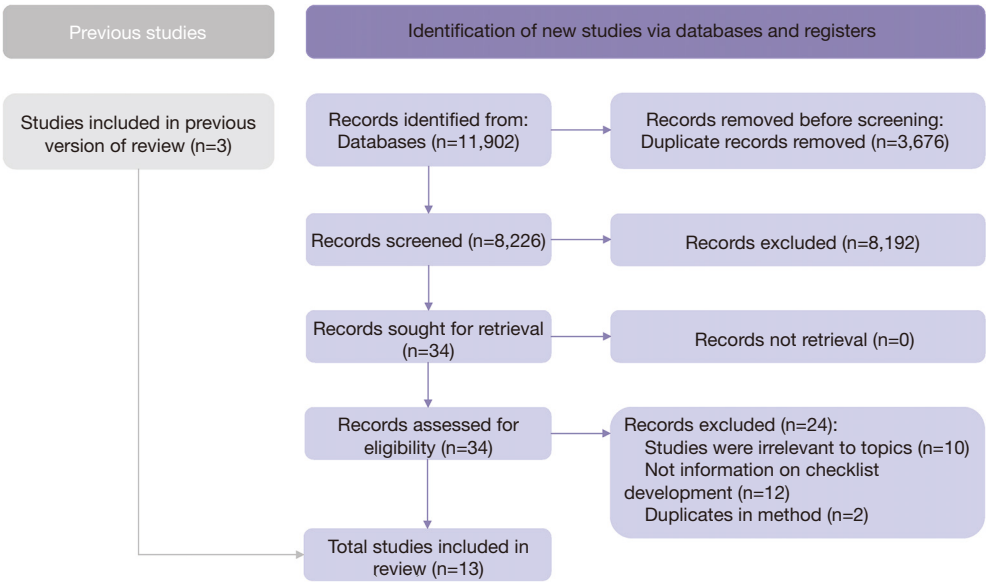

Figure 1 Study flow diagram.

Study selection and data extraction

After eliminating duplicates, two reviewers screened all titles, abstracts and full texts independently. Discrepancies were settled by discussion or consultation with a third reviewer. We developed a standardized form for data extraction including basic information on first author, publication year, research sites, target conditions, settings, checklist developers, basis and process. Information regarding the basis of the checklist development could be guidelines (36), standards (37) or others that mentioned in the publications.

Statistical analysis and checklist development methodology

By summarizing the methodological information of the checklist development, that is, we would take some procedures as key steps if more studies have adopted them. Meanwhile, according to the recommendations of the *WHO Handbook for Guideline Development* (38) and the "6S" model, which is often used to identity the best evidence for a clinical issue (39), we further verified and supplemented the initial key steps, and proposed the final key steps eventually. The "6S" pyramid model was proposed by Dicenso *et al.* in 2009, which includes six levels from top to bottom: systems (e.g., computerized decision support systems), summaries (e.g., evidenced-based clinical practice guidelines), synopses of syntheses (e.g., ACP Journal Club), syntheses (e.g., systematic reviews), synopses of studies (e.g., evidenced-based abstraction journals) and studies (e.g., original articles published in journals). Clinical evidence can be indexed sequentially until reliable and valid evidence was obtained at a certain level (39).

Results

Literature search results

The database search yielded 11,902 publications. After deduplication (n=3,676), the abstracts of 8,226 articles were reviewed and 34 potentially eligible articles were identified. Thirteen articles (20,21,30,40-49) were included eventually (*Figure 1*).

Basic information

Basic information of the included studies is presented in *Table 1*. Ten out of the 13 included studies were conducted in high income countries: four in the United States, two in Australia, and one in Germany, Canada, the Netherlands and Qatar each. Two studies were conducted in upper-middle income countries (China and South Africa). One study was from lower-middle income country (India).

The number of conditions reported in each study varied between one and four. The heath conditions covered

Table 1 Checklist develop information of included studies

Study	Country	Health conditions	Setting	Checklist developers (who)	Basis (what)	Process (how)
Burger *et al.* 2020, (21)	South Africa	Hypertension	Public healthcare facilities	Clinical experts, primary healthcare providers and program managers	Guidelines	Review
Langer *et al.* 2018, (41)	German	Acute diarrhea	Community pharmacies	The German Federal Chamber of Pharmacists	Guidelines & working aids	Not reported
Sylvia *et al.* 2015, (42)	China	Dysentery, angina	Village clinicians in rural China	Not reported	Checklist from previous studies	Adaptation
Das *et al.* 2012, (43)	India	Unstable angina, asthma, child dysentery	Public healthcare facilities	The National Rural Health Mission	Guidelines	Not reported
Glassman *et al.* 2000, (30)	United States	Lower back pain, chronic obstructive pulmonary disease, coronary artery disease, diabetes mellitus	Outpatient settings	Generalist & condition-specific specialists	Guidelines	Review
Tamblyn *et al.* 1997, (44)	Canada	Chronic hip pain due to early osteoarthritis, anti-inflammatory drugs related gastropathy	Not reported	Expert panel (eight academically affiliated physicians: representing rheumatology, geriatrics, clinical pharmacology, internal medicine, and family medicine)	Guidelines	Expert consensus
Zolezzi *et al.* 2019, (20)	Qatar	Cardiovascular disease	Community pharmacies (hospital and health center pharmacies excluded)	Pharmacy graduates	Guidelines & literature	Not reported
Gerner *et al.* 2010, (45)	Australia	Childhood overweight and mild obesity	General practices in Melbourne	Not reported	Standards	Not reported
Gordon *et al.* 1989, (46)	Australia	Urinary tract infection, tension headache, bronchitis	Casualty department	General practitioners, physicians & house staff members	Standards	Expert consensus
Rethans *et al.* 1991, (47)	Netherlands	Headache, diarrhoea, shoulder pain and diabetes	General practices in the Netherlands	Experts in general practice as well as general practitioners “in the field”	Standards	Expert consensus
Zabar *et al.* 2018, (48)	United States	Depression	Urban primary care clinics	Not reported	Standard & systemic review	Not reported
Calhoun *et al.* 1987, (40)	United States	3 chronic medical condition patients: (I) hypertension and chronic obstructive pulmonary disease exacerbated by a paralyzed diaphragm secondary to trauma; (II) diabetes mellitus, rheumatic heart disease, and weight loss; (III) lymphocytic leukemia and possible infection	Internal medicine house	Internal medical faculty members	The medical data base	Not reported
Day *et al.* 1993, (49)	United States	Lumbar pain and headaches	Outpatients in the general internal medicine clinic	Physicians & investigators	Partly based on existing tools	Discussion and validation

Table 2 Five key steps to develop treatment quality assessment checklist for studies using SPs

Step	Main content
1	Forming a multidisciplinary team
	• Checklist development group ("Group A")
	• Clinical expert consensus group ("Group B")
2	Selecting and evaluating relevant references
	• Literature reviewers
	• Databases to retrieve
	• Time frame of literature search
	• Search terms
	• Inclusion criteria
	• Literature screening
	• Quality evaluation and inclusion of final literature
3	Extracting medical information and forming the basic items
4	Clinical expert consensus on the items
5	Piloting the item pool and determining final items

SPs, standardized patients.

the circulatory system (hypertension, coronary artery disease, cardiovascular disease and angina), respiratory system (chronic obstructive pulmonary disease, asthma and bronchitis), digestive system (diarrhea, dysentery and gastropathy), endocrine system (childhood obesity and diabetes) and urinary system (urinary tract infection). Other conditions included headache, depression, lower back pain and shoulder pain. Most of the conditions were common diseases without complications, except for three complex chronic diseases in the study by Calhoun *et al.* (40) and the other four conditions in the study by Glassman *et al.* (30). The latter one developed two case scenarios—basic and complex for each condition. The most common conditions assessed in the studies were headache and diarrhea.

Case checklist development methodology

Ten of the 13 studies reported the specific checklists developer, but the number and background of the panelists varied. The majority of the experts were medical professionals, such as PHC providers, general practitioners, medical workers, and condition-specific specialists. Three studies did not report the checklist developers (*Table 1*).

Six out of the 13 studies reported having employed condition specific guidelines as the basis for diagnosis and treatment checklist development. Four studies developed the checklist according to corresponding standards. For the remaining three articles, one used a medical database as the basis of the checklist, while the other two based on the existing tools. Six studies did not report the process of how the team formulated the checklist. The rest of the studies developed the checklists through expert consensus, review, adaptation, discussion and validation, etc.

Methods and process of checklist development

Based on the results of the review, we proposed five key steps for development of quality assessment checklists in studies using SPs (*Table 2*).

Forming a multidisciplinary team

We recommend to set up a multidisciplinary team. The expert team should include methodologists and clinicians with expertise in public health and evidence-based medicine, and the number of team members can be adjusted according to specific cases. They should work as the following two working groups:

- Group A is the diagnostic and treatment checklist development group (three to four experts). The group should include one or two instructors to coordinate the project related issues, and two methodologists with expertise in evidence-based medicine who are responsible for retrieving the literature, evaluating the quality of the retrieved studies and extracting the checklist of target conditions.
- Group B, the clinical expert consensus group (six to seven experts) should include specialists in all relevant areas of the target disease, whose education, experience, region and gender also need to be considered. They are responsible for discussing and reaching consensus on the checklist items proposed by Group A.

Selecting and evaluating relevant references

We recommend that groups A and B work together to discuss and formulate a scheme for the retrieval, selection and evaluation of the literature on the target disease. The following principles are recommended for literature selection and evaluation.

Table 3 Justifications for selecting each database

Database	Reasons for selection
WanFang Data	Since 2008, the database exclusively includes all journals of the Chinese Medical Association (51,52)
Medlive	The largest clinical disease information service platform in China, which brings together the latest domestic and international guidelines and expert consensus recommendations, and provides links for the latest domestic and international guidelines and interpretations as well as translations of guidelines
MEDLINE	MEDLINE provides bibliographic indexes and abstracts of more than 4,300 major biomedical journals from more than 70 countries around the world since 1950, and also provides partially free or paid full-text links (50)
GIN NICE WHO	All three databases contain international authoritative guidelines
DynaMed UpToDate	Two authoritative databases of evidence-based medicine in the world

Literature reviewers

Two of the researchers in Group A should search the literature independently and then cross-check their results. Discrepancies can be settled by discussion or consultation with a third reviewer.

Databases to retrieve

According to the "6S" model, we recommend to search the following international databases: MEDLINE (50), Guidelines International Network (GIN), National Institute for Health and Care Excellence (NICE), World Health Organization (WHO), DynaMed, and UpToDate (*Table 3*). Since the context is to develop checklists to be used in China, WanFang Data (51,52) and Medlive are also recommended to be searched. The latest version of relevant textbooks should also be searched.

Time frame of literature search

We recommended to limit the period of literature search to the last five years to assure the timeliness of the included evidence. The time frame can be adjusted according to specific target conditions and the corresponding literatures.

Search terms

The search should include search terms relevant for the target diseases; search terms to identify guidelines, systematic reviews and diagnostic tests; as well as other relevant terms, such as language, year and type of the article.

Inclusion criteria

Inclusion criteria for the literature should contain eligible guidelines and consensus, systematic reviews, and studies on diagnostic and treatment for target diseases sequentially. Followed by the quality evaluation of included studies, as shown in the "*quality evaluation and inclusion of final literature*" below. Literature can be retrieved according to the general principles: If eligible high-quality guidelines can be retrieved, then there is no need to search consensus, systematic reviews or the lower-level literatures unless the upper-level literatures cannot answer the diagnostic and treatment problems of target diseases. Otherwise, it is suggested to continue to search for other levels of eligible studies in order until the diagnosis and treatment of the target disease can be answered.

Literature screening

Two researchers in Group A should screen the literature the following way. After removal of duplicates, they should screen all titles and abstracts independently. When studies are deemed eligible, the researchers should then obtain full texts and perform further screening. Disagreements can be settled by discussion or consultation with a third reviewer.

Quality evaluation and inclusion of final literature

Appraisal of Guidelines for REsearch & Evaluation II (AGREE II) (53), A MeaSorement Tool to Assess systematic Reviews (AMSTAR) (54) and QUality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) (55) should be used for evaluating the methodological quality of the preliminary included guidelines, systematic reviews and diagnostic tests, respectively. Although some studies may meet the initial "inclusion criteria", if they are found to be of low-quality or there are large defects, they should be eventually excluded and continue to search for other high-quality studies that meet the requirements.

Extracting medical information and forming the basic items

Two researchers in Group A should extract the information of the included literature independently using a pre-defined information extraction form, which includes details related to diagnosis and treatment of the target disease. After the summary, a preliminary pool of items for the diagnosis and treatment checklist should be formed.

Clinical expert consensus on the items

We recommend to conduct two to three rounds of group discussion and expert consultation to achieve consensus among clinical experts on the items of target disease diagnosis

and treatment checklist. The process can be made more efficient by using electronic questionnaires and telephones or internet communication (e.g., e-mail, WeChat).

Pre-testing the item pool and determining final items

After gathering all expert opinions, the researchers can select two to three community medical institutions for pre-testing the checklist. During the pre-test, investigators need to make detailed records on facts such as whether the item was needed to be explained, how it was explained, whether the clinician could understand Mandarin, whether the item was understood ambiguously, and whether and how many times the item was needed to be repeated. According to the results of the pre-test, the researchers can further refine the items and then form the final checklist.

Discussion

SPs have long been used in assessing the competency of medical students/workers and in evaluating the performance of healthcare providers' real practice (56-58). However, gaps in the developing of checklists for SPs examinations of medical workers' performance still remain. We found 13 articles that met our selection criteria, and our systematic review investigated and summarized the procedure and methodology of the current SPs checklist development. Based on the recommendations of the *WHO Handbook for Guideline Development* (38), we proposed five key steps for the checklist development for the diagnosis and treatment of common conditions in PHC settings.

SPs can be used as a needs-assessment tool, or a summative method to assess the performances of interns, residents, fellows, or specialists (59). When conducting a SPs project, checklist development is the most important step, evidence-based high-quality checklist is the prerequisite for writing a case script and setting scoring criteria (60). Our review found that, while some researches had based the contents of checklists on literature data, it was not clear whether a systematic literature search had been conducted (20,21,40-43,46). When checklists are based on data from the literature, we recommend to use evidence-based data. Unfortunately, only part of the included studies described this development process, most of them did not mention it in details. In the second step proposed in this study, we have pointed out that we need to search relevant references systematically, and use quality evaluation tools to evaluate the risk of bias of the included studies accordingly, e.g., AGREE II (53) and AMSTAR (54) could be used for evaluating the methodological quality of the included guidelines and systematic reviews respectively. Only on the basis of evidence of low-bias-risk, can we get a high-quality checklist. Standardization is not only important in SPs studies, but also in real practice. Standardized diagnosis and treatment checklists can help to evaluate unnecessary and inappropriate diagnoses and treatment, and promote the effective use of primary medical resources and optimize the time and costs of diagnosis and treatment.

Further, in the five-step methods we proposed, the information on diagnosis and treatment based on high-quality research should be extracted first, and the final checklist items should be formed through a discussion of a multidisciplinary expert group. *WHO handbook for guideline development* (38) and previous studies about the development of guidelines suggested that the guideline development working group should be multidisciplinary and representative (61). In multidisciplinary aspects: members must include clinical experts and guideline methodologists, as well as experts in relevant fields such as epidemiology, health economics, ethics, law and other related fields according to different guidelines. In representational aspects: gender, geography, institution and qualifications of the members should be taken into account (61). Due to the complexity of clinical practice, we recommend to involve multidisciplinary focus group discussion method and Delphi method to conduct clinical expert consensus of the target conditions.

This is the first article about the methods to develop quality assessment checklists for commonly-occurring conditions in the context of PHC. However, checklist validation and weighted scoring method in practice was out the scope of this study. Future researches need to explore whether and how to weight the items of the checklist. Although this research aims to develop quality assessment checklists for real practice in PHC using SPs. This study also has some limitations. Due to the few SP-related keywords in indexing, some articles may not be identified. Another limitation of our study was that it only included articles published in English and Chinese, which could introduce publication bias.

Actually, our checklist development method can also be applied to other application areas of SPs. For example, when using SPs to measure competence in medical education, our proposed five-steps method could also be used. Through the synthesis of high-quality evidence and consensus of relevant professionals, a more scientific evaluation checklist can be developed to a certain extent. Nevertheless, the checklist may be influenced by various

ways in practice, which are not only related to the technical aspects of diagnosis and treatment, but also involve social, economic, legal and other factors (62). Therefore, in order to establish more standardized quality assessment checklists in SPs studies, we should mobilize all forces of the research community and adopt a more comprehensive approach to ensure the practicality and accessibility of checklists.

Conclusions

In conclusion, SP has been proven to be an effective method to assess performance in practice. Gaps in the developing of case-specific checklists for SPs examinations of medical workers' performance still remain, including the failure to conduct systematic literature searching and evaluation, and lack of a multidisciplinary expert group for consensus and so on. To ensure the validity and reliability of medical workers' performances that use SPs, the development processes for the checklists need to be more standardized and procedural. Only in this way can researchers use the SPs to truly and reliably assess the quality of primary medical services, thereby providing decision-making basis for health policy makers and promoting the rational use of health services.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at https://dx.doi.org/10.21037/apm-21-712

Peer Review File: Available at https://dx.doi.org/10.21037/apm-21-712

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/apm-21-712). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Pettigrew LM, De Maeseneer J, Anderson MI, et al. Primary health care and the Sustainable Development Goals. Lancet 2015;386:2119-21.
2. Li X, Lu JP, Hu S, et al. The primary health-care system in China. Lancet 2017;390:2584-94.
3. Jiang GY, Kong FZ, Huang XT. Common diseases in primary health care units in rural areas of North China: status quo and countermeasures. Chinese Rural Healthy Service Administration 2014;34:774-7.
4. Cao YC. Study on the Health Service Quality Improvement Strategy for Hepan Township Health Hospital [dissertation]. Lanzhou University; 2011.
5. Das J, Woskie L, Rajbhandari R, et al. Rethinking assumptions about delivery of healthcare: implications for universal health coverage. BMJ 2018;361:k1716.
6. Su M, Zhang QL, Lu JP, et al. Protocol for a nationwide survey of primary health care in China: the China PEACE (Patient-centered Evaluative Assessment of Cardiac Events) MPP (Million Persons Project) Primary Health Care Survey. BMJ Open 2017;7:e016195.
7. Yip WC, Hsiao WC, Chen W, et al. Early appraisal of China's huge and complex health-care reforms. Lancet 2012;379:833-42.
8. Su M, Zhang QL, Bai XK, et al. Availability, cost, and prescription patterns of antihypertensive medications in primary health care in China: a nationwide cross-sectional survey. Lancet 2017;390:2559-68.
9. World Health Organization. Strengths and weaknesses of available methods for assessing the nature and scale of harm caused by the health system: literature review [Internet]. World Health Organization; 2020 [cited 2020 Jun 30]. Available online: https://www.who.int/patientsafety/research/P_Michel_Report_Final_version.pdf
10. Hermida J, Nicholas DD, Blumenfeld SN. Comparative validity of three methods for assessment of the quality of primary health care. Int J Qual Health Care 1999;11:429-33.
11. Simmons R, Elias C. The study of client-provider interactions: a review of methodological issues. Stud Fam

Plann 1994;25:1-17.
12. Siminoff LA, Rogers HL, Waller AC, et al. The advantages and challenges of unannounced standardized patient methodology to assess healthcare communication. Patient Educ Couns 2011;82:318-24.
13. Rocha KB, Rodríguez-Sanz M, Pasarín MI, et al. Assessment of primary care in health surveys: a population perspective. Eur J Public Health 2012;22:14-9.
14. Hoppe RB, Farquhar LJ, Henry R, et al. Residents' attitudes towards and skills in counseling: using undetected standardized patients. J Gen Intern Med 1990;5:415-20.
15. Rethans JJ, Saebu L. Do general practitioners act consistently in real practice when they meet the same patient twice? Examination of intradoctor variation using standardised (simulated) patients. BMJ 1997;314:1170-3.
16. Levine AI, Swartz MH. Standardized patients: the "other" simulation. J Crit Care 2008;23:179-84.
17. Epstein RM, Hundert EM. Defining and assessing professional competence. JAMA 2002;287:226-35.
18. Gerzina HA, Stovsky E. Standardized Patient Assessment of Learners In Medical Simulation. Treasure Island (FL): StatPearls; 2020.
19. Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. AAMC. Acad Med 1993;68:443-51.
20. Zolezzi M, Abdallah O, Kheir N, et al. Evaluation of community pharmacists' preparedness for the provision of cardiovascular disease risk assessment and management services: A study with simulated patients. Res Social Adm Pharm 2019;15:252-9.
21. Burger R, Christian CS, Gerdtham UG, et al. Use of simulated patients to assess hypertension case management at public healthcare facilities in South Africa. J Hypertens 2020;38:362-7.
22. Harris-Hayes M, Van Dillen LR. The inter-tester reliability of physical therapists classifying low back pain problems based on the movement system impairment classification system. PM R 2009;1:117-26.
23. Strayer SM, Martindale JR, Pelletier SL, et al. Development and evaluation of an instrument for assessing brief behavioral change interventions. Patient Educ Couns 2011;83:99-105.
24. Franz CE, Epstein R, Miller KN, et al. Caught in the act? Prevalence, predictors, and consequences of physician detection of unannounced standardized patients. Health Serv Res 2006;41:2290-2302.
25. Milone JM, Burg MA, Duerson MC, et al. The effect of lecture and a standardized patient encounter on medical student rape myth acceptance and attitudes toward screening patients for a history of sexual assault. Teach Learn Med 2010;22:37-44.
26. Fallucco EM, Hanson MD, Glowinski AL. Teaching pediatric residents to assess adolescent suicide risk with a standardized patient module. Pediatrics 2010;125:953-9.
27. Henry BW, Smith TJ. Evaluation of the FOCUS (Feedback on Counseling Using Simulation) instrument for assessment of client-centered nutrition counseling behaviors. J Nutr Educ Behav 2010;42:57-62.
28. Swanson DB, van der Vleuten CP. Assessment of clinical skills with standardized patients: state of the art revisited. Teach Learn Med 2013;25:S17-25.
29. Norman GR, Tugwell P, Feightner JW. A comparison of resident performance on real and simulated patients. J Med Educ 1982;57:708-15.
30. Glassman PA, Luck J, O'Gara EM, et al. Using standardized patients to measure quality: evidence from the literature and a prospective study. Jt Comm J Qual Improv 2000;26:644-53.
31. Rethans JJ, Gorter S, Bokken L, et al. Unannounced standardised patients in real practice: a systematic literature review. Med Educ 2007;41:537-49.
32. Kwan A, Daniels B, Bergkvist S, et al. Use of standardised patients for healthcare quality research in low- and middle-income countries. BMJ Glob Health 2019;4:e001669.
33. HealthStream. Understanding the Importance of Healthcare Competency Assessment Programs [Internet]. HealthStream; 2020 [cited 2020 Jun 30]. Available online: https://www.healthstream.com/resources/blog/blog/2018/07/17/understanding-the-importance-of-healthcare-competency-assessment-programs
34. Xu DR, Hu MY, He WJ. et al. Assessing the Quality of Primary Healthcare in Seven Chinese Provinces With Unannounced Standardised Patients: Protocol of a Cross-Sectional Survey. BMJ Open 2019;9:e023997.
35. Liao J, Chen YL, Cai YY, et al. Using smartphone-based virtual patients to assess the quality of primary healthcare in rural China: protocol for a prospective multicentre study. BMJ Open 2018;8:e020943.
36. Institute of Medicine. Clinical practice guidelines we can trust. Washington DC: the National Academies Press; 2011.
37. International Renewable Energy Agency. What are Standards? [Internet]. International Renewable Energy Agency; 2020 [cited 2020 Jun 30]. Available online: https://www.irena.org/inspire/Standards/What-are-Standards
38. World Health Organization. WHO handbook for guideline development. 2nd ed. World Health

Organization; 2014.
39. Dicenso A, Bayley L, Haynes RB. Accessing pre-appraised evidence: fine-tuning the 5S model into a 6S model. Evid Based Nurs 2009;12:99-101.
40. Calhoun JG, Woolliscroft JO, ten Haken JD. Internal medicine house officers' performance as assessed by experts and standardized patients. J Med Educ 1987;62:754-60.
41. Langer B, Bull E, Burgsthaler T, et al. Assessment of counselling for acute diarrhoea in German pharmacies: a simulated patient study. Int J Pharm Pract 2018;26:310-7.
42. Sylvia S, Shi YJ, Xue H, et al. Survey using incognito standardized patients shows poor quality care in China's rural clinics. Health Policy Plan 2015;30:322-33.
43. Das J, Holla A, Das V, et al. In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. Health Aff (Millwood) 2012;31:2774-84.
44. Tamblyn R, Berkson L, Dauphinee WD, et al. Unnecessary prescribing of NSAIDs and the management of NSAID-related gastropathy in medical practice. Ann Intern Med 1997;127:429-38.
45. Gerner B, Sanci L, Cahill H, et al. Using simulated patients to develop doctors' skills in facilitating behaviour change: addressing childhood obesity. Med Educ 2010;44:706-15.
46. Gordon JJ, Saunders NA, Sanson-Fisher RW. Evaluating interns' performance using simulated patients in a casualty department. Med J Aust 1989;151:18-21.
47. Rethans JJ, Sturmans F, Drop R, et al. Assessment of the performance of general practitioners by the use of standardized (simulated) patients. Br J Gen Pract 1991;41:97-9.
48. Zabar S, Hanley K, Watsula-Morley A, et al. Using Unannounced Standardized Patients to Explore Variation in Care for Patients with Depression. J Grad Med Educ 2018;10:285-91.
49. Day RP, Hewson MG, Kindy P Jr, et al. Evaluation of resident performance in an outpatient internal medicine clinic using standardized patients. J Gen Intern Med 1993;8:193-8.
50. National Center for Biotechnology Information [Internet]. All Resources. [cited 2020 Jun 30]. Available online: https://www.ncbi.nlm.nih.gov/guide/all/#databases
51. Chinese Medical Association. The Chinese Medical Association signed an exclusive cooperation agreement with Wanfang Data to deposit a series of journal databases. Chinese Journal of Pediatrics 2009;47:82.
52. Chinese Medical Association. The Chinese Medical Association and Beijing Wanfang Data Co., Ltd. renewed the "Chinese Medical Association Series Magazine Database" exclusive cooperation agreement. Chinese Journal of Urology 2010;31:445.
53. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 2010;182:E839-42.
54. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
55. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.
56. Pieters HM, Touw-Otten FW, De Melker RA. Simulated patients in assessing consultation skills of trainees in general practice vocational training: a validity study. Med Educ 1994;28:226-33.
57. Colliver JA, Swartz MH. Assessing clinical performance with standardized patients. JAMA 1997;278:790-1.
58. Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of 3 methods for measuring quality. JAMA 2000;283:1715-22.
59. Ryall T, Judd BK, Gordon CJ. Simulation-based assessments in health professional education: a systematic review. J Multidiscip Healthc 2016;9:69-82.
60. Gorter S, Rethans JJ, Scherpbier A, et al. Developing case-specific checklists for standardized-patient-based assessments in internal medicine: A review of the literature. Acad Med 2000;75:1130-7.
61. Chen YL, Ma YF, Zhou Q, et al. Who Should Participate in the Development of Clinical Practice Guidelines? Medical Journal of Peking Union Medical College Hospital 2019;(5):524-30.
62. Yang ZY, Kong LB, Yang Z, et al. The influence factors and mode establish of standardization diagnosis and treatment behavior pattern. Chinese Journal of Behavioral Medical Science 2006;15:234-6.

Cite this article as: Wang J, Zhao S, Xu DR, Estill J, Lv M, Zhao S, Zhang M, Cai Y, Liao J, Lu Y, Wang R, Wang X, Chen Y. Developing evidenced-based quality assessment checklist for real practice in primary health care using standardized patients: a systematic review. Ann Palliat Med 2021;10(7):8232-8241. doi: 10.21037/apm-21-712

Supplementary

Search strategy

MEDLINE (N=2800)

#1 "standardized patient*"[Title/Abstract]
#2 "simulated patient*"[Title/Abstract]
#3 OR/1-2

Cochrane Library (N=1326)

#1 "standardized patient*"[ti, ab, kw]
#2 "simulated patient*"[ti, ab, kw]
#3 OR/1-2

Epistemonikos (N=189)

#1 "standardized patient*"[Title/Abstract]
#2 "simulated patient*"[Title/Abstract]
#3 OR/1-2

CNKI (N=2853)

#1 " 标准化病人 "[主题]
#2 " 模拟病人 "[主题]
#3 " 标准化患者 "[主题]
#4 " 模拟患者 "[主题]
#5 OR/1-4

CBM (N=1987)

#1. " 标准化病人 "[常用字段 : 智能]
#2. " 模拟病人 "[常用字段 : 智能]
#3. " 标准化患者 "[常用字段 : 智能]
#4. " 模拟患者 "[常用字段 : 智能]
#5. OR/1-4

Wanfang (N=2747)

#1 " 标准化病人 "[主题]
#2 " 模拟病人 "[主题]
#3 " 标准化患者 "[主题]
#4 " 模拟患者 "[主题]
#5 OR/1-4

https://dx.doi.org/10.21037/apm-21-712