BMJ Open

Screening and assessment tools for early detection of malnutrition in hospitalised children: a systematic review of validation studies

Petra Klanjsek,[1] Majda Pajnkihar,[1] Natasa Marcun Varda,[2,3] Petra Povalej Brzan[1,3,4]

[1]Faculty of Health Sciences, University of Maribor, Maribor, Slovenia
[2]Department of Paediatrics, University Medical Centre Maribor, Maribor, Slovenia
[3]Faculty of Medicine, University of Maribor, Maribor, Slovenia
[4]Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

Correspondence to
Assistant Professor Petra Povalej Brzan; petra.povalej@um.si

ABSTRACT

Objective The aim of the present study was to identify all currently available screening and assessment tools for detection of malnutrition in hospitalised children, and to identify the most useful tools on the basis of published validation studies.

Design Systematic review.

Data sources PubMed, CINAHL and MEDLINE were searched up to October 2017.

Eligibility criteria for selecting studies Studies in English that reported sensitivity, specificity and positive/negative predictive values (PPVs/NPVs) in the paediatric population were eligible for inclusion.

Data extraction and synthesis Two authors independently screened all of the studies identified, and extracted the data. The methodological qualities of the studies included were assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool.

Results The 26 validation studies that met the inclusion criteria for this systematic review used eight screening and three assessment tools. The number of participants varied from 32 to 14 477. There was considerable variability in the chosen reference standards, which prevented direct comparisons of the predictive performances of the tools. Anthropometric measurements were used as reference standards in 16 of the identified studies, and full nutritional assessment in 5. The Pediatric Yorkhill Malnutrition Score (PYMS) screening tool performed better than Screening Tool for the Assessment of Malnutrition and Screening Tool for Risk On Nutritional status and Growth when compared in terms of anthropometric measurements, especially for body mass index (Se=90.9, Sp=81.9) and triceps skinfold thickness (Se=80.0, Sp=75.0). However, low PPVs indicated the problem of overprediction of positive cases, which was typical for all of the studies that used anthropometric measurements as the reference standard.

Conclusions This systematic review identifies the need for definition of the gold standard for validation of screening tools. Anthropometry measurements using WHO or Centers for Disease Control and Prevention growth charts should be considered as the possible reference standard in future validation studies. We would recommend the use of PYMS for hospitalised paediatric patients without chronic conditions, in combination with full nutritional assessment.

PROSPERO registration number CRD42017077477.

Strengths and limitations of this study

► This systematic review was based on a comprehensive search and includes a large number of screening/assessment tools for evaluation of the malnutrition risk in hospitalised children, along with their validation studies.

► Only the studies in English that reported sensitivity, specificity and positive/negative predictive values or data enabling manual calculation of them were included.

► The methodological quality of the validation studies included was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool.

► This systematic review highlights the heterogeneity of both the tools available and their validation studies, along with the challenges that result from this heterogeneity.

► Although our search included multiple electronic databases and grey literature, relevant data that have not been reported may be missed.

INTRODUCTION

Over the last decade, several studies have shown that the prevalence of malnutrition in hospitalised children varies from 6.1% to 55.6% worldwide.[1–6] The importance of early detection of malnutrition in hospitalised paediatric patients has led to the development of several nutritional screening and assessment tools. Screening tools are designed to provide early identification of children at risk of nutritional impairment, and they have the potential to improve health outcomes and to reduce healthcare costs. All patients considered at risk during such screening should be referred for nutritional assessment and possible intervention. However, currently, there is no consensus on the appropriate screening tool to identify these children who are at risk of developing malnutrition during hospitalisation.[7–9]

Due to the absence of a gold standard,[8 10–14] screening/assessment tools are

usually validated using the following reference standards: dietetic/nutritional assessment; anthropometric measures, as defined by the WHO[15 16]; growth charts of the Centers for Disease Control and Prevention (CDC)[17]; national growth charts or Roher's Ponderal Index for newborns.

The CDC and National Center for Health Statistics growth charts for the USA were released in the year 2000.[18] The WHO released international growth charts for children up to 5 years of age in 2006. In addition, in 2007, the WHO developed a growth reference for school-age children and adolescents (≤19 years).[19] According to the WHO and CDC recommendations, the WHO growth charts are more appropriate for children aged from 0 to 2 years. The methods used to create the CDC and WHO growth charts were similar for children from 2 to 5 years of age. The CDC growth charts can be used for children <5 years.[20] The WHO growth charts had already been adopted in 125 out of 219 countries by the end of April 2011.[21] National growth charts have also been used in some countries, or parts of countries, including China[22] and the United Arab Emirates.[23]

Dietetic/nutritional assessment is the systematic process of collecting and interpreting information to make decisions on the nature and causes of nutrition-related health issues that affect an individual. Full nutritional assessments also include biochemical parameters.[24] However, dietetic/nutritional assessments vary across different countries due to differences in educational standards.[3]

Anthropometric measures such as weight for age (WFA), height for age (HFA), weight for height (WFH) and body mass index (BMI) for age, as SD scores (ie, Z-scores) are usually used for identification of malnutrition.[25] Malnutrition can be acute (ie, wasting) or chronic (ie, stunting). Moderate acute malnutrition is usually defined using WFH, as Z-scores between −3 and −2. A WFH Z-score less than −3 indicates severe acute malnutrition. Chronic malnutrition is defined using HFA, as Z-scores between −3 and −2 for moderate chronic malnutrition, and Z-scores less than −3 for severe chronic malnutrition. Furthermore, mid-upper arm circumference (MUAC) is used for the identification of malnutrition in infants and children aged 3 months to 5 years, with the cut-off Z-score of less than −2.[25 26]

The aim of the present study was to systematically review the available publications on the screening and/or assessment tools for hospitalised children, with a focus on the ability of these tools to predict the risk or presence of malnutrition, in order to identify the most useful tool for use in the clinical environment.

METHODS

Design

This systematic review of published validation studies was registered with PROSPERO (online supplementary file 1). The findings are reported according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement[27] (see PRISMA checklist; online supplementary file 2).

We focused on answering the following research questions (RQs) (see online supplementary file 3): RQ1: What are the currently available screening and assessment tools for detecting malnutrition in hospitalised paediatric patients? RQ2: What is the validity of the screening and assessment tools versus the reference standard?

Search strategy

To identify all relevant publications, we performed systematic searches in the following bibliographic databases: PubMed, CINAHL and MEDLINE. The searches were carried out on October 20, 2017. The keyword combinations used included the following: premature*, immature*, child*, baby, infant*, newborn*, neonate*, kid*, babies, adolescent*, pediatric*, paediatric*, screen*, assess*, tool*, undernutrition*, undernourish*, malnutrition, malnourish*. The search strategies are outlined in online supplementary file 4, and they were adapted to each database and kept consistent across all searchers. The reference lists of the identified studies were manually searched for potentially relevant studies.

Study selection

All potentially relevant titles and abstracts were blinded for author, journal and year of publication, and then screened for eligibility by two reviewers (PK and PPB), independently. Differences in judgement were resolved through a consensus procedure. All of the studies obtained from the bibliographic databases were entered into EndNote X8, and duplicates were excluded. The inter-rater agreement between the reviewers based on Cohen's kappa statistic was 0.79 for 576 studies. The full texts of the selected studies—no longer blinded to authors and journals—were obtained for further review by two reviewers, independently (PK and PPB), to judge for eligibility. The Cohen's kappa coefficient here was 0.93 for 64 studies, which indicated a high level of agreement. In cases of doubt, a decision was made by a third reviewer (NMV).

The studies were reviewed to ensure their focus was aligned with the purpose of the literature review. Those that were clearly inappropriate to answer the RQs and/or did not fit the predefined inclusion criteria were excluded. The flowchart of the complete search and selection process is shown in figure 1.

Inclusion/exclusion criteria

The studies eligible for inclusion were validation studies in English that reported sensitivity (Se), specificity (Sp) and positive/negative predictive values (PPVs, NPVs, respectively) in paediatric populations. All of the predefined inclusion and exclusion criteria were used as outlined in online supplementary file 5. The list of studies (n=38) not meeting the selection criteria after reading the full text is presented in online supplementary file 6.
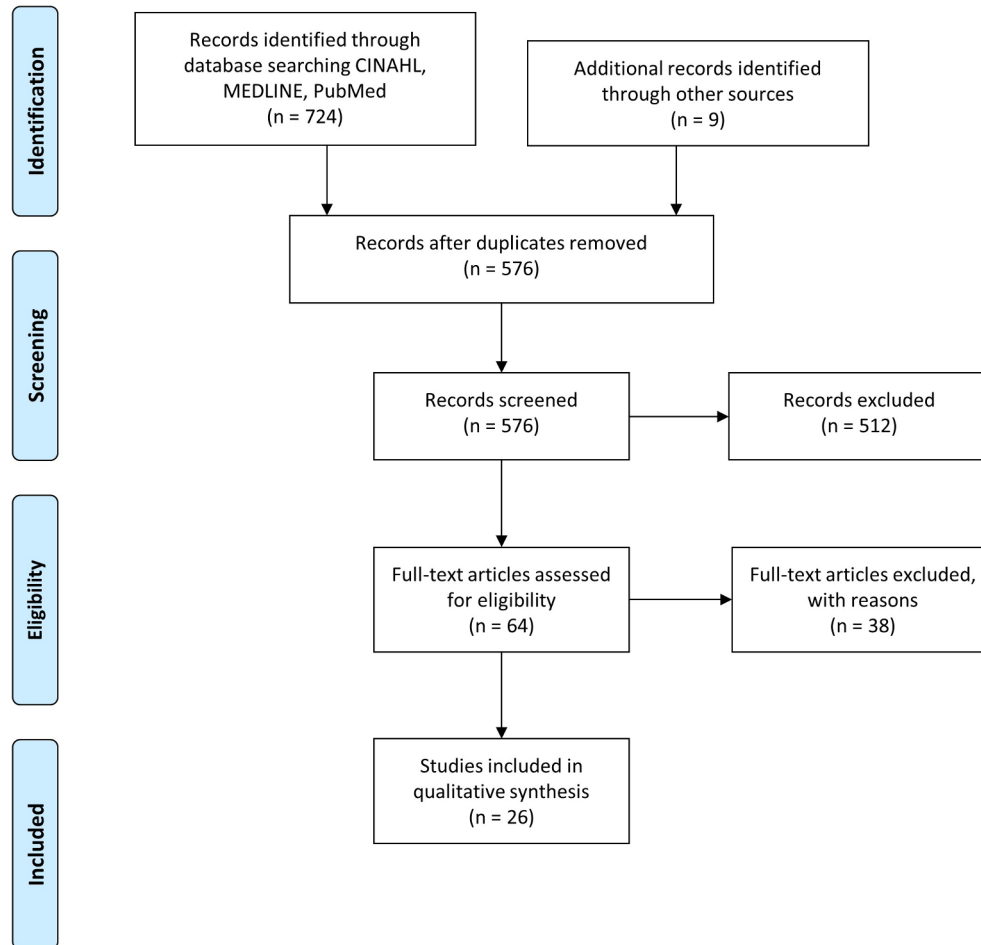
Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram of the search and study selection process.

Quality appraisal

The methodological quality assessment of the studies was performed using Review Manager V.5.3,[28] with a revised tool for the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2.[29] The QUADAS-2 tool uses four key domains to rate the risk of bias and the applicability of primary diagnostic accuracy studies. The key domains are as follows: patient selection (sampling, inclusion/exclusion criteria, sampling bias, adequacy), index tests (the validated tool, correct use and interpretation, possible bias), reference standards (the reference tool, correct use and interpretation, possible bias) and the flow and timing (the sequence, time interval, correct performance of reference standard and index test, possible bias). The results from QUADAS-2 can be expressed as high/unclear/low risk of bias and as high/unclear/low applicability concerns.

Data extraction and synthesis

The data were extracted by two reviewers (PK and PPB) and checked by a third reviewer (MP) using the predefined data-extraction criteria, which included the following: authors and country, nutritional screening/assessment tool used, study type, sample size, age of participants and reported clinical performance. To evaluate the clinical performance and diagnostic accuracy of the screening tools, the following criteria were considered: Se, Sp, PPV and NPV. Studies that did not report Se, Sp, PPV and/or NPV, but that provided the data that enabled calculation of these values, were also included in the study. These metrics were subsequently calculated manually by the authors and are indicated as such in the results tables.

The validation of the reproducibility and reliability of the screening/assessment tools was also considered, using data from the agreement analysis between the assessed tools and the chosen reference, as well as the inter-rater agreement shown in the studies.

For reasons of clarity, we have rated the results of each study as good, moderate/fair or poor validity. The kappa values were rated by the classification system proposed by Landis and Koch.[30] Although the literature does not provide general cutoffs for Se and Sp, as they greatly depend on the clinical consequences, we have rated the values to maintain transparency and clarity, as in van Bokhorst-de van der Schueren et al.[31] All of the cutoff points are described in table 1.

Patient and public involvement

No patients or public were involved in the present study.

Table 1 Cutoffs applied to assess the validity of the nutritional screening and monitoring tools

Assessment	Code	Rating	Cutoff
Sensitivity (Se)/	g	Good	Se and Sp≥80%
Specificity (Sp)	f	Fair	Se or Sp<80%, but both >50%
	p	Poor	Se or Sp≤50%
Kappa[30]	pe	Almost perfect	0.81–1.00
	su	Substantial	0.61–0.80
	m	Moderate	0.41–0.60
	f	Fair	0.21–0.40
	s	Slight	0–0.20
	n	No agreement	<0

RESULTS

Search results

Figure 1 shows the flow diagram of the search and study selection process. We identified 724 studies initially, of which 26 met all of the pre-established eligibility criteria and were included in the critical appraisal.

Paediatric nutritional screening tools identified

During this systematic review, we identified eight validated pediatric nutritional screening tools and three nutritional assessment tools for hospitalised paediatric patients. These can be classified based on their specialties, as follows:

▶ Medical and Surgical Department: Screening Tool for the Assessment of Malnutrition in Paediatrics (STAMP),[32 33] Paediatric Yorkhill Malnutrition Score (PYMS),[7 34] Screening Tool for Risk On Nutritional status and Growth (STRONG$_{kids}$)[2] and Paediatric Nutrition Screening Tool (PNST).[9]
▶ Surgical Department: Subjective Global Assessment (SGA)[35–38] and Subjective Global Nutritional Assessment (SGNA).[13]
▶ Oncology Department: Nutrition Screening tool for childhood Cancer (SCAN),[39] which was developed specifically for children with cancer.
▶ Pulmonology Department: Nutrition screening tool for pediatric patients with cystic fibrosis (CF),[40] and Nutritional risk screening tool in CF[41] for pediatric patients with CF.
▶ Neonatal Intensive Care: Neonatal Nutrition Screening Tool (NNST)[42] for infants in the Neonatal Intensive Care Unit.
▶ Clinical Assessment of Nutritional Status (CANS) score,[43] to differentiate malnourished from appropriately nourished babies.

However, six additional screening tools were identified in the eligibility step of PRISMA, although the associated studies were excluded in the final step as the inclusion criteria were not met:

▶ Medical Department: Nutrition Risk Score (NRS).[44]

▶ Medical and Surgical Department: Simple Paediatric Nutrition Risk Score (SPNRS)[45] and Pediatric Digital Scaled MAlnutrition Risk screening Tool (PeDiSMART).[46 47]
▶ Psychiatric Department: St Andrew's Nutrition Screening Instrument (SANSI).[48]
▶ Screening tool to predict malnutrition among children under 2 years old in Zambia.[49]
▶ Nutrition screening for infants and young children with special health care needs: A Look at Your Child's Nutrition.[50]

Characteristics of the studies included in the systematic review

The characteristics of the 26 studies included in this systematic review are outlined in table 2. Sample sizes varied from 32 to 14 477 participants. Eleven studies (42.3%) excluded patients with length of hospital stay (LOS) of <24 hours.[2 9 14 32 39 51–56] The studies often excluded children <1 year (57.7%),[2 6 7 32 38–41 52 53 55–59] intensive care unit (ICU) patients (34.6%)[2 14 38 51 52 58 60–62] and patients with unstable/specific conditions, such as oncology patients,[53–55 61] conditions that affected hydration,[9 39] cardiology, renal and orthopaedic specialties,[7] fever, diarrhoea,[6] obesity[53] and others. Some of the studies included only patients with particular conditions, where a specially designed screening/assessment tool was usually used. However, STAMP was designed for clinical and surgical patients and was validated also on patients with spinal cord injury (SCI)[6 14] and with inflammatory bowel disease (IBD).[59] Additionally, STAMP was also used in outpatients in two studies.[57 59] The SGA assessment tool was originally designed for adults; however, in one of the selected studies, it was tested on children.[38]

There was relatively high heterogeneity in the choice of the reference standards. Anthropometric measurements were used in 18 of the selected studies (69.2%)[2 6 9 38 42 51–55 58–65] and full dietetic/nutritional assessments were used in 5 studies (19.2%).[7 14 32 41 57] Three studies (12.5%) used SGNA as the reference standard.[9 39 56] The Cystic Fibrosis Foundation (CFF) Consensus Report criteria were used as the reference standard in both of the studies on patients with CF.[40 41] Validation with other screening tools along with the reference standards were reported in four studies (16.7%).[14 40 59 63] Nine studies (37.5%) reported validation of two or more screening tools with the same reference standard and on the same patients.[9 53 54 56 57 59 61 63 65]

Risk of bias and applicability concerns

The results of the quality appraisal analysis using the QUADAS-2 tool are presented in online supplementary file 7.

The patient selection was considered as high risk or unclear because of the non-specific description of the patient selection process in three studies.[7 41 58] Possible bias from conducting non-blinded index tests with respect to the results of the reference standard

Table 2 Characteristics of the studies included in the systematic review

Study	Tool	Country	Time frame	Population	Sample size	Age range	Exclusion criteria	Reference standard
Chourdakis et al[63]	PYMS; STAMP; STRONGkids	12 countries: Croatia, Israel, Denmark, Italy, France, Poland, Germany, Greece, The Netherlands, UK, Romania	Feb 2010 to July 2011	Clinical and surgical patients	2567	1 month to 18 years	Accident and Emergency Department of Day Care Unit	WHO (BMI, MUAC, TSFT)
Durakbaşa et al[60]	STRONGkids	Turkey	April to July 2012	Surgical patients	494	>1 month	Age <30 days, admitted to clinics other than paediatric surgery ward or had another operation in the preceding 30 days	WHO (WFH, BMI) for acute malnutrition, WHO (HFA) for chronic malnutrition
Galera-Martínez et al[61]	STRONGkids; STAMP	Spain	May 2013	Clinical and surgical patients	223	>1 month	Paediatric ICU, oncology, day surgery ward, patients who could not be weighed and measured at admission	WHO (BMI SDS)
Gerasimidis et al[7]	PYMS	UK	23 June to 28 Oct 2008	Clinical and surgical patients	247	1–16 years	Cardiology, renal, orthopaedic specialties, critical care	Full dietetic assessment: dietary history, anthropometric measurements. Body composition characteristics
Hulst et al[2]	STRONGkids	The Netherlands	26–28 Nov 2007	Clinical and surgical patients	417	>1 month to 18 years	ICU, LOS <1 day	WHO (HFA, WFH) on The Netherlands National Standards[16]
Huysentruyt et al[51]	STRONGkids	Belgium	Dec 2010 to April 2011	Clinical and surgical patients	368	1 month to 16 years	<24 hours LOS, ICU, readmitted for same condition within 7 days	WHO (WFH) for acute malnutrition, WHO (HFA) for chronic malnutrition
Johnson et al[42]	NNST	UK	2010	Neonatal patients in ICU	909		Infants <2 weeks at discharge	WHO (WFA) with UK-WHO growth chart
Ling et al[52]	STAMP; STRONGkids	UK	Two consecutive days in August 2009	Inpatients	56	6 weeks to 16 years	Age <1 month, LOS <1 day, height measurement not accurately obtained, ICU, day surgery ward	WHO (WFH, BMI if HFA); WFH: height<120 cm; BMI: height>120 cm; WHO (HFA) for chronic malnutrition
Mahdavi et al[38]	SGA	Iran	June 2008 to Aug 2008	Clinical and surgical patients	140	2–12 years	Emergency Department, newborn ICU, newborn Special Care Units	Anthropometric, biochemical measurements (WFA, HFA, WFH, TSF, MAC, serum albumin, transferrin). Patients classified as undernourished when at least two parameters are subnormal
Mărginean et al[53]	STRONGkids; modified STRONGkids	Romania	1 May 2011 to 30 Jan 2012	Tertiary paediatric teaching hospital	326	1–17 years	Obese, active malignancy, LOS <24 hours	WHO (WFH, HFA); total serum protein (normal: 6.6–8.7 mg/dL)

Continued

Table 2 Continued

Study	Tool	Country	Time frame	Population	Sample size	Age range	Exclusion criteria	Reference standard
Martinez-Nadal et al[64]	CANS score	Spain	2003–2014	Neonatal patients	14477	Newborns	Not reported	Roher's Ponderal Index (PI<2.2g/cm³)
McCarthy et al[32]	STAMP	UK	April to June 2004	Clinical and surgical patients	238	2–17 years	LOS <24hours, unobtainable weight or height measurements	Full nutritional assessment: historical records, nutritional measurements, anthropometry
McDonald[41]	Nutrition Risk Screening Tool in CF	USA	Not reported	Paediatric patients with CF diagnosis	85	2–20 years	Not reported	CFF Consensus Report criteria; full dietetic assessment: MUAC and/or TSF less than 25th percentile for sex and age, suboptimal serum vitamin and/or trace mineral levels, pulmonary exacerbation, suboptimal dietary intake
Moeeni et al[54]	PYMS; STRONGkids; STAMP	Iran	26 Dec 2010 to 19 Jan 2011	Tertiary paediatric teaching hospital	119	1–17 years	Malignancy, LOS <24hours	WHO (WFH, HFA, BMI if WFH or HFA missing)
Moeeni et al[55]	STRONGkids – simplified	New Zealand	30 Oct to 30 Dec 2012	Clinical and surgical patients	162	1 month to 17 years	LOS <24hours, oncological patients	WHO (BMI, WFA, HFA, WFH)
Murphy et al[39]	SCAN	Australia	Not reported	Paediatric oncology patients	32	5–18 years	LOS <24hours, clinically unstable, conditions that affect hydration, non-English speaking	SGNA
Rub et al[57]	STAMP; STAMP-modified	Israel	Not reported	Outpatients	60	1–6 years	Children cared for in outpatient clinics	Full nutritional assessment: dietary history, anthropometric measurements, body composition characteristics
Soundarya et al[62]	CANS score	India	Not reported	Maternity hospital	300	0–48 hours	Newborns with congenital anomalies, newborns <37 completed weeks gestation, multiple pregnancies, NICU care, mother's gestational diabetes mellitus, unreliable estimation of gestational age	Anthropometry based on Indian growth charts (Roher's Ponderal Index, head circumference to length ratio, chest circumference or MUAC and/or MUAC to head circumference ratio, BMI)
Souza Dos Santos et al[40]	Nutrition screening tool for pediatric patients with CF	Brazil	Not reported	Pediatric patients with CF diagnosis	82	6–18 years	Not reported	CFF Consensus Report criteria; Nutritional risk screening tool proposed by McDonald[41]
Spagnuolo et al[58]	STRONGkids	Italy	Oct 2012 to Nov 2012	Clinical and surgical patients	144	1–18 years	ICU patients	WHO (BMI) for acute malnutrition, WHO (HFA) for chronic malnutrition
Thomas et al[65]	STAMP; PMST (STAMP-modified); PYMS	UK	Dec 2014 to March 2015	Children in tertiary hospital acute units	300	0–17.6years	Not reported.	WHO (HFA, WFH, BMI)

Continued

Table 2 Continued

Study	Tool	Country	Time frame	Population	Sample size	Age range	Exclusion criteria	Reference standard
Wang et al[6]	STAMP	China	July 2014 to July 2015	Inpatients and outpatients with SCI	45	1–12 years	Fever or diarrhoea on admission, history of hepatopathy, nephropathy or formal dietetic assessment and nutritional therapy that may have introduced bias towards nutrition status from medical records	WHO (BMI for age, WFA, HFA)
White et al[9]	PNST; SGNA	Australia	Sept 2012 to June 2013	Clinical and surgical patients	295	0–16 years	LOS <24 hours, clinically unstable, conditions that affect hydration, non-English speaking	WHO (BMI for age, HFA, WFA) for 0–2 years; CDC 2000 for 2–20 years; SGNA
Wiskin et al[59]	STAMP; STRONG_kids; SPNRS; PYMS	UK	Dec 2009 to June 2010	Paediatric patients with IBD, outpatients and inpatients	46	3–17 years	Not reported.	ICD-10; mild: weight SDS <-2; moderate: SDS between -2 and -3; severe: SDS ≥-3
Wong et al[14]	STAMP	UK	Jan to Dec 2010	Paediatric patients with SCI	51	6 months to 18 years	LOS <24 hours, ICU	Full dietetic assessment: clinical, nutritional, historical records
Wonoputri et al[56]	PYMS; STAMP; STRONG_kids	Indonesia	Jan to Feb 2014	Clinical patients	116	1–15 years	LOS <24 hours	SGNA, WHO growth chart (results not reported)

BMI, body mass index; CANS, Clinical Assessment of Nutritional Status; CDC, Centers for Disease Control and Prevention; CF, cystic fibrosis; CFF, Cystic Fibrosis Foundation; HFA, height for age; IBD, inflammatory bowel disease; ICD, International Statistical Classification of Diseases and Related Health Problems; ICU, intensive care unit; LOS, length of hospital stay; MAC, mid-arm circumference; MUAC, mid-upper arm circumference; NICU, Neonatal Intensive Care Unit; NNST, Neonatal Nutrition Screening Tool; PNST, Paediatric Nutrition Screening Tool; PYMS, Paediatric Yorkhill Malnutrition Score; SCAN, Nutrition Screening Tool for Childhood Cancer; SCI, spinal cord injuries; SDS, SD score; SGA, Subjective Global Assessment; SGNA, Subjective Global Nutritional Assessment; SPNRS Simple Paediatric Nutrition Risk Score; STAMP, Screening Tool for the Assessment of Malnutrition in Paediatrics; STRONG_kids, Screening Tool for Risk on Nutritional Status and Growth; TSFT, triceps skin fold thickness; WFA, weight for age; WFH, weight for height.

and/or vice versa was considered as high risk in seven studies,[7 9 38 39 56 59 62] and as a possibility for bias in 15 studies.[2 6 14 32 40–42 51 53 58 60–62 64 65] Similarly, non-blinded interpretation of the reference standard with regards to the index test results was considered as a possibility for bias in 17 studies.[2 14 32 39–42 51–53 56 58 60–62 64 65] The information about the patient flow and timing was considered to be unclear in 15 studies,[2 7 14 32 38 40 42 53 54 56 58–60 62 65] as the intervals between the index tests and the reference standard measurements were not reported. In one study,[62] the anthropometric measurements were used as the index test and compared with the reference standard CANS score, which in our opinion introduces possible bias.

The reference standard was the second key domain of concern regarding applicability. Murphy *et al*[39] used the SGNA tool for the reference standard, which is an assessment method rather than a reference standard.[7] Similarly, Wonoputri *et al*[56] defined the WHO anthropometric grow chart as the reference standard; however, the results presented only showed the comparisons of PYMS, STAMP and STRONG$_{kids}$ with SGNA. As mentioned before, Soundarya *et al*[62] used the CANS score as the reference standard; however, in our report we present the results as reported in the study and also as calculated from the reported data in an inverted manner. The inclusion or exclusion criteria were not clearly defined in five studies.[7 41 56 58 65]

Evaluation of the screening/assessment tools

The diagnostic accuracy of the validated screening/assessment tools for the chosen reference standards in the selected studies is presented in table 3.

Nine studies (34.61%) did not report Se, Sp, PPV and NPV; however, the data reported enabled the calculation of these validation metrics. Additionally, two studies did not report PPV and NPV, which were subsequently calculated by the authors of the present study. The calculated values are highlighted with dagger (†) in table 3. In one study, only the Se was reported, with no data for the calculation of the other validation metrics.[55]

The sensitivities of the screening/assessment tools ranged from 20%[63] to 100%.[6 39 52 56 59] SCAN with oncological patients and STAMP and STRONG$_{kids}$ with clinical patients showed the best results versus SGNA in terms of Se.[39 56] Additionally, STAMP performed with 100% Se versus the anthropometric measurements for inpatients and outpatients with SCI.[6] STAMP, STRONG$_{kids}$, SPNRS and PYMS obtained 100% Se versus International Statistical Classification of Diseases and Related Health Problems (ICD)-10; however, the Sp was poor (0%), except for PYMS (53.5%).[59]

The specificities ranged from 0% (STAMP, STRONG$_{kids}$, SPNRS vs ICD-10)[59] to 96.4% (STRONG$_{kids}$ vs PYMS).[39] The combination of Se and Sp was evaluated as good only in four studies[7 40 61 63] according to classification outlined in table 1.

Relatively high NPVs were seen for most of the studies, which ranged from 0%[58] to 100% (SCAN vs SGNA;

STAMP vs SGNA; STRONG$_{kids}$ vs SGNA; PYMS vs ICD-10; STAMP vs nutritional intervention; STRONG$_{kids}$ vs nutritional intervention).[6 39 52 59] On the contrary, the observed PPVs were a lot lower in most of the studies; these started from 2.5% (STAMP vs TSFT)[63] and reached as high as 100% (STAMP, STRONG$_{kids}$, SPNRS, PYMS, among each other).[59]

Agreement between the nutritional screening/assessment tool and the reference standard or other screening tool was verified in 12 studies (46.1%).[6 7 14 32 38 40 53 56 57 59 63 65]

As presented in table 3, all of the abovementioned validation metrics differed greatly when different cutoff values were used. The studies also included different populations and different sample sizes, and therefore direct comparisons of the results are not possible.

Seven studies (26.9%)[7 14 32 41 51 55 61] also reported interobserver agreements, which varied from 0.53 for PYMS completed by two dietitians compared with the nursing staff,[7] to 0.9 for STAMP completed by dietitians and nursing staff.[32] Only two studies reported intraobserver agreement, where there was substantial agreement with the kappa value of 0.6 for STRONG$_{kids}$[51] and 0.6 for STAMP.[14]

DISCUSSION

This section discusses the results and the main findings of the present study. Recommendations for new studies that focus on validation of nutritional screening and assessment tools are proposed.

Two RQs were proposed in the present study, as follows.

RQ1: What are the currently available screening and assessment tools for detecting malnutrition in hospitalised paediatric patients?

Currently, there are 14 nutritional screening tools and 3 nutritional assessment tools. In this systematic review, we identified validation studies of eight nutritional screening tools (SCAN, Nutritional screening tool for paediatric patients with CF, Nutritional risk screening tool in CF, NNST, PYMS, STRONG$_{kids}$, STAMP and PNST) and three malnutrition risk assessment tools (SGA, SGNA and CANS score). Six screening tools were not included in this systematic review (NRS, SPNRS, PeDiSMART, SANSI, Screening tool to predict malnutrition among children under 2 years old in Zambia and Nutrition screening for infants and young children with special health care needs: A Look at Your Child's Nutrition), as the studies identified did not include validation of the screening tool.

It is important to emphasise that the nutritional screening and assessment tools were developed and validated for different populations of children, which were mainly focused on age limits and different exclusion criteria in terms of the admission diagnosis/status or other chronic illnesses.

Table 3 Sensitivity (Se), specificity (Sp), positive and negative predictive values (PPV and NPV) and agreement of studies included in the systematic review

Study	Tool	Reference standard	Se (%)	Sp (%)	Rating	PPV (%)	NPV (%)	Inter-/intra observer agreement (kappa)	Agreement with reference standard/other screening tools	Rating kappa
Chourdakis et al[63]	PYMS	WHO BMI <-2SD*	90.91†	81.97†	g	21.90†	99.39†			
	STAMP		77.27†	81.21†	f	18.61†	98.47†			
	STRONG$_{kids}$		45.45*	91.74†	p	23.44†	96.80†			
	PYMS	MUAC <-2SD*	66.67†	75.19†	f	7.55†	98.67†			
	STAMP		81.82†	78.31†	f	9.89†	99.33†			
	STRONG$_{kids}$		41.67†	92.80†	p	15.15†	98.10†			
	PYMS	TSFT <-2SD*	80.00†	75.07†	f	4.08†	99.65†			
	STAMP		40.00†	78.09†	p	2.50†	98.93†			
	STRONG$_{kids}$		20.00†	92.78†	p	3.7 0†	98.82†			
	PYMS	STAMP*	58.70†	88.37†	f	59.31†	88.11†		STAMP: k=0.47 (95% CI 0.42 to 0.53)	m
	STAMP	STRONG$_{kids}$*	77.30†	84.11†	f	36.82†	96.87†		STRONG$_{kids}$: k=0.47 (95% CI 0.42 to 0.53)	m
	STRONG$_{kids}$	PYMS*	31.40†	96.54†	p	74.51†	81.38†		PYMS: k=0.35 (95% CI 0.28 to 0.42)	f
Durakbaşa et al[60]	STRONG$_{kids}$	WHO								
		Acute malnutrition: WFH, BMI‡	48.00†	65.77†	p	13.64†	91.82†			
		Chronic malnutrition: HFA‡	52.17†	45.18†	p	6.82†	96.54†			
Galera-Martinez et al[61]	STRONG$_{kids}$ experts	WHO BMI <-2SD*	37.50†	91.16†	p	13.64†	97.51†	Interobserver: k=0.72 (95% CI 0.63-0.80)^m		
	STRONG$_{kids}$ non-experts		37.50†	94.42†	p	20.00†	97.60†			
	STAMP experts		87.50†	82.79†	g	15.90†	99.44†	Interobserver: k=0.74 (95% CI 0.67-0.81)^m		
	STAMP non-experts		62.50†	79.53†	f	10.20†	98.27†			
Gerasimidis et al[7]	PYMS^c	Full dietetic assessment*	59	92	f	47	95	Interobserver: k=0.53 (95% CI 0.38 to 0.67)^b	AR^c: k=0.46 (95% CI 0.27 to 0.64)	m
									STAMP^c: k=0.47 (95% CI 0.34 to 0.61)	m
									SGNA^c: k=0.12 (95% CI −0.11 − −0.34)	s
	PYMS^d		85	87	g	44	98		AR^d: k=0.51 (95% CI 0.40 to 0.70)	m
	STAMP^d		81	78	f	31	97		STAMP^d: k=0.34 (95% CI 0.20 to 0.50)	f
	SGNA^d		15	100	p	100	91		SGNA^d: k=0.24 (95% CI 0.10 to 0.50)	f

Table 3 Continued

Study	Tool	Reference standard	Se (%)	Sp (%)	Rating	PPV (%)	NPV (%)	Inter-/intra observer agreement (kappa)	Agreement with reference standard/other screening tools	Rating kappa
Hulst et al[2]	STRONG_kids	WHO (HFA, WFH)‡	75.32†	41.47†	p	22.57†	88.13†			
Huysentruyt et al[51]	STRONG_kids	WHO (WFH, HFA)‡	71.9^e	49.1^e	p	11.9^e	94.8^e	Interobserver: k=0.61		
			69^f	48.4^f	p	10.4^f	94.8^f	Intraobserver: k=0.66		
Johnson et al[42]	NNST	UK-WHO growth chart (WFA)	89.6	75.1	f	32.9	98.1			
Ling et al[52]	STAMP	Nutritional intervention‡	100†	40†	p	41.94†	100†			
	STRONG_kids		100†§	38.89†§	p	35.29†§	100†§			
Mahdavi et al[38]	SGA	Anthropometric and biochemical measurements WHO (WFA, HFA, WFH, TSFT and serum transferrin)	88.24	45.83	p	60.61	80.49		AR^i: k=0.336	f
Mârginean et al[53]	STRONG_kids	WHO (WFH, HFA)‡	97.00†	50.29†	f	53.30†	96.63†		AR: k=0.61	su
	Modified STRONG_kids		97.00†	65.50†	f	62.18†	97.39†		AR: k=0.716	su
Martinez-Nadal et al[64]	CAN score	Roher's Ponderal index	48.90†	95.47†	p	44.60†	96.16†			
McCarthy et al[32]	STAMP	Full nutritional assessment*	70 (95% CI 51 to 84)	91 (95% CI 86 to 94)	f	54.8 (95% CI 38.8 to 69.8)	94.9 (95% CI 90.5 to 97.4)	Interobserver: k=0.92 (95% CI 0.76 to 1.00)	AR: k=0.54 (95% CI 0.46 to 0.62)	m
McDonald[41]	Nutrition Risk Screening	CFF Consensus Report criteria‡	84	75	f	82†	77†	Interobserver: k=0.85^a		
	Tool in CF	Full dietetic assessment‡	86	78	f	84†	80†			
Moeeni et al[54]	PYMS	WHO (WFH, HFA, BMI if WFH or HFA missing)‡	90.00†	31.46†	p	30.68†	90.32†			
	STRONG_kids		83.33†	49.44†	p	35.71†	89.80†			
	STAMP		90.00†	37.08†	p	32.53†	91.67†			
Moeeni et al[55]	STRONG_kids	WHO (BMI, WFA, HFA, WFH)	84^k		/			Interobserver: k=0.65^j		
	simplified		90^l		/					
Murphy et al[39]	SCAN	SGNA	100 (95% CI 76 to 100)	39 (95% CI 17 to 64)	p	56 (95% CI 35 to 76)	100 (95% CI 59 to 100)			
Rub et al[57]	STAMP	Full nutritional assessment‡	47.6 (95% CI 28.3 to 67.6)	94.9 (95% CI 81 to 99)	p	83.3	77.1		AR: k=0.47 (95% CI 0.35 to 0.79)	m
	STAMP-modified		76.19 (95% CI 54.91 to 89.37)	82.05 (95% CI 67.33 to 91.02)	f	69.6	86.5		AR: k=0.57 (95% CI 0.35 to 0.79)	m
Soundarya et al[62]	CAN Score	Roher's Ponderal index	65.2	86.4	f	60.2	88.7			
			60.26†	88.74†	f	65.28†	86.40†			

Continued

Table 3 Continued

Study	Tool	Reference standard	Se (%)	Sp (%)	Rating	PPV (%)	NPV (%)	Inter-/intra observer agreement (kappa)	Agreement with reference standard/other screening tools	Rating kappa
		BMI	84.7	73.6	f	50.4	93.8			
		Roher's Ponderal index+BMI	50.41†	93.85†	f	84.72†	73.68†			
			84.7	72.8	f	49.5	93.7			
		Mid arm circumference/head	41.6	77.6	p	37.0	80.8			
		circumference ratio	37.04†	80.82†	p	41.67†	77.63†			
Souza Dos Santos et al[40]	Nutrition screening tool for pediatric patients with CF	CFF Consensus Report criteria‡	72.4	71.7	f	58.33†	82.61†	AR: k=0.418 (p<0.001)		m
		Nutritional risk screening tool proposed by McDonald‡	85	95.2	g	94.44†	86.96†	AR: k=0.804 (p<0.001)		su
Spagnuolo et al[58]	STRONGkids	WHO (BMI, HFA for acute and chronic malnutrition)‡	71 (95% CI 48 to 89)	53 (95% CI 43 to 63)	f	21 (95% CI 17 to 25)	85 (95% CI 85 to 90)			
Thomas et al[65]	STAMP	WHO (HFA, WFH, BMI)‡	63.2	36.3	p	35.6	63.8		AR: k=−0.005	n
	PMST (STAMP-modified)		94.4	29.0	p	40.5	91.1		AR: k=0.177	s
	PYMS		26.1	67.1	p	34.3	58.0		AR: k=−0.71	n
Wang et al[6]	STAMP	WHO (WHA, HFA, BMI)‡	100	73.3	f	65.2	100		AR: k=0.603	m
White et al[9]	PNST	WHO and CDC 2000								
		BMI ≤ −2SD‡	89.3	66.2	f	22.5	98.4			
		WFA ≤ −2SD‡	89.5	65.0	f	15.3	98.9			
		HFA ≤ −2SD‡	55.6	62.4	f	4.5	97.8			
		SGNA‡	77.8	82.1	f	69.3	87.6			
	SGNA	WHO and CDC 2000								
		BMI ≤ −2SD‡	96.5	72.5	f	27.7	99.5			
		WFA ≤ −2SD‡	85.7	69.7	f	17.8	98.5			
		HFA ≤ −2SD‡	46.2	66.5	p	6.0	96.4			
Wiskin et al[59]	STAMP	ICD-10‡	100†	0†	p	6.52†	NA†		AR: k=−0.014	n
	STRONGkids		100	0†	p	6.52†	NA†		AR: k=−0.013	n
	SPNRS		100†	0†	p	6.52†	NA†		AR: k=−0.013	n
	PYMS		100†	53.49†	f	13.04†	100†		AR: k=0.079	s
	STAMP	STRONGkids‡	100†	NA†	p	100†	NA†		STRONGkids: k=0.774	su
	STRONGkids	SPNRS‡	100†	NA†	p	100†	NA†		SPNRS: k=0.732	su

Continued

Table 3 Continued

Study	Tool	Reference standard	Se (%)	Sp (%)	Rating	PPV (%)	NPV (%)	Inter-/intra observer agreement (kappa)	Agreement with reference standard/ other screening tools	Rating kappa
		PYMS‡	50†	NA†	p	100†	0†		PYMS: k=0.332	f
SPNRS	SPNRS‡		100†	NA†	p	100†	NA†		SPNRS: k=0.600	m
	PYMS‡		50†	NA†	p	100†	0†		PYMS: k=0.270	f
	PYMS‡		50†	NA†	p	100†	0†		PYMS: k=0.236	f
Wong et al[14]	STAMP	Full dietetic assessment‡	83.3	66.7	f	78.1	73.6	Interobserver^g: k=0.752	AR: k=0.507 (95% CI 0.646 to 1.000)	m
		PYMS‡						Intraobserver^h: k=0.635	PYMS: k=0.314 (95% CI 0.076 to 0.552)	f
Wonoputri et al[56]	PYMS	SGNA‡	95.31 (95% CI 0.87 to 0.98)	76.92 (95% CI 63 to 86)	f	83.56 (95% CI 73 to 90)	93.02 (95% CI 81 to 97)		AR^e: k=0.348 (95% CI 0.191 to 0.506)	f
									AR^f: k=0.125 (95% CI 0 to 0.299)	s
	STAMP		100 (95% CI 0.94 to 1)	11.54 (95% CI 5 to 23)	p	58.2 (95% CI 48 to 67)	100 (95% CI 61 to 100)		AR^e: k=0.018 (95% CI 0 to 0.140)	s
									AR^f: k=0 (95% CI 0 to 0.140)	s
	STRONG_kids		100 (95% CI 94 to 100)	7.7 (95% CI 3 to 18)	p	57.14 (95% CI 47.9 to 65.9)	100 (95% CI 51 to 100)		AR^e: k=0.028 (95% CI 0 to 0.149)	s
									AR^f: k=0 (95% CI 0 to 0.144)	s

*low and medium malnutritional risk versus high.
†not reported by authors (ie, calculated in the present study).
‡low versus medium and high malnutritional risk.
§error in reported data.
aamong six registered dietitians on 18 patients.
btwo research dietitians and ward nursing staff.
cnurse-rated malnutrition risk screening tool.
dresearch–dietitian-rated malnutrition risk screening tool.
ereference standard: acute malnutrition (WFH <-2SD).
freference standard: chronic malnutrition (HFA <-2SD).
gregistered dietitian versus nurses.
hregistered dietitian.
ion 93 patients.
jpaediatrician versus nurses.
knurse.
lpaediatrician.
mregistered dietitian/physician specialised in paediatric nutrition versus nurse non-specialised in nutrition or paediatric resident.
AR, agreement with the reference standard used; for other abbreviations and acronyms, see table 1.

RQ2: What is the validity of the screening and assessment tools versus the reference standard?

This systematic review evaluated 26 studies in all. The methodological quality of all of these was considered moderate. The main methodological problems were related to the lack of detailed descriptions of the study protocols, and to the non-blinded interpretation of index tests with regards to the reference standards, and/or vice versa. In particular, the study by Soundarya et al[62] was inadequate in the reporting of information about the study protocol. There was no information about the study flow, the time frame and the number of people involved in the evaluation process. Additionally, the CANS score was used as the reference standard and not as the index test, as would be expected. Apart from the abovementioned study, four additional studies did not report the time frame of the study.[39–41 57]

Direct comparisons of the screening tools in terms of Se, Sp, PPV and NPV are not possible, as different cutoff values were used in the different studies. Also, the three malnutrition risk groups (ie, low, medium and high) were not uniquely combined into two groups for comparisons. Some studies used the combination of low to medium risk compared with high risk, while other studies used low compared with medium to high risk.

STAMP and STRONG$_{kids}$ were the most often validated screening tools in the investigated studies. Anthropometry measurements were used as the reference standard in five studies for validation of STAMP, and in nine studies for validation of STRONG$_{kids}$. Clinical and surgical paediatric patients (n=223) were included in a study using BMI measurements as the reference standard, while comparing low and medium risk groups versus the high risk group.[61] STAMP was validated with a good rating when performed by experts, but rated only fair with non-experts. Similar results were obtained on a larger sample in a multicenter study (n=2567) when validated with BMI and MUAC, but these were poor when validated with triceps skin fold thickness (TSFT).[63] However, for comparisons of low risk versus medium and high risk groups, the validation studies reported poor ratings, with the exception of the study that included only a specifically small group of paediatric patients with SCI (n=45).[6] When compared with full dietetic assessment, STAMP obtained fair ratings in three studies[7 14 32] and poor ratings in the study on 60 outpatients.[57] The modified version of STAMP in the last study here obtained a fair rating. Additionally, when validated with SGNA[56] and ICD-10,[58] the ratings were poor.

STRONG$_{kids}$ obtained fair ratings when compared with anthropometric measurements in only two studies.[53 58] The validation of STRONG$_{kids}$ on the same group of patients as STAMP resulted in lower agreement with the BMI measurements regardless of the expertise of the assessor.[61] Similar conclusions can be drawn from the larger multicenter study on 2567 patients.[63]

When validated with SGNA, STRONG$_{kids}$ obtained the highest Se (100%); however, the Sp was very low (7.7%).[56] Unfortunately, the results of the validation with anthropometry were not reported. The authors report only the prevalence of malnutrition based on the WHO criteria. Although the number of patients included was not so small (n=116), the authors reported a high percentage of children with oncological disorders (43.1%) and infectious diseases (14.6%), which resulted in a higher percentage of positive samples (28.4%). In the same study, PYMS compared with SGNA obtained the best results (Se=95.3%, Sp=76.9%, PPV=83.6%, NPV=93.0%). These results deviated quite a lot compared with STAMP and STRONG$_{kids}$ in the same patient group. However, no conclusions can be drawn here as the results were reported only when compared with other screening/assessment tools, and not using an anthropometry measurement as the reference standard.[56] Similar results were found in the study with only 46 patients with IBD.[59]

On the contrary, the CANS score was validated using anthropometry measurements (Roher's Ponderal Index) on the largest group of newborns (n=14477), and it gave a Sp of >90% and a Se of <50%.[64] The CANS score showed a better performance compared with Roher's Ponderal Index and BMI on a much smaller sample of 300 newborns.[62] The authors of both studies concluded that the CANS score is useful for the identification of fetal malnutrition in newborns; however, the results from the larger sample did not confirm these statements. The second malnutrition screening tool for newborns, NNST, was also validated on a relatively large group of patients (n=909) using anthropometry measurements, where the UK growth charts obtained fair results (Se=89.6%, Sp=75.1%).[42] However, the NNST was designed for a specific group of patients: neonatal patients in ICU who were >2 weeks old. The results showed good performance, although the low PPV indicates that two-thirds of the patients were unnecessarily wrongly predicted as at risk of malnutrition.

The PNST was validated in only one study with anthropometric measures (BMI ≤−2SD; Se=89.3%, Sp=66.2%).[9] However, the SGNA was also validated with anthropometric measures (BMI≤−2SD) in the same study on the same group of patients. These results showed even better agreement (Se=96.5%, Sp=72.5%) although the very low PPVs indicated high overprediction of positive cases. The SGNA was also used as the reference standard for validation of PNST, which obtained a fair performance.

The PYMS screening tool was validated with anthropometric measurements in three studies,[54 63 65] and it obtained a good rating compared with BMI in a multi-country study that included 2567 paediatric patients,[63] with a poor rating in the other two smaller studies (n=300, n=119). When compared with STAMP and STRONG$_{kids}$, PYMS obtained the best results in the validation with BMI, for Se and Sp. Very low PPVs can be a source of concern, which was indeed noted by the authors of the tool in their first validation study using full dietetic assessments as the reference standard.[7]

When the PYMS, STAMP, STRONG$_{kids}$ and PNRS were validated with ICD-10 as the reference standard, they

showed poor agreement.[59] This study included a small (n=46) and very heterogeneous group of children with IBD, and it did not reach any conclusions about the reasonability of the use of these tools for children with IBD.

The PYMS was also validated with the SGNA, which obtained a fair rating regarding the Se and Sp.[56] High PPVs and NPVs also indicated that overprediction is not a cause for concern. As reported by the authors, PYMS obtained the best agreement with SGNA for the detection of acute malnutrition (k=0.3).

Three malnutrition screening tools specially designed for children with specific chronic illnesses were identified in this study[39–41] (SCAN, Nutrition Risk screening tool in CF, Nutrition screening tool for paediatric patients with CF). None of these tools were validated with anthropometry measurements. SCAN was validated with SGNA,[39] and obtained a poor rating due to weak identification of negative cases (Sp=39%); however, identification of positive cases was 100%. The PPV showed that >40% of the identified children were actually not exposed to the risk. The Nutrition Risk screening tool in CF was validated using full dietetic/nutritional assessment and the CFF consensus report criteria as the reference standard, which obtained a fair rating with high PPVs and NPVs, thus indicating a good performance.[41] The second tool for patients with CF (ie, Nutrition screening tool for paediatric patients with cystic fibrosis) was also validated with the CFF consensus report criteria and obtained fair agreement, but with lower PPVs and NPVs, which thus indicated more problems with the overprediction of positive cases (>40%).[40] Good agreement was reported when this tool was validated with the Nutrition Risk screening tool in CF (k=0.8).

Comparisons with similar systematic reviews

Similar reviews of validation studies on paediatric malnutrition screening tools can be found in the literature.[8 66 67] However, the present systematic review includes the greatest number of paediatric malnutrition screening/assessment tools identified, and also in terms of the validation studies. The systematic reviews by Moeeni and Day[68] and Hartman *et al*[69] only focused on the description of the available paediatric malnutrition screening tools (six and five tools, respectively). Six paediatric malnutrition screening tools were described in the study by Joosten and Hulst,[8] with eight validation studies included. The authors defined two tools as the most practical and reliable: STRONG$_{kids}$ and PYMS. They proposed that STRONG$_{kids}$ is the most reliable for assessing nutritional risk, and PYMS for assessing nutritional risk and actual nutritional status. A systematic review of studies that validated malnutrition screening tools for hospitalised children and included a meta-analysis was published by Huysentruyt *et al*.[66] The systematic review included four malnutrition screening tools (PYMS, STRONG$_{kids}$, STAMP and PNRS) and 15 validation studies. As also observed in the present study, the authors were confronted with several problems

when comparing the validation results from several studies. This was in particular due to the heterogeneity of the reference standards, the different cutoff points used and the small sample sizes (only one study had more than 100 participants). Their conclusions demonstrated that at the time there was insufficient evidence to choose one screening tool over another. The most recent systematic review was conducted by Teixeira and Viana,[67] and this included five malnutritional screening/assessment tools (PYMS, STRONG$_{kids}$, STAMP, PNST and SGNA). The authors concluded that STRONG$_{kids}$ and STAMP showed the best clinical performances in the studies included.

The results of the present systematic review are not consistent with the conclusions of Josten and Hulst[8] and Teixseira and Viana,[67] who recommended STRONG$_{kids}$ as the most reliable screening tool; here, PYMS showed better performance (table 3).

Strengths and limitations

There are some key limitations to the present study that have to be emphasised. The most important limitation comes from the lack of a gold standard for evaluation of the malnutrition risk of hospitalised children. Consequently, the studies used different reference standards, most often as anthropometric measures or dietetic assessments. As reported in a number of studies,[70–72] the accuracy of the anthropometric measurements was often poor, which resulted in questionable uniformity of the validation of these screening/assessment tools. Full dietetic/nutritional assessment also varied due to the different methods used and the different educational standards for dietetics in different countries.[3] Some studies used another screening/assessment tool for the reference standard, which introduces a certain source of bias. As observed in the present study, the PPVs in the validation studies that compared one screening tool to another were higher than for the same screening tool compared with the anthropometry measures. However, this can be expected if both of the screening tools overpredict positive cases when validated with the same anthropometry measurements as reference standard. The choice of the reference standard, therefore, represents a source of bias to the original studies, and consequently also to the present systematic review.

Another limitation comes from the different inclusion and exclusion criteria used in the studies, and the lower power of the studies with small sample sizes. The protocols of the studies were not uniquely defined, so the studies differed in the type and number of assessors using the nutritional screening tool studied. Similarly, different types and numbers of assessors were involved in evaluation of the malnutrition risk, according to the study-defined reference standard. Also, almost one-third of the studies evaluated (30.8%)[6 14 39–41 52 57 59] had a sample size of <100 patients. These small numbers of paediatric patients involved mandate caution when generalising the results to the full population.

However, one of the main strengths of the present systematic review is that we focused only on studies that included validation of the screening/assessment tools used based on the chosen reference standard. We are confident that all of the currently available nutritional screening/assessment tools for paediatric patients are included in this review. This study thus presents a complete review of the success of these tools in the prediction of malnutrition risk.

Recommendations

The results of this study show the need for the definition of a gold standard. We propose that an expert group is formed to discuss and define which reference standard should be used worldwide for the evaluation of screening tools. Based on the deficiencies identified in validation studies during the present systematic review, we recommend that the reference standard should never be another screening/assessment tool. We propose that the anthropometric measurements defined in the WHO growth charts/Anthro/AnthroPlus software or the CDC 2000 growth charts/Epi Info 7 should be considered as the basic reference standard for the purpose of fair comparisons. The American Society for Parenteral and Enteral Nutrition (ASPEN)[73] recommendations included the recording of weight, height, BMI and MUAC, and considered the TSFT and mid-arm muscle circumference on admission, with reference to the appropriate growth chart. Head circumference must also be obtained in infants younger than <2 years. The CDC[74] and ASPEN[73] recommend using the WHO charts for children up to 2 years of age. For children and adolescents from 2 to 20 years of age, they recommend using the CDC 2000 charts. The newest versions of growth charts should also be used. Anthropometric measurements should be performed using calibrated equipment, according to the examination protocols described by the CDC.[75] As validation of malnutrition screening tools using anthropometry measurements as the reference standard tend to produce a lot of false-positive results, our recommendation is to use full dietetic/nutritional assessment in the second stage only for the positively identified cases. However, the dietetic/nutritional assessment process and evaluation should also be standardised first. Furthermore, specific disease conditions can cause energy and/or protein imbalances, and therefore these should be considered in the final evaluation of the nutritional status.[73] Validation studies should also test several malnutrition screening tools on the same population, to avoid bias due to different patient populations, disease backgrounds or age groups.[31]

It is important that healthcare professionals who perform nutritional screening are appropriately educated and trained in the measurement of the anthropometric parameters, and they should use the appropriate growth charts or computer software, and the chosen screening tool.

The study protocol should be carefully designed and followed, with particular attention paid to the following:

► Patient selection: The sample of patients included in studies should be as large as possible, with consecutive or random sampling used, and inappropriate exclusions avoided. The inclusion and exclusion criteria should be carefully defined. All of the patients included in a study should be included in the validation procedures.

► Flow and timing: To obtain an evaluation of the nutritional status that is as objective as possible, the evaluator using the screening test and the reference standard should not be the same person. Additionally, at the time of performing the evaluation of the nutritional status, the evaluators should not be acquainted with the results of other evaluations. The exact sequence of the performing of the screening tests and the reference standards has to be defined in the study protocol, while taking into consideration that both the screening test and the reference standard should be evaluated on admission[73] and on the same day (or as close as possible). This will avoid possible changes in the patient health condition in between these evaluations. The exact flow and timing of these should also be reported.

► Reporting results: All of the traditional evaluation metrics should be reported, such as Se, Sp, PPV and NPV. When the malnutrition risk is evaluated using three categories (ie, low/moderate/high), the moderate and high risks should both be treated as 'at risk'. A table showing the cross-classification of the malnutrition risk on the screening test compared with the reference standard should also be reported. When several evaluators use the screening test or the reference standard, the inter-rater or intra-rater agreement should also be reported.

CONCLUSIONS

This systematic review shows that several paediatric nutritional screening/assessment tools have been developed; however, due to the lack of a gold standard, it is very difficult (if not impossible) to compare them at present. The validation results show that nutritional screening tools perform better when designed for specific groups of patients who suffer from chronic or specific conditions. An exception is seen for SCAN, which is designed for oncological patients. The only validation study of SCAN that was found for inclusion in the present systematic review used SGNA as the reference standard; therefore, additional validation studies are needed for correct validation here. Low PPVs were seen for almost all of the studies that used anthropometry as the reference standard, which indicates the problems associated with overprediction of positive cases. It is true that it is better to include more false positives than false negatives, but this also leads to unnecessary exposure of the children to more invasive assessments, an increased workload for the health staff and an additional financial burden. However, very low PPVs should be treated with caution.

It is particularly difficult to recommend any one screening/assessment tool on the basis of the results of all of these published studies, due to their heterogeneity. However, PYMS appears to perform better than STAMP and STRONG$_{kids}$ when compared with anthropometric measurements, especially in terms of BMI and TSFT. Therefore, we would recommend the use of PYMS in the hospital setting for paediatric patients without chronic conditions. Due to its tendency to overpredict positive cases, we also recommend the use of full dietetic/nutritional assessments in the second stage for the positively identified cases.

For fair comparisons here, there is the need for more studies that are aimed at the validation of different screening/assessment tools for the same group of patients using the same reference standard. We also recommend that a unified standard for full nutritional assessment should be developed, and that this should then be used in combination with the cited growth charts.

Thus, we recommend further studies to validate nutritional screening/assessment tools with the aim being to provide health experts with fair comparisons, and consequently easier decisions, in terms of which tool(s) to use.

REFERENCES

1. Campanozzi A, Russo M, Catucci A, *et al*. Hospital-acquired malnutrition in children with mild clinical conditions. *Nutrition* 2009;25:540–7.
2. Hulst JM, Zwart H, Hop WC, *et al*. Dutch national survey to test the STRONG$_{kids}$ nutritional risk screening tool in hospitalized children. *Clin Nutr* 2010;29:106–11.
3. Joosten KF, Hulst JM. Prevalence of malnutrition in pediatric hospital patients. *Curr Opin Pediatr* 2008;20:590–6.
4. Okoromah CA, Ekure EN, Lesi FE, *et al*. Prevalence, profile and predictors of malnutrition in children with congenital heart defects: a case-control observational study. *Arch Dis Child* 2011;96:354–60.
5. Pawellek I, Dokoupil K, Koletzko B. Prevalence of malnutrition in paediatric hospital patients. *Clin Nutr* 2008;27:72–6.
6. Wang YJ, Zhou HJ, Liu PJ, *et al*. Risks of undernutrition and malnutrition in hospitalized pediatric patients with spinal cord injury. *Spinal Cord* 2017;55:247–54.
7. Gerasimidis K, Keane O, Macleod I, *et al*. A four-stage evaluation of the Paediatric Yorkhill Malnutrition Score in a tertiary paediatric hospital and a district general hospital. *Br J Nutr* 2010;104:751–6.
8. Joosten KF, Hulst JM. Nutritional screening tools for hospitalized children: methodological considerations. *Clin Nutr* 2014;33:1–5.
9. White M, Lawson K, Ramsey R, *et al*. Simple nutrition screening tool for pediatric inpatients. *JPEN J Parenter Enteral Nutr* 2016;40:392–8.
10. Elia M, Stratton RJ. Considerations for screening tool selection and role of predictive and concurrent validity. *Curr Opin Clin Nutr Metab Care* 2011;14:425–33.
11. Elia M, Stratton RJ. An analytic appraisal of nutrition screening tools supported by original data with particular reference to age. *Nutrition* 2012;28:477–94.
12. Erkan T. Methods to evaluate the nutrition risk in hospitalized patients. *Turk Pediatri Ars* 2014;49:276–81.
13. Secker DJ, Jeejeebhoy KN. Subjective global nutritional assessment for children. *Am J Clin Nutr* 2007;85:1083–9.
14. Wong S, Graham A, Hirani SP, *et al*. Validation of the Screening Tool for the Assessment of Malnutrition in Paediatrics (STAMP) in patients with spinal cord injuries (SCIs). *Spinal Cord* 2013;51:424–9.
15. WHO. Growth reference data for 5-19 years. 2007 http://www.who.int/growthref/en/ (Accessed 4 Jan 2018).
16. WHO. The WHO Child Growth Standards, Documentation. 2016 http://www.who.int/childgrowth/standards/en/ (Accessed 4 Jan 2018).
17. Centers for Disease Control and Prevention. CDC Growth Charts. 2016 https://www.cdc.gov/growthcharts/cdc_charts.htm (Accessed 2 Feb 2018).
18. Kuczmarski RJ, Ogden CL, Guo SS, *et al*. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11* 2002;11:1–190.
19. de Onis M, Onyango AW, Borghi E, *et al*. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007;85:660–7.
20. Centers for Disease Control and Prevention. WHO Growth Standards Are Recommended for Use in the U.S. for Infants and Children 0 to 2 Years of Age. 2010 https://www.cdc.gov/growthcharts/who_charts.htm (Accessed 4 Jan 2018).
21. de Onis M. Update on the implementation of the WHO child growth standards. *World Rev Nutr Diet* 2013;106:75–82.
22. Zong XN, Li H. Construction of a new growth references for China based on urban chinese children: Comparison with the WHO growth standards. *PLoS One* 2013;8:e59569.
23. Aburawi EH, Nagelkerke N, Deeb A, *et al*. National growth charts for United Arab Emirates children with Down syndrome from birth to 15 years of age. *J Epidemiol* 2015;25:20–9.
24. The British Dietetic Association. *Model and Process for Nutrition and Dietetic Practice*. Birmingham: The British Dietetic Association, 2016. https://www.bda.uk.com/publications/professional/model_and_process_for_nutrition_and_dietetic_practice_. (Accessed 23 Dec 2018).
25. WHO. *Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. WHO Technical Report Series No. 854.* : World Health Organ Tech Rep Ser, 1995:854: 1–452. http://apps.who.int/iris/bitstream/10665/37003/1/WHO_TRS_854.pdf. (Accessed 08 Mar 2018).
26. Cogill B. Anthropometric Indicators Measurement Guide. 2003 https://www.fantaproject.org/sites/default/files/resources/anthropometry-2003-ENG.pdf (Accessed 22 Apr 2018).
27. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8:336–41.
28. *Review Manager (RevMan) [Computer program]. Version 5.3.* Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014. https://community.cochrane.org/help/tools-and-software/revman-5. (Accessed 1 Oct 2017).
29. 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.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
31. van Bokhorst-de van der Schueren MA, Guaitoli PR, Jansma EP, *et al*. Nutrition screening tools: does one size fit all? A systematic review of screening tools for the hospital setting. *Clin Nutr* 2014;33:39–58.
32. McCarthy H, Dixon M, Crabtree I, *et al*. The development and evaluation of the Screening Tool for the Assessment of Malnutrition in Paediatrics (STAMP©) for use by healthcare staff. *J Hum Nutr Diet* 2012;25:311–8.
33. McCarthy H, McNulty H, Dixon M, *et al*. Screening for nutrition risk in children: the validation of a new tool. *Journal of Human Nutrition and Dietetics* 2008;21:395–6.
34. Gerasimidis K, Macleod I, Maclean A, *et al*. Performance of the novel Paediatric Yorkhill Malnutrition Score (PYMS) in hospital practice. *Clin Nutr* 2011;30:430–5.

35. Baker JP, Detsky AS, Wesson DE, *et al*. Nutritional assessment: a comparison of clinical judgement and objective measurements. *N Engl J Med* 1982;306:969–72.

36. Detsky AS, Baker JP, Mendelson RA, *et al*. Evaluating the accuracy of nutritional assessment techniques applied to hospitalized patients: methodology and comparisons. *JPEN J Parenter Enteral Nutr* 1984;8:153–9.

37. Detsky AS, McLaughlin JR, Baker JP, *et al*. What is subjective global assessment of nutritional status?. *JPEN J Parenter Enteral Nutr* 1987;11:8–13.

38. Mahdavi AM, Ostadrahimi A, Safaiyan A. Subjective global assessment of nutritional status in children. *Matern Child Nutr* 2010;6:374–81.

39. Murphy AJ, White M, Viani K, *et al*. Evaluation of the nutrition screening tool for childhood cancer (SCAN). *Clin Nutr* 2016;35:219–24.

40. Souza Dos Santos Simon MI, Forte GC, da Silva Pereira J, *et al*. Validation of a Nutrition Screening Tool for Pediatric Patients with Cystic Fibrosis. *J Acad Nutr Diet* 2016;116:813–8.

41. McDonald CM. Validation of a nutrition risk screening tool for children and adolescents with cystic fibrosis ages 2–20 years. *J Pediatr Gastroenterol Nutr* 2008;6:438–46.

42. Johnson MJ, Pearson F, Emm A, *et al*. Developing a new screening tool for nutritional risk in neonatal intensive care. *Acta Paediatr* 2015;104:e90–e93.

43. Metcoff J. Clinical assessment of nutritional status at birth. Fetal malnutrition and SGA are not synonymous. *Pediatr Clin North Am* 1994;41:875–91.

44. Reilly HM, Martineau JK, Moran A, *et al*. Nutritional screening-evaluation and implementation of a simple Nutrition Risk Score. *Clin Nutr* 1995;14:269–73.

45. Sermet-Gaudelus I, Poisson-Salomon AS, Colomb V, *et al*. Simple pediatric nutritional risk score to identify children at risk of malnutrition. *Am J Clin Nutr* 2000;72:64–70.

46. Apostolou A, Printza N, Karagiozoglou-Lampoudi T, *et al*. Nutrition assessment of children with advanced stages of chronic kidney disease-A single center study. *Hippokratia* 2014;18:212–6.

47. Karagiozoglou-Lampoudi T, Daskalou E, Lampoudis D, *et al*. Computer-based malnutrition risk calculation may enhance the ability to identify pediatric patients at malnutrition-related risk for unfavorable outcome. *JPEN J Parenter Enteral Nutr* 2015;39:418–25.

48. Rowell A, Long C, Chance L, *et al*. Identification of nutritional risk by nursing staff in secure psychiatric settings: reliability and validity of St Andrew's Nutrition Screening Instrument. *J Psychiatr Ment Health Nurs* 2012;19:722–8.

49. Hasegawa J, Ito YM, Yamauchi T. Development of a screening tool to predict malnutrition among children under two years old in Zambia. *Glob Health Action* 2017;10:1339981.

50. Gilliam J, Laney SO. Nutrition screening for infants and young children with special health care needs: Spoken country, Washington. 2008 http://www.doh.wa.gov/Portals/1/Documents/Pubs/970-116_NutritionScreeningForInfantsAndYoungCSHCN.pdf (Accessed 25 Jan 2018).

51. Huysentruyt K, Alliet P, Muyshont L, *et al*. The STRONG(kids) nutritional screening tool in hospitalized children: a validation study. *Nutrition* 2013;29:1356–61.

52. Ling RE, Hedges V, Sullivan PB. Nutritional risk in hospitalised children: An assessment of two instruments. *E Spen Eur E J Clin Nutr Metab* 2011;6:e153–e157.

53. Mărginean O, Pitea AM, Voidăzan S, *et al*. Prevalence and assessment of malnutrition risk among hospitalized children in Romania. *J Health Popul Nutr* 2014;32:97–102.

54. Moeeni V, Walls T, Day AS. Assessment of nutritional status and nutritional risk in hospitalized Iranian children. *Acta Paediatr* 2012;101:e446–e451.

55. Moeeni V, Walls T, Day AS. The STRONG(kids) nutritional risk screening tool can be used by paediatric nurses to identify hospitalised children at risk. *Acta Paediatr* 2014;103:e528–e531.

56. Wonoputri N, Djais JT, Rosalina I. Validity of nutritional screening tools for hospitalized children. *J Nutr Metab* 2014;2014:1–6.

57. Rub G, Marderfeld L, Poraz I, *et al*. Validation of a nutritional screening tool for ambulatory use in pediatrics. *J Pediatr Gastroenterol Nutr* 2016;62:771–5.

58. Spagnuolo MI, Liguoro I, Chiatto F, *et al*. Application of a score system to evaluate the risk of malnutrition in a multiple hospital setting. *Ital J Pediatr* 2013;39:81.

59. Wiskin AE, Owens DR, Cornelius VR, *et al*. Paediatric nutrition risk scores in clinical practice: children with inflammatory bowel disease. *J Hum Nutr Diet* 2012;25:319–22.

60. Durakbaşa ÇU, Fettahoğlu S, Bayar A, *et al*. The prevalence of malnutrition and effectiveness of strongkids tool in the identification of malnutrition risks among pediatric surgical patients. *Balkan Med J* 2014;31:313–21.

61. Galera-Martínez R, Moráis-López A, Rivero de la Rosa MD, *et al*. Reproducibility and inter-rater reliability of 2 paediatric nutritional screening tools. *J Pediatr Gastroenterol Nutr* 2017;64:e65–e70.

62. Soundarya M, Basavaprabhu A, Raghuveera K, *et al*. Comparative assessment of fetal malnutrition by anthropometry and can score. *Iran J Pediatr* 2012;22:70–6.

63. Chourdakis M, Hecht C, Gerasimidis K, *et al*. Malnutrition risk in hospitalized children: use of 3 screening tools in a large European population. *Am J Clin Nutr* 2016;103:1301–10.

64. Martínez-Nadal S, Demestre X, Raspall F, *et al*. [Assessment of foetal nutrition status at birth using the CANS score]. *An Pediatr* 2016;84:218–23.

65. Thomas PC, Marino LV, Williams SA, *et al*. Outcome of nutritional screening in the acute paediatric setting. *Arch Dis Child* 2016;101:1119–24.

66. Huysentruyt K, Devreker T, Dejonckheere J, *et al*. Accuracy of nutritional screening tools in assessing the risk of undernutrition in hospitalized children. *J Pediatr Gastroenterol Nutr* 2015;61:159–66.

67. Teixeira AF, Viana KD. Nutritional screening in hospitalized pediatric patients: a systematic review. *J Pediatr* 2016;92:343–52.

68. Moeeni V, Day AS. Nutritional risk screening tools in hospitalised children. *Int J Child Health Nutr* 2012;1:39–43.

69. Hartman C, Shamir R, Hecht C, *et al*. Malnutrition screening tools for hospitalized children. *Curr Opin Clin Nutr Metab Care* 2012;15:303–9.

70. Geeta A, Jamaiyah H, Safiza MN, *et al*. Reliability, technical error of measurements and validity of instruments for nutritional status assessment of adults in Malaysia. *Singapore Med J* 2009;50:1013–8.

71. Moon RJ, Wilson P, Kirkham FJ, *et al*. Growth monitoring following traumatic brain injury. *Arch Dis Child* 2009;94:699–701.

72. Sullivan PB. Malnutrition in hospitalised children. *Arch Dis Child* 2010;95:489–90.

73. Mehta NM, Corkins MR, Lyman B, *et al*. Defining pediatric malnutrition: a paradigm shift toward etiology-related definitions. *JPEN J Parenter Enteral Nutr* 2013;37:460–81.

74. Grummer-Strawn LM, Reinold CM, Krebs NF, *et al*. Use of world health organization and cdc growth charts for children aged 0-59 months in the united states. In: Department of health and human services, centers for disease control and prevention. 2010. Centers for Disease Control and Prevention (2012) https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5909a1.htm (Accessed 5 Feb 2018).

75. *National Health and Nutrition Examination Survey (NHANES) anthropometry procedures manual*. USA: CDC, 2009. https://www.cdc.gov/nchs/data/nhanes/nhanes_09_10/lab.pdf. (Accessed 5 Feb 2018).