Clinical measurement properties of malnutrition assessment tools for use with patients in hospitals: a systematic review

Yue Camille Xu 1,2* and Joshua I. Vincent 1

Abstract

Background: The use of malnutrition outcome measures (OM) by registered dietitians (RD) with inpatients in hospitals has increased promoting the achievement of nutritional care goals and supporting decision-making for the allocation of nutritional care resources in hospitals. There are 3 commonly used OMs: Subjective Global Assessment (SGA), Patient Generated-Subjective Global Assessment (PG-SGA) and Mini Nutritional Assessment (MNA). The purpose of this current study was to systematically review the evidence of the clinical measurement properties of malnutrition assessment tools for use with patients admitted in hospitals.

Methods: MEDLINE, Cinahl, EMBASE, and PubMed were searched for articles published between 2000 and 2019. Research articles were selected if they established reliability, validity, and responsiveness to change properties of the SGA, PG-SGA and MNA tools, were written in English, and used any of these OMs as an outcome measure. Abstracts were not considered. The risk of bias within studies was assessed using the Quality Appraisal for Clinical Measurement Study (QA-CMS).

Results: Five hundred five studies were identified, of which 34 articles were included in the final review: SGA (n = 8), PG-SGA (n = 13), and MNA (n = 13). Of the 34 studies, 8 had a quality score greater than 75%; 23 had a quality score of 40–75% and 3 studies had a quality score of less than 40%. PG-SGA was found to have excellent diagnostic accuracy (ROC: 0.92–0.975; Sensitivity: 88.6–98%; Specificity: 82–100%), sufficient internal consistency (Cronbach’s alpha: 0.722–0.73), and strong test-retest reliability (r = 0.866). There was insufficient evidence to suggest adequate diagnostic accuracy and good inter-rater reliability for SGA. Only one study examined the minimum detectable change of MNA (MDC = 2.1).

Conclusions: The evidence of validity for the existing malnutrition assessment tools supports the use of these tools, but more studies with sound methodological quality are needed to assess the responsiveness of these OMs to detect the change in nutritional status.

Keywords: Malnutrition assessment tool, Reliability, Validity, Responsiveness, Outcome measure, Systematic review

© The Author(s). 2020 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
**Introduction**

According to the World Health Organization [1], malnutrition can be caused by many different factors including under-nutrition, over-nutrition, inadequate nutrient intake, and an unhealthy diet, resulting in chronic illnesses such as diabetes, stroke, and hypertension. A recently published prospective cohort study conducted in 18 Canadian hospitals from July 2010 to February 2013 found that 45% of patients were malnourished at admission [2]. Malnutrition in inpatients is associated with adverse health outcomes, such as the development of infectious diseases, respiratory failure, and pressure ulcers [3, 4]. Moreover, the impact of malnutrition on health outcomes for patients with stroke can be significant and increase mortality and delay functional recovery [5, 6]. Patients that developed malnutrition during hospitalization required longer hospital stays, could not independently perform daily activities, and became high-cost care users at discharge [7, 8].

**What are malnutrition outcome measures?**

The use of outcome measures in the health sector promotes the achievement of care goals, facilitates patient-healthcare professional communication, matches the delivery of health care to the patient’s needs, and supports the decision-making for the allocation of healthcare resources [9]. A well-developed outcome measure should consist of three primary psychometric properties: reliability, validity, and responsiveness to change [10, 11] (Table 1). In the context of malnutrition management, there are two broad categories of outcome measures used in tertiary care facilities: malnutrition screening and malnutrition assessment. Malnutrition screening is a quick and simple process in which the screening can be readily performed by nursing staff to identify patients at risk of malnutrition and inform practitioners if further clinical nutritional interventions are warranted [19]. The screening process at hospital admission is a crucial step to improve safe patient care; moreover, using a validated screening tool triggers appropriate referrals to dietitians that can assess and treat malnutrition in a timely manner to reduce overspending of resources from preventable misdiagnosis and poor patient outcomes [20].

Malnutrition assessment is different from malnutrition screening in that an in-depth and comprehensive evaluation of nutritional status is performed; therefore, professional training is required to conduct malnutrition assessment and this process is usually completed by a trained registered dietitian (RD) [21]. Moreover, in the field of nutritional practice, the assessment tool should

| Clinical measurement property | Rating method | Definition [10, 12, 13] and/or Cut-off value [10, 14, 15] |
|-----------------------------|--------------|--------------------------------------------------------|
| Inter-rater reliability     | A measure of the consistency for the scores obtained between two raters that are measuring the same subject. |
| Test-retest reliability     | A property of which the stability of the result using the same outcome measure with the same group of subjects in a repeated test. |
| Internal consistency        | A reliability measure; ensuring all measurement items measure the same context. |
| Kappa\(^a\)                 | A preferred statistical measure for rating the inter-rate reliability; cut-off value: \(≥ 0.70\) – acceptable [10, 16, 17]. |
| Cronbach’s \(\alpha\)\(^b\) | A statistical measure for rating the internal consistency; cut-off value: \(≥ 0.70\) – acceptable [10, 18]. |
| ICC\(^c\)                   | A statistical measure for rating the reliability; cut-off value: ≥ 0.70 – acceptable [10]. |
| Concurrent validity         | A type of criterion validity; examining the consistency of the score obtained by the testing instrument compared with a gold standard. |
| Predictive validity         | A type of criterion validity; measuring a correlation between the testing score and future events. |
| Construct validity          | An essential measurement validity which assesses whether a testing instrument measures what it was made to measure by comparing hypothetical and non-criterion parameters. |
| Diagnostic accuracy         | An ability to discriminate the population with targeted health condition. |
| ROC\(^d\)                   | A statistical measure for expressing how well the measure can distinguish between two populations; cut-off values: > 0.75 – excellent; 0.50 to 0.75 – moderate; 0.25 to 0.49 – mild; < 0.25 – weak [14]. |
| Responsiveness to change    | An ability to detect the changes in health outcomes following the intervention. |
| MDC\(^e\)                  | A statistical measure for rating the response to changes which is not due to the measurement error. |

\(^a\)Kappa = Cohen’s Kappa; \(^b\)Cronbach’s \(\alpha\) = Cronbach’s alpha; \(^c\)ICC = Intraclass Correlation Coefficient; \(^d\)r = Pearson’s correlation coefficient; \(^e\)ROC = Receiver Operator Cure; \(^f\)MDC = minimal detectable change

Note: \(\alpha\) = Cronbach’s alpha; ICC = Intraclass Correlation Coefficient; \(r\) = Pearson’s correlation coefficient; ROC = Receiver Operator Cure; MDC = minimal detectable change
not only be used to diagnose malnutrition at the initial visit, but the same tool should also be used by RDs to compare the effect of nutritional intervention and to measure nutritional outcomes at re-assessment. There are three well-studied malnutrition assessment tools available for this purpose: the Mini Nutritional Assessment (MNA) [22, 23], the Subjective Global Assessment (SGA) [24] and the Patient-Generated Subjective Global Assessment (PG-SGA) [25].

**Mini nutritional assessment (MNA)**
This tool was originally developed in 1990 to assess the nutritional status of elderly patients [22]. The full form of MNA consists of 18 scored questions that are divided into 4 categories: 1) anthropometric measurements; 2) global assessment; 3) dietary history and 4) metabolic stress [22]. The MNA generates a total score of 30. The total scores are interpreted as follows: 24–30 (normal nutritional status); 17–23.5 (at risk of malnutrition); less than 17 points (malnourished) [22].

**Subjective global assessment (SGA)**
This instrument was originally developed by Detsky et al. in 1987 to predict malnutrition in patients undergoing gastrointestinal surgery [24]. The SGA consists of two assessment features: 1) history and 2) physical examination relevant to malnutrition status [24]. The history part includes patterns of weight change, dietary history, gastrointestinal signs and symptoms, physical functionality and underlying inflammatory disease [24]. Additionally, three categories of physical assessment in relation to malnutrition are used in SGA: loss of subcutaneous fat, muscle wasting and the presence of fluid retention [24]. After the assessment, the patient is classified as either SGA-A (well-nourished), SGA-B (moderately malnourished), or SGA-C (severely malnourished) [24].

**PatientGenerated-subjective global assessment (PG-SGA)**
The PG-SGA was originally developed as an extension of the SGA tool to assess the nutritional status of patients with cancer [25]. The tool includes all SGA components and involves patients to self-report their nutritional histories [25]. Additionally, there are two new features in this instrument: first, for each item of the PG-SGA, a score of 0–4 is added, the more severe the symptoms in relation to malnutrition the higher the assigned value. Second, PG-SGA can generate a numerical score in addition to summarizing a global rating of A (well-nourished), B (moderately malnourished), and C (severely malnourished). A total score between 0 and 35 quantitatively informs the severity of malnutrition and types of intervention needed: 0–1 point indicates no need for any intervention; 2–3 points suggest education needs for the patients and family; 4–8 points indicate the need for a referral to a dietitian; and a score of 9 or more recommends an action of critical nutritional intervention [25].

**What are gaps identified in the current literatures?**
A survey distributed to 125 stroke-specific health care institutions in Canada during 2008–2009 revealed that the majority of RDs did not use validated screening tools to assess the nutritional status of patients with stroke; the author also suggested that these results can be extrapolated to dietitian practice in other patient areas [26]. Encouraging the use of outcome measures among RDs is a necessary step that will move the profession toward an outcome-based practice. However, it is important that these existing malnutrition assessment tools are adequately examined for their validity prior to recommending their use in a hospital setting. Many recently published systematic reviews either invested their interests on the validation of the malnutrition screening tools [19], which is different from the malnutrition assessment tools, or did not include all three malnutrition assessment tools in their reviews [27, 28]. One systematic review study with meta-analysis assessed the validity of using SGA, PG-SGA, and MNA in the community [21]. This study applied some comprehensive searching and study appraisal strategies; however, the authors did not appraise the appropriateness of the statistical methodologies used to express the criterion validity, and not all psychometric properties were included for review in this article [21]. Therefore, the objectives of this paper are to systematically review the literature available on the clinical measurement properties of three malnutrition assessment tools, SGA, PG-SGA, and MNA, used with patients in hospitals and summarize the advantages and limitations of each assessment tool.

**Methodology**
PRISMA [29, 30] checklist and flow diagram were used as a reporting guideline (Additional file 1).

**Literature search**
Systematic searches were performed between the months of August 2019 and November 2019 in 4 databases: PubMed, CINAHL, ProQuest, and MEDLINE (via Ovid). Hand-searching of the key journals: Clinical Nutrition Journal, Nutrition Journal, Journal of Parenteral and Enteral Nutrition, and Nutrition in Clinical Practice were also completed. The reference lists of relevant articles were also searched to ensure a comprehensive search. The search terms were developed by following three steps. First, the author (YX) reviewed Medical Subject Heading (MeSH) terms associated with malnutrition, health care institution, and the nutritional
assessment tools and ensured that these search terms were relevant to the context of this research study. Second, in consultation with another investigator (JV), essential terms of psychometric properties were added to maximize the specificity of the literature search. Third, in the review of other systematic reviews, additional context-specific terms in malnutrition assessment were considered [19, 21, 31]. Therefore, the key search terms applied in this paper included “SGA”, “subjective global assessment”, “PG-SGA”, “patient-generated subjective global assessment”, “MNA”, “mini nutrition assessment”, “nutritional assessment”, “nutrition outcome”, “malnutrition”, “protein-energy malnutrition”, “undernutrition”, “hospital”, “rehabilitation”, “subacute care”, “reliability, “validity”, and “responsiveness to change”. These terms were combined using the Boolean operator OR and AND in the search. We kept the clinical measurement terms broader to obtain maximal yield.

Study selection
The study selection in this paper was performed by YX and JV following two steps. First, the titles and abstracts of articles were screened to identify potentially eligible studies. Second, the full articles were retrieved and carefully reviewed to meet the study eligibility criteria.

Studies were included if they met the following including criteria: Study published in English after 2000 and conducted in tertiary care facilities, including inpatient, outpatient, sub-acute care, and rehabilitation; participants included in this study review were adults over 18 years old; studies should have assessed one or more clinical measurement properties of the studied outcome measures. Studies were excluded from this review if they were completed in primary care, community, long-term care, and intensive care units set-up, which require a nutritional assessment that is different from general practice within a hospital; if malnutrition screening tools were used (e.g. the short form of MNA, which is a screening tool, not a comprehensive assessment tool) [22]; if modified versions of the outcome measures were used (e.g. 7-point SGA, Taiwanese-specific version of MNA-T1 and T2, Thai-version PG-SGA); and if they were systematic reviews.

Data extraction
Data collection included the collection of the key characteristics of the literature, such as population, settings, and the location where each study was conducted. Moreover, other relevant variables, including psychometric properties examined, statistical methodologies used, and the main findings of the validation result for each article are summarized in the Additional file 2.

Risk of bias in individual studies
The Quality Appraisal for Clinical Measurement Study (QA-CMS) was chosen to evaluate the internal validity of individual studies [12]. The QA-CMS consists of 12 items spread across 5 categories: study question, study design, measurement description, statistical analyses used, and study recommendation (Additional file 3). The quality score for each item is on a scale of 0–2, giving a total score out of 24, which is converted into a percentage [12]. Two reviewers (YX and JV), who were blinded to each other’s evaluation, independently performed the study appraisal in this review. An initial calibration review was completed, in which both the reviewers independently reviewed at least 2 articles. Then each item was discussed to clarify the meaning and interpretation of the items on the QA-CMS. After completion of the independent critical appraisal, both reviewers discussed each specific item on the QA-CMS for all the articles to obtain consensus (Additional file 4).

Summary of statistical measures use
The cut-off points of each measure used for determining the adequacy of validation results in this review are summarized in Table 1.

Results
Search results and study characteristics
Five hundred and five studies were originally identified in the literature search: 102 in CINAHL, 89 in MEDLINE via Ovid, 153 in ProQuest, and 161 in PubMed. Additionally, 18 articles were added from the hand-searching of key journals and the reference list of identified systematic reviews. One hundred and fifty-nine articles were removed as duplicates and 214 papers were excluded after abstract screening. Following the full article review, 34 literature articles remained eligible, of which 8 investigated SGA, 13 investigated PG-SGA, and 13 investigated MNA. A flowchart of the study selection process is presented in Fig. 1.

Studies took place in Australia (n = 11), Turkey (n = 5), Japan (n = 3), Iran (n = 2), Korea (n = 2), Malaysia (n = 2), USA (n = 1), UK (n = 1), India (n = 1), Norway (n = 1), Taiwan (n = 1), Poland (n = 1), Greece (n = 1), Libya (n = 1), and Mexico (n = 1). The mean age of study participants was 63.26 years ranging between 19 to 92 years. The types of study designs identified in this review were prospective cohort studies (n = 32), a retrospective study (n = 1), and a randomized control trial study (n = 1). The psychometric properties examined in these 34 articles included test–retest reliability (n = 3), inter-rater reliability (n = 5), internal consistency (n = 5), predictive validity (n = 6), construct validity (n = 22), diagnostic accuracy (n = 7), and responsiveness to change (n = 3). The data
extraction including pertinent study characteristics and result findings are presented in the Additional file 2.

Risk of bias within studies

When discrepancies of scoring of specific items existed, the two reviewers (YX and JV) revisited and discussed the full article to resolve the score by mutual agreement. Of the 34 articles evaluated using QA-CMS [12], only 5 (14.70%) papers described a thorough literature review of the studied tools in their introduction, including the currently known measurement properties and the gaps identified from the current literature review that resulted in the development of relevant research questions, and calculations to determine the optimum numbers of study subjects to participate in the study were performed in 6 papers (17%). However, most papers (28 out of the 34 studies) included an adequate description of the studied instruments, scoring, and statistical methodologies applied to examine the clinical measurement properties. Thirteen articles evaluated two or more psychometric properties, of which Lin et al. [32], Ghazi et al. [33], Soy- sal et al. [34], and Bauer et al. [35] provided a comprehensive review of the studied tool by concurrently examining three or more indicators of reliability and validity [12]. The overall administrative procedure to perform the outcome measure in an unbiased way was evaluated primarily based on 1) if a professionally trained dietitian or experienced clinician was hired to perform patient assessments, 2) if a standardized testing procedure was followed to perform anthropometric measures, 3) if the timing when these measures were performed was consistent for each study participant, and 4) if the time-interval to assess test-retest reliability or responsiveness to change was appropriate [12].

All final scores for each eligible study are summarized in the Additional file 4. Moreover, a cut-off value of 60% was arbitrarily determined by the two authors to identify papers with an acceptable level of quality that examined...
the psychometric properties of the studied tools in this review.

**Synthesis of measurement properties for MNA, SGA, and PG-SGA**

A summary of clinical measurement properties extracted from 18 eligible articles that obtained scores of 60% and higher based on QA-CMS are presented in Table 2.

**Discussion**

A comprehensive systematic review was performed to validate the clinical measurement properties of three malnutrition assessment tools: SGA, PG-SGA, and MNA. A broad range of clinical measurement properties was studied in this systematic review, including internal consistency, inter-rater reliability, test-retest reliability, construct validity, criterion validity, and responsiveness to change. Furthermore, studies conducted in diverse ethnicity and patient population groups were identified.

Approximately 30 years ago, the MNA tool was developed in recognition of the high prevalence and specificity of malnutrition among institutionalized geriatric patients [22, 23]. Nowadays, MNA is well known by health care providers for evaluating the adequacy of nutritional status in elderly patients [59], and its reliability is confirmed in this review by an acceptable internal consistency (Cronbach’s alpha coefficient: 0.70 to 0.70) [33, 34] and test-retest reliability (ICC = 0.91, 95% CI, 0.85 to 0.94) [32]. Moreover, MNA was strongly correlated with PG-SGA for the assessment of hospitalized patients affected by stroke in Australia [48], and MNA correlated with objective measures including anthropometric and laboratory tests for the assessment of malnutrition among patients with end-stage renal failure, liver disease, gerontological conditions, and stroke in European and Asian populations [44, 47, 50–52]. Moreover, a lower MNA score predicted a longer duration of dialysis for patients with end-stage renal disease [44]. The most interesting finding about the MNA tool was its responsiveness to change (MDC = 2.1) which indicates a higher accuracy to assist with clinician’s clinical judgment amid nutritional re-evaluation of patient outcome [32]. Finally, as a diagnostic tool, MNA was also found to have a good to excellent diagnostic accuracy (ROC: 0.71 to 0.90; Sensitivity: 58 to 83.30%; Specificity: 74.40 to 91.20%) in Asian, Australian, and European patient populations [33, 34, 36].

There are many advantages to using MNA in clinical practice. For example, the tool is easy to use and a full assessment can be completed within 10 min [53]; moreover, the tool is accessible online at no cost. The study by Guigoz et al. [54] also indicated acceptable results of this MNA tool for use among 30,000 elderly patients in various health care facilities and different countries. Furthermore, this review revealed that many recent studies have expanded their interests to validate the use of the MNA tool for patients other than geriatrics; this has expanded its popularity for use in patients with complex medical needs. However, the challenge exists in the applicability of Body Mass Index (BMI) measurements for patients in tertiary-care settings [53]. For example, there are disagreements in BMI thresholds based on different age-group; body weight is sensitive to the change in fluid status, which is commonly seen in inpatients; cancer tumors can also significantly increase the body mass, thus creating bias of BMI interpretation [55]. On the other hand, MNA contains questions to self-evaluate nutritional and health status, and this may reduce its applicability in patients with declined cognition or impaired speech capacity [53]. Finally, questions that address food choices, portion size, and the mode of feeding may not be appropriate for patients who are nutritionally stable but are receiving enteral feeding as an alternate route of food intake [56].

Among the three malnutrition assessment tools studied in this review, SGA was the first tool developed and validated for use in healthcare. Moreover, it was recommended as an acceptable assessment tool by the European Society for Clinical Nutrition and Metabolism (ESPEN) [57]. Compared with MNA and PG-SGA, the portion of patient-reported items is lower in SGA. However, multifaceted anthropometric measures are added into the assessment to examine the muscle mass loss, subcutaneous fat loss, and signs of fluid overload in 11 areas of body parts; these features of SGA were reported to improve the accuracy of malnutrition assessment [23, 58], and the tool has been validated for use in a variety of patient populations [59]. However, in the current review, SGA was not found to have sufficient sensitivity for the identification of severe malnutrition among patients with renal failure; one explanation about this conflicting finding could be that fluid overload as a result of end-stage renal disease may mask the sign of subcutaneous fat loss, which can interfere with clinicians’ subjective judgments on nutritional status [37].

One commonly reported limitation for the clinical application of SGA is its accuracy in relation to the rater’s experiences [60, 61]. Steenson et al. [62] found that dietitians with more than 5 years of clinical experience generated the highest inter-rater reliability with the benchmark compared with other groups of dietitians who had fewer years of working experiences after graduation. Therefore, inter-rater reliability is a very pertinent psychometric property to validate SGA in clinical measurement studies. Interestingly, in this review, the agreement between trained dietitians was acceptable (Cohen kappa = 0.96, p < 0.001) [63]; however, an inadequate agreement was found between a renal dietitian and a
Table 2 Summary of clinical measurement properties of MNA, SGA and PG-SGA tools

| Measurement property | Patient population applied and measure extracted | MNA | SGA | PG-SGA |
|----------------------|-------------------------------------------------|-----|-----|--------|
| Diagnostic accuracy  | Geriatric patients [36]:                         |     |     |        |
|                      | End-stage renal failure [37]:                    |     |     |        |
|                      | Parkinson’s disease [33]:                       |     |     |        |
|                      | Pre-frail elderly [34]:                         |     |     |        |
|                      | Frail elderly [34]:                             |     |     |        |
|                      | Test-retest reliability                         |     |     |        |
|                      | Internal consistency                            |     |     |        |
|                      | Inter-rater reliability                         |     |     |        |
|                      | Construct validity                              |     |     |        |

**Geriatric patients [36]:**
- ROC = 0.85 ± 0.05 (95% CI 0.75 to 0.96, p < 0.0001); sensitivity: 83.30%; specificity: 74.40%

**End-stage renal failure [37]:**
- SGA-C can discriminate patients who are not malnourished (specificity: 0.93-0.94), but cannot accurately identify patients with severe malnutrition (sensitivity: 0.05-0.14)

**Parkinson’s disease [33]:**
- Cut-off score = 22.5
- ROC = 0.7 (p < 0.001); sensitivity: 58%; specificity: 82%

**Pre-frail elderly [34]:**
- Cut-off score = 25.5
- ROC = 0.83 (95% CI 0.79 to 0.87); sensitivity: 72.10%; specificity: 91.20%

**Frail elderly [34]:**
- Cut-off score = 25.5
- ROC = 0.90 (95% CI 0.88 to 0.93); sensitivity: 72.10%; specificity: 91.20%

**Test-retest reliability**
- Stroke [32]: None
- Parkinson’s disease [33]: None
- General medicine [42]:
  - Patients rated by two dietitians (Cohen’s kappa = 0.96, p < 0.001) versus dietitian and a trained allied health assistance (Cohen kappa = 0.84, p < 0.001)

**Internal consistency**
- Parkinson’s disease [33]: Cronbach’s α = 0.70 (95% CI, 0.62 to 0.77)

**Inter-rater reliability**
- None

**Construct validity**
- Dialysis [44]:
  - Moderately correlated with a female’s body weight (r = 0.64) and BMI (r = 0.62)
- Digestive disease [45]:
  - Weak correlation with BMI (r = 0.36), percentage of arm muscle circumference (r = 0.33), percentage of triceps skin fold thickness (r = 0.26), albumin (r = 0.41), total cholesterol (r = 0.16), and lymphocyte count (r = 0.24).
- Gynecological cancer [39]:
  - ROC = 0.92 (95% CI, 0.84 to 1.01, p < 0.001)

**Pre-frail elderly [34]:**
- Cut-off score = 25.5
- ROC = 0.83 (95% CI 0.79 to 0.87); sensitivity: 72.10%; specificity: 91.20%

- ROC = 0.90; sensitivity: 88.60%; specificity: 100%

**Cancer inpatient [35]:**
- Cut-off score = 9
- Sensitivity: 98%; specificity: 82%
nephrologist (Cohen kappa = 0.60) [37]. On the other hand, the studied tool was found to correlate with anthropometric measures and biochemical measures in a variety of patient populations in Asia and Europe [45, 64]. Moreover, SGA with category B or C predicts longer hospital stays [45] and mortality [65].

The advantage of SGA is its clinical utility. It is simple, quick, inexpensive, and has been widely accepted as a criterion to validate new tools developed for nutritional screening and assessment [19]. Most interestingly, SGA was recognized as a nutritional screening tool used to increase the Diagnosis Related Group based health care reimbursement in Europe [66] and the coding of malnutrition on a casemix-based hospital funding system in Singapore [67]. However, in addition to the need for ongoing training and practice to maintain the high accuracy of malnutrition diagnosis by SGA [62], this subjective assessment tool lacks responsiveness to change to detect the changes in nutritional status following intervention [60]. This disadvantage may limit its use in clinical practice to measure the effect of malnutrition treatment and it may reduce its selection for use as an outcome measure in future nutritional studies.

The features of PG-SGA included a shift from clinician-centered to patient-centered assessment approach and enhancing patient-clinician interaction, thus promoting collaborative decision making and streamlining care delivery [68]. Moreover, the Oncology Nutrition Dietetic Practice Group of the American Dietetic Association has recognized the PG-SGA as the recommended malnutrition assessment tool for use in a patient with cancer [69]. PG-SGA was found to have outstanding diagnostic accuracy (ROC: 0.92 to 0.98; Sensitivity: 88.60 to 98%; Specificity: 82 to 100%) for the identification of malnutrition in patients with cancer or those undergoing surgery upon admission to the hospital [35, 38, 39].

Moreover, PG-SGA is a reliable tool as evidenced by several findings that have indicated acceptable internal consistency (Cronbach’s alpha = 0.72 to 0.73) in a variety of patient populations and acceptable test-retest reliability ($r = 0.87$) with a 14-day-period reassessment [40, 41]. PG-SGA correlates with various nutritional parameters and nutritionally associated outcomes, such as global quality of life, to mostly assess patients with cancer in Australian and Asian populations [35, 41, 46, 70]. Furthermore, a higher PG-SGA score can predict longer

### Table 2 Summary of clinical measurement properties of MNA, SGA and PG-SGA tools (Continued)

| Measurement property | Patient population applied and measure extracted | MNA | SGA | PG-SGA |
|----------------------|--------------------------------------------------|-----|-----|--------|
| **Cancer [48]:**     | Highly correlated with PG-SGA at the baseline ($r = -0.76$), at the 4–6 weeks follow up ($r = -0.73$) and at the 8–12 weeks follow up ($r = -0.83$). | | |
| **Parkinson’s disease [33]:** | Mildly correlated with weight ($r = 0.43$), BMI ($r = 0.35$), mid-arm circumference ($r = 0.268$), calf circumference ($r = 0.29$) | | |
| **Stroke [32]:** | Mildly correlated with quality of life ($r = 0.32$) | | | |
| **Predictive validity** | **Dialysis [44]:** | Moderately correlated with the duration of dialysis ($r = -0.53$) | | |
| **Digestive disease [45]:** | Weak correlation with the length of hospital stays ($r = 0.29$) | | |
| **Cancer inpatient [35]:** | Weak correlation with the length of stays ($r = 0.36$) | | |
| **Post-appendectomy surgery [38]:** | Weak predictive validity for the length of stays ($r = 0.38$) | | |
| **Patient with cancer had undergone radiotherapy [46]:** | a moderate correlation between the change in PG-SGA score and change in global quality of life after 4 weeks of radiotherapy ($r = -0.55$) | | |
| **Responsiveness to change and minimal detectable change (MDC)** | **Stroke [32]:** | None | | |
| | MDC = 2.1 | | | |
hospital stays for patients with stroke [71], cancer [35], and undergoing gastrointestinal surgery [38]. Isenring et al. [46] reported that a change in PG-SGA score of 9 (95% CI, 7.20 to 10.90) was required to achieve improvement or deterioration of nutritional status; however, the authors did not explain if this score was validated or a standard error of measurement (SEM) was calculated. Therefore, the sensitivity to change of PG-SGA is still unclear based on the result of this review.

In addition to the patient-involved assessment feature of the PG-SGA tool, other advantages, such as the extensive consideration of nutritionally relevant disease diagnosis, the clarity of metabolic stress contributors and physical examinations, and the scoring system of this instrument have made it stand out from the other tools. Moreover, the numerical PG-SGA points feature allows the detection of changes over time; therefore, it seems to be a favorable tool used in recent nutritional studies that assessed improvement of nutritional status following the studied interventions [68, 72]. Furthermore, PG-SGA was used by Kellett et al. to identify appropriate coding of malnutrition, which allowed the estimation of unclaimed financial reimbursement based on the Diagnosis Relate Group (DRG) hospital funding criteria in Australia [73, 74]. Marshall’s study [36] also discovered a substantial agreement between the PG-SGA and the International Statistical Classification of Diseases and Health-Related Problems, 10th revision, Australian Modification (ICD-10-AM) criteria based on which Australian hospitals receive their funding reimbursements [75]; however, the method of Cohen’s kappa was misused in this study as it measures the agreement between raters, not the measures [10]. Unlike the MNA, which is a one-page streamlined assessment, the design of the PG-SGA tool is segmented and tedious. Therefore, the limitations of the use of PG-SGA in clinical practice may be that it is time-consuming to complete one assessment and labor-intensive to calculate the rating scores. Similar to the SGA tool, this scored PG-SGA also contains extensively subjective measures of physical examination on 13 areas of body parts; therefore, ongoing training and practice are also required for raters to maintain the high accuracy of malnutrition assessment and diagnosis.

Limitations and recommendations for future studies
Although this is a comprehensive review in which we systematically appraised the quality of psychometric properties of three malnutrition assessment tools in a diverse patient population, one pertinent limitation could be the narrow timeline in the exclusion criteria. For example, Persson’s finding of the inter-rater reliability of PG-SGA was not included in this review because this study was conducted prior to 2000 [76]. Many studies that were identified in this review used a combined criterion to validate the testing tools, and these nutritional parameters are anthropometric assessments, dietary intake assessments, and biochemical measurements. However, challenges exist in nutritional studies, including recall bias, variation in appetite and nutritional symptoms, and confounding factors such as fluid overload and inflammatory effects on nutrition-sensitive markers [5, 6]. Future studies are required to include a critical appraisal tool that addresses nutrition-related confounders, bias, and sources of errors.

Finally, RDs use malnutrition assessment tools for both diagnostic and outcome measure purposes; however, only one article identified in this review appropriately applied statistical methodology to identify the responsiveness to change of the testing tool. For further studies, investigation of outcome measurements related properties, such as minimal detectable change (MDC) and minimal clinically important difference (MCID) are particularly important for both clinicians and researchers to study the treatment effect of nutritional intervention.

Conclusions
A critical review of the clinical measurement properties of three malnutrition assessment tools for use with patients in hospitals was performed. A total of 34 studies were eligible for review, of which 18 were rated to have an acceptable quality of clinical measurement study design. The reliability and validity of all three tools: SGA, MNA, and PG-SGA were assessed; all of them were easy to use, non-invasive, and cost-effective for assessing the malnutrition status of patients.

MNA was the most validated for a variety of measurement properties, whereas SGA was the least studied tool in the last 20 years. Both MNA and PG-SGA possess acceptable test-retest reliability and internal consistency; moreover, PG-SGA had excellent diagnostic accuracy for the identification of malnutrition in various patient populations, and one study that properly examined the responsiveness to change of the MNA tool (MDC = 2.1) was identified [32]. None of the three tools showed a consistently strong correlation with other nutritional parameters and health adverse outcomes, and the inter-rater reliability for both SGA and PG-SGA was not consistently acceptable among the studies identified in this review.

Because of the lack of a gold standard to define malnutrition, this review did not find sufficient evidence to suggest the criterion (concurrent) validity for these studied tools; however, region-specific malnutrition criteria selected to identify coding of malnutrition in inpatients which informs hospital reimbursement funding may be
used as a benchmark in such circumstances to validate the tool use. A future study using sound methodological quality is needed to evaluate the responsiveness to change of these malnutrition assessment tools for the detection of a change in nutritional status.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12973-020-00813-0.

Additional file 1. PRISMA checklist.

Additional file 2. Data extraction of the identified literatures in this review [32–48, 50–52, 62–65, 70, 71, 77–84].

Additional file 3. Quality Appraisal for Clinical Measurement Research Reports.

Additional file 4. Critical Appraisal of Eligible Studies using the Quality Appraisal for Clinical Measurement Study (QA-CMS) [32–48, 50–52, 62–65, 70, 71, 77–84].

Acknowledgments

The authors would like to acknowledge Dr. Linda Greene-Finestone, Ms. Norine Foley, and Dr. Jenna Smith for their support and comments in the preparation of this project and the revision of the final manuscript.

Authors’ contributions

YX and JV conducted the literature search, study selection, and critical appraisal of individual studies. YX completed the data extraction and drafted the manuscript. JV reviewed and approved the manuscript. The author(s) read and approved the final manuscript.

Funding

No funding was received to conduct this systematic review.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 26 May 2020 Accepted: 25 August 2020
Published online: 21 September 2020

References

1. World Health Organization. Malnutrition; 2018 16 [cited 2019 Dec 21]; [about 1 screen]. Available from: https://www.who.int/news-room/factsheets/detail/malnutrition.
2. Allard JP, Keller H, Jeejeebboy KN, Laporte M, Duersken DR, Gramlich L, et al. Malnutrition at hospital admission—contributor and effect on length of stay: A prospective cohort study from the Canadian malnutrition task force. JEPN. 2016;40(4):248–97.
3. Cederholm T, Bosaeus J, Barazzoni R, Bauer J, Van Gossum A, Klek S. Etc. diagnostic criteria for malnutrition—an ESPEN consensus statement. Clin Nutr. 2015;34(3):335–40.
4. Soeters P, Bosetti F, Cyonber L, Forbes A, Shenkin A, Sobotka L. Defining malnutrition: a plea to rethink. Cln Nutr. 2017;36(3):896–901.
5. Finestone HM, Green-Finestone LS, Wilson ES, Teasell RW. Prolonged length of stay and reduced functional improvement rate in malnourished stroke rehabilitation patients. Arch Phys Med Rehabil. 1996;77(4):340–5.
6. Foley NC, Teasell R, Richardson M, Wiener J, Finestone H. Nutritional Interventions Following Stroke. Evid Based Rev Stroke Rehabil. 2018; [Retrieved Dec 20, 2019]. Available from http://www.ebrsr.com/sites/default/files/v18-SREBR-CH16-ET.pdf.
7. Agarwal E, Ferguson M, Banks M, Batterham M, Bauer J, Capra S, et al. Malnutrition and poor food intake are associated with prolonged hospital stay, frequent readmission, and greater in-hospital mortality: Results from the nutrition care day survey 2010. Clin Nutr. 2013;32(5):737–43.
8. Amaral TF, Matos LC, Tavares MM, Subtil A, Martins R, Nazare M. Etc. the economic impact of disease-related malnutrition at hospital admission. Clin Nutr. 2007;26(6):778–84.
9. Veillard J, Fekri O, Dhalli I, Klatzina N. Measuring Outcomes in the Canadian Health Sector: Driving Better Value from Healthcare. Toronto: the C.D. Howe Institute; 2015. 2015Nov [cited 2020 Jan 3]. Commentary 438. Available from: https://www.cdhowe.org/sites/default/files/attachments/research_papers/mixed/Commentary_438.pdf.
10. Fawcett AL. Principles of Assessment and Outcome Measurement for Occupational Therapists and Physiotherapists: Theory, Skills and Application. Chichester: Wiley; 2007. Apr [Cited Jan 3, 2020]. Available from: https://dl.usw.ac.uk/bitstream/handle/13151/1/Principles%20of%20Assessment%20and%20Outcome%20Measurement%20for%20Occupational%20Therapists%20and%20Physiotherapists%20Theory%20%20Skills%20and%20Practice.pdf.
11. Riddle DL, Stratford PW. Is this change real? Interpreting patient outcomes in physical therapy. Philadelphia: F A Davis Company; 2013.
12. MacDermid J. Quality appraisal for clinical measurement studies - evaluation form and guidelines. In: Law MC, MacDermid J, editors. Evidence-based rehabilitation: a guide to practice. 3rd ed. Thorofare: Slack; 2014.
13. BolanirwA OA. Principle and methods of validity and reliability testing of questionnaires used in social and health science researches. Med J. 2015; 22(4):195–201.
14. Portney LG, Watkins MP. Foundations of clinical research: application to practice. New Jersey: Pearson Education, Inc; 2009.
15. Simundic AM. Measure of diagnostic accuracy: basics definition. EJIFCC. 2009(22)(4):203–11.
16. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968;70(4):213–20.
17. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
18. Cronbach LJ. Coefficient alpha and internal structure of tests. Psychometrika. 1951;16(3):297–334.
19. Power L, Mallylo D, Gibney ER, Clarke M, Visser M, VoIkert D, et al. A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings - a MalnuEstu. Clinical Nutrition ESPEN. 2018;24:1–13.
20. Kelley HH, McCullough J, Davidson B, Vesnaver E, Laporte M, Gramlich L, et al. The Integrated Nutrition Pathway for Acute Care (INPAC): Building consensus with a modified Delphi. Nutr J. 2015;14(63). https://doi.org/10.1186/s12937-015-0051-y.
21. Marshall S, Craven D, Kelly J, Isenring E. A systematic review and meta-analysis of the criterion validity of nutrition assessment tools for diagnosing protein-energy malnutrition in the older community setting (the MACRo study). Clin Nutr. 2018;37(6 Pt A):1902–12.
22. Laporte M, Keller HH, Payette H, Allard JP, Duersken DR, Bernier P, et al. validity and reliability of the new Canadian nutrition screening tool in the ‘real-world’ hospital setting. Eur J Clin Nutr. 2015;69(5):558–64.
23. Field LB, Hand RK. Differentiating malnutrition screening and assessment: a nutrition care process perspective. J AcadNutr Diet. 2015;115(5):824–8.
24. Guigoz Y, Vellas B, Garry PJ. Mini nutritional assessment: a practical assessment tool for grading the nutritional state of elderly patients. In: Vellas B, editor. The mini nutritional Assessment (MNA), Suppl 2: 1960;20:37.
25. Ottery FD. Rethinking nutritional support of the cancer patient: the new field of nutritional oncology. SemOncol. 1994;21(6):770–8.
26. Peters L, O’Connor C, Giouso I, Teasell R, Foley N. Screening and assessment of nutritional status following stroke: results from a national survey of registered dietitians in Canada. Disabil Rehabil. 2015;37(26):2413–7.
27. Van Bokhorst-de van der Schueren MA, Guaitoli PR, Jansma EP, de Vet HCW. Nutrition screening tools: does one size fit all? A systematic review of screening tools for the hospital setting. Clin Nutr. 2014;33(1):39–58.
28. Muller C, Compber C, Mary Ellen D. ASPEN Broad of Directors. ASPEN Clinical Guidelines: Nutrition screening, assessment, and intervention in adults. JPN J. 2011;35(1):16–24.
29. Shamseer L, Moher D, Clarke M, Ghereh D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P): 2015: elaboration and explanation. BMJ. 2015;349:g7647.

30. Moher D, Shamseer L, Clarke M, Ghereh D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst. Rev. 2015;4:1.

31. Foley NC, Saltan K, Robertson J, Teasell RW, Woodbury MG. Which reported estimate of the prevalence of malnutrition after stroke is valid? Stroke. 2009; 40(3):e66–74.

32. Lin SC, Lin KH, Lee YC, Peng HY, Chiu EC. Test-retest reliability of the mini nutritional assessment and its relationship with quality of life in patients with stroke. PLoS One. 2019;14(6):e0201874. https://doi.org/10.1371/journal.pone.0201874.

33. Ghazi L, Fereidunnejad SM, Fard SA, Sadeghi M, Shahidi GA, Lokk J. Mini nutritional assessment (MNA) is rather a reliable and valid instrument to assess nutritional status in Iranian healthy adults and elderly with a chronic disease. Ecologie of Food and Nutrition. 2015;54(4):342–57.

34. Soysal P, Isik AT, Arif F, Kalam U, Eyvaz A, Veronese N. Validity of the mini-nutritional assessment scale for evaluating frailty status in older adults. JAMDA. 2019;20(2):183–7.

35. Bauer J, Capra S, Ferguson M. Use of the scored patient-generated subjective global assessment (SGA) in critically ill patients with cancer. Eur J Clin Nutr. 2002;56(8):779–85.

36. Marshall S, Young A, Bauer J, Isenring E. Malnutrition in geriatric rehabilitation: prevalence, patient outcomes, and criterion validity of the scored patient-generated subjective global assessment and the mini nutritional assessment. J AcadNutr Diet. 2016;116(5):785–94.

37. Cooper BA, Bartientt LH, Aslani A, Allen BJ, Ibeis L, Pollock CA. Validity of nutritional assessment of patients with end-stage renal disease using nutritional screening with subjective global assessment predicts hospital stay in patients with digestive diseases. Nutrition. 2007;23(9):634–9.

38. Huang TH, Chi CC, Liu CH, Chang CC, Kuo LM, Hsieh CC. Nutritional status estimate of the prevalence of malnutrition after 20 years of research and clinical practice. –

39. Tsilika E, Parpa E, Panagiotou I, Roumeliotou A, Kouloulias V, Gennimata V. New clinicians require ongoing training to improve adherence to the ESPEN guidelines on enteral nutrition terminology, definitions and general topics. Clin Nutr. 2006;25(2):180–6.

40. Baker JP, Detsky AS, Wesson DE, Wolman SL, Stewart S, Whitevell J, et al. Nutritional assessment: A comparison of clinical judgment and objective measurements. N Engl J Med. 1982;306(16):969–72.

41. Da Silva FJ, Daniel de Mello P, Daniel de Mello E. Subjective global assessment of nutritional status – a systematic review of the literature. ClinNutr. 2015;34(5):785–92.

42. Barbosa-Silva MC, Barros AJ. Indications and limitations of the use of subjective global assessment in clinical practice: an update. CurrOpin Clin NutrMetab Care. 2006;9(3):263–9.

43. Abd Aziz NA, Teng NM, Abdul Hamid MR, Ismail NH. Assessing the nutritional status of hospitalized elderly. Clin Inerv Aging. 2017;4(12):1615–25.

44. Steenson J, Vivanti A, Isenring E. New clinicians require ongoing training to ensure high inter-rater reliability of the subjective global assessment. Nutrition. 2013;29(1):361–2.

45. Newman CS, Cornwell PL, Young AM, Ward EC, Moceri A. Accuracy and confidence of allied health assistants administering the subjective global assessment on inpatients in a rural setting: a preliminary feasibility study. Nutr Diet. 2016;75(1):129–36.

46. Thoresen L, Fjeldstad I, Krogstad K, Kaasa S, Faltmker UG. Nutritional status of patients with advanced cancer: the value of using the subjective global assessment of nutritional status as a screening tool. Palliat Med. 2002;16(1):33–42.

47. Kirushnan BB, Rao BS, Annigeri R, Balasubramanjan S, Seshadri R, Prakash KC. Etc. Impact of malnutrition and atherosclerosis on the outcome in hemodialysis patients. Indian J Nephrol. 2017;27(4):277–83.

48. Ockenga J, Freudeneich M, Zakonsny R, Norman K, Ilicich M, Lochs H. Nutritional assessment and management in hospitalized patients: Implication for DRG-based reimbursement and health care quality. Clin Nutr. 2005;24(6):913–9.

49. Raja R, Lim AV, Lim YP. Malnutrition screening in hospitalized patients and its implication on reimbursement. Intern Med J. 2004;34(4):176–85.

50. Jager-Wittenaar H, Ottery FD. Assessing nutritional status in cancer: role of the patient-generated subjective global assessment. Curr Opin Clin NutrMetab Care. 2017;20(5):322–9.

51. Parasa K, Avaru K. Assessment of nutritional status of Cancer patients using scored PG-SGA tool. J Dent Med Sci. 2016;15(8):37–40.

52. Na BG, Han SS, Cho YA, Wie GA, Kim JY, Lee JM, et al. Nutritional status of Patients with Cancer: A prospective cohort study of 1,588 hospitalized patients. Nutr Cancer. 2018;70(8):1228–36.

53. Martinou E, Bauer J, Isenring E, Cohen S. Malnutrition determined by the patient-generated subjective global assessment is associated with poor outcomes in acute stroke patients. Clin Nutr. 2005;24(6):1073–7.

54. Lee HO, Han SR, Choi SI, Lee JI, Kim SH, Ahn HS, Etc. Effects of intensive nutrition education on nutritional status and quality of life among postgastrectomy patients. Ann Surg Treat Res. 2016;90(2):79–88.

55. Kellet J, Kyle G, Istitoupolis C, Naumont M, Luff N. Malnutrition: the importance of identification, documentation and coding in the acute care setting. J Nutr Metab. 2016. https://doi.org/10.1155/2016/902698.
74. Byron A, McCathie C. Casemix: the allied health response. Med J Aust. 1998;169(S1):46–7.
75. Agarwal E, Ferguson M, Banks M, Baue J, Capra S, Isenring E. Malnutrition coding shortfalls in Australian and New Zealand hospitals. Nutr Diet. 2014;72(1):69–73.
76. Persson C, Spoden PO, Glimelius B. The Swedish version of the patient-generated subjective global assessment of nutritional status: gastrointestinal vs urological cancer. Clin Nutr. 1999;18(2):71–7.
77. Alkan S, Artac M, Rakicioglu N. The relationship between nutritional status and handgrip strength in adult cancer patients: a cross-sectional study. Support Care Cancer. 2018;26:2441–51.
78. Bauer JD, Capra S. Nutrition intervention improves outcomes in patients with cancer cachexia receiving chemotherapy – a pilot study. Support Care Cancer. 2004;13(4):270–4.
79. Fereshtehnejad SM, Ghazi L, Shafieesabet M, Shahidi GA, Delbari A, Lokk J. Motor, psychiatric and fatigue features associated with nutritional status and its effects on quality of life in Parkinson’s disease patients. PLoS One. 2014;9(3):e91153. https://doi.org/10.1371/journal.pone.0091153.
80. Lim HJ, Choue R. Nutritional status assessed by the patient-generated subjective global assessment (PG-SGA) is associated with quality of diet and life in Korean cerebral infarction patients. Nutrition. 2010;26(7–8):766–71.
81. Murphy MC, Brooks CN, New SA, Lumbers ML. The use of the mini-nutritional assessment (MNA) tool in elderly orthopaedic patients. Eur J Clin Nutr. 2000;54:555–62.
82. Ongun N. Does nutritional status affect Parkinson’s disease features and quality of life? PLoS One. 2018;13(10):e0205100. https://doi.org/10.1371/journal.pone.0205100.
83. Pandey D, Buzgeia M, Badr SAE, Senussi FG, El-Mokasabi HI, El-Shahorni AM. Assessing malnutrition among chemotherapy and/or radiotherapy cancer Benghazi outpatients. Nutr Food Sci. 2011;41(5):298–307.
84. Valente KP, Almeida BL, Lazarini TR, Souza VF, Ribeiro TSC, Guedes de Moraes RA, etc. Association of Adductor Pollicis Muscle Thickness and Handgrip Strength with nutritional status in cancer patients. PLoS One 2019 Aug; 14(8): e0220334. doi:https://doi.org/10.1371/journal.pone.0220334.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.