BJS, 2025, znaf013
https://doi.org/10.1093/bjs/znaf013
Original Article

Multicentre prospective study on the diagnostic and prognostic validity of malnutrition assessment tools in surgery

Georgia Petra[1], Evangelos I. Kritsotakis[2], Nikolaos Gouvas[3], Dimitrios Schizas[4], Konstantinos Toutouzas[5], Michael Karanikas[6], George Pappas-Gogos[7], Georgios Stylianidis[8], George Zacharioudakis[9], Aggelos Laliotis[10], Grigorios Christodoulidis[11], Ioannis Kehagias[12] and Konstantinos Lasithiotakis[1,*] on behalf of the MATS Study Group

[1]Department of General Surgery, University Hospital of Heraklion, Medical School, University of Crete, Heraklion, Crete, Greece
[2]Laboratory of Biostatistics, Medical School, University of Crete, Heraklion, Crete, Greece
[3]Department of Surgery, University of Cyprus, School of Medicine, Nicosia, Cyprus
[4]1st Department of Surgery, Laikon Hospital, National and Kapodistrian University of Athens, Athens, Greece
[5]1st Propaedeutic Department of Surgery, Hippokrateio Hospital, National and Kapodistrian University of Athens, Athens, Greece
[6]1st Surgical Department, Faculty of Medicine, School of Health Sciences, Democritus University of Thrace, Alexandroupolis, Greece
[7]2nd Surgical Department, Faculty of Medicine, School of Health Sciences, Democritus University of Thrace, Alexandroupolis, Greece
[8]2nd Department of Surgery, Evangelismos General Hospital, Athens, Greece
[9]5th Surgical Department of Aristotle University of Thessaloniki, Ippokrateio General Hospital of Thessaloniki, Thessaloniki, Greece
[10]Department of Surgery, Venizeleio Hospital, Heraklion, Crete, Greece
[11]Department of Surgery, School of Medicine, University of Thessaly, Larissa, Greece
[12]Department of Surgery, School of Medicine, University of Patras, Patras, Greece

*Correspondence to: Konstantinos Lasithiotakis, Department of General Surgery, University Hospital of Heraklion, School of Medicine, University of Crete, Voutes, Heraklion, Crete, 71110, Greece (e-mail: k.lasithiotakis@uoc.gr)

Members of the MATS Study Group are co-authors of this study and are listed under the heading Collaborators.

Abstract

Background: Malnutrition is a risk factor for postoperative morbidity but the optimal tool for the assessment of malnutrition is unclear.

Methods: This is a prospective multicentre cohort study. Consecutive patients undergoing elective or emergency major abdominal surgery for benign or malignant disease in 12 Greek hospitals between January 2022 and December 2023 were included. Patients unable to provide nutrition history and/or informed consent were excluded. Subjective global assessment (SGA) was used as a reference standard for malnutrition diagnosis. GLIM (global leadership initiative on malnutrition), MNA-SF (mini nutrition assessment short form), MST (malnutrition screening tool), MUST (malnutrition universal screening tool), NRI (nutritional risk index), NRS-2002 (nutrition risk scale 2002), PONS (perioperative nutrition screen) and SNAQ (short nutrition assessment questionnaire) tools were applied for malnutrition risk assessments. Indicators of diagnostic accuracy (sensitivity, specificity, diagnostic odds ratio, areas under the receiver operating characteristic curve—AUC), construct validity (convergent associations with relevant variables) and prognostic validity (logistic regression) were appraised.

Results: 1649 patients were included (58% colorectal, 21% upper gastrointestinal, 14% hepatobiliary operations). SGA defined 562 (34.1%) patients as malnourished with excellent construct and prognostic validity. Malnutrition risk assessments varied from 24.0% using NRS-2002 to 58.6% with the MNA-SF. On their ordinal scales, MNA-SF (AUC = 0.83, 95% c.i. 0.81 to 0.85) and MUST (AUC = 0.79, 95% c.i. 0.77 to 0.82) had the best discriminatory abilities with minimal between-centre heterogeneity. As binary classifiers, MNA-SF (OR = 30.2; 95% c.i. 20.2 to 45.1) and MUST (OR = 16.1; 95% c.i. 12.4 to 21.1) had the highest diagnostic ORs but only MUST had sensitivity and specificity close to 80%. MUST performed well in construct and prognostic validity appraisals.

Conclusion: This study supports the use of the MUST as it is the most valid nutritional screening tool in patients after major abdominal surgery.

Introduction

Malnutrition is a deficiency of nutrient intake, assimilation and metabolism leading to altered body composition and organ dysfunction[1,2]. Malnutrition is common in general surgery, affecting almost two of every three patients who present with oncologic or gastrointestinal disease[3,4]. It is frequently the result of poor oral intake due to dysfunction of the gastrointestinal tract in combination with the acute and chronic inflammation arising from the disease and surgical trauma. Malnutrition can affect every organ system. Muscle function declines even after short periods of starvation, even before any measurable alteration in body composition. This leads to decreased functional capacity and cardiorespiratory reserves[5]. Prolonged malnutrition affects pancreatic function, intestinal blood flow and permeability with negative effects on immune function and wound healing.

Received: October 05, 2024. **Revised:** December 06, 2024. **Accepted:** January 05, 2025

Consequently, malnourished patients have higher rates of postoperative complications (pneumonia, wound infections and anastomotic dehiscence) and mortality compared to well-nourished patients[3,4,6,7].

Malnutrition and its associated risks are reversible. Perioperative nutritional supplementation decreased postoperative complications in patients undergoing gastrointestinal cancer surgery and elective surgery[8,9]. Multimodal prehabilitation programmes comprising nutritional conditioning, physical exercise and psychological support[10] reduce severe complications and length of hospital stay in high-risk and frail patients undergoing major abdominal surgery. Hence, integrating nutritional interventions into the management of surgical patients is recommended by international guidelines[11].

Despite the availability of several nutritional screening instruments, none has been universally accepted for preoperative screening of malnutrition[11]. Most tools were developed and validated in a cohort of surgical and non-surgical hospitalized as well as community patients. The predictive ability of these screening instruments for postoperative outcomes is highly variable[12,13]. An accurate nutritional screening tool may help in optimizing perioperative and nutritional care and as such improve postoperative outcomes.

Methods

Study design

Malnutrition assessment tools in surgery (MATS) was a prospective multicentre, population-based cohort study. Consecutive patients undergoing major abdominal surgery were recruited from 12 hospitals in Greece and Cyprus. Hospitals participated on a rolling basis and enrolled the patients over 8–12 months, between January 2022 and December 2023. Inclusion criteria were age >18 years, elective or emergency surgery, operations involving the gastrointestinal tract, and magnitude of operation classified as major/major+ by POSSUM[14] (*Appendix*). Exclusion criteria were the inability to obtain information about nutritional history and lack of informed consent. The patients were followed until hospital discharge and/or 30 days after surgery to document in-hospital mortality and 30-day major complications (classified grade III or greater by the Clavien–Dindo system)[15]. Follow-up appointments at the clinic were provided when required for postoperative assessments.

Sample size

A precision-based calculation determined that 1437 patients were required to ensure, with probability 0.95, that anticipated 80% sensitivity and 80% specificity to detect malnutrition would be estimated with precision ±4% in a 95% confidence interval. The calculation assumed an underlying true malnutrition prevalence of 30% and was based on Student's *t* distribution (rather than the Normal) to incorporate uncertainty about anticipated estimates[16]. Because loss of precision may occur due to intraclass correlation (that is data from the same hospital are more similar than those from different hospitals), the sample size was inflated by 15%, assuming a design effect of 1.15 (intraclass correlation 0.10) and moderately heterogeneous distributions among hospitals[17]. The aim was to recruit 1650 patients in the study.

Nutritional status and malnutrition screening

Nutritional status was assessed with the 7-point subjective global assessment (SGA), a recognized semi-gold standard for malnutrition diagnosis[18–20]. Within 72 h before surgery, trained surgical registrars assessed each patient for five components of medical history (weight loss, dietary intake, gastrointestinal symptoms, functional status, disease affecting nutritional requirements) and three components of physical examination (muscle wasting, fat depletion, nutrition-related oedema). A detailed list of the registered variables and their definitions is shown in the *Appendix*. The data were collected as part of a wide assessment of perioperative clinical and nutritional data so the registrars had no access to summative ratings of nutritional status from any of the tools. Before the study onset, the registrars received training materials on performing nutritional assessments, somatometric techniques and physical activity tests. Group presentations and in-person discussions via emails, telephone and video calls were held to standardize the data collection process. After completion of data collection, each participant received an overall SGA score for nutritional status: A (well nourished), B (mild or moderately malnourished) or C (severely malnourished). In addition, the registrars collected data required for malnutrition risk ratings by the GLIM (global leadership initiative on malnutrition), MNA-SF (mini nutrition assessment short form), MST (malnutrition screening tool), MUST (malnutrition universal screening tool), NRI (nutritional risk index), NRS-2002 (nutrition risk scale 2002), PONS (perioperative nutrition screen) and SNAQ (short nutrition assessment questionnaire) tools (described in *Table S1*). Of note, GLIM is intended for malnutrition diagnosis but was applied as a screening tool in this study, without a validated body composition measuring technique for muscle mass (for example MRI, DXA or CT)[18]. Rather, hand-grip strength (<16 kg females, <27 kg males), assessed by a CAMRY Digital Hand Dynamometer (Camry Scale, USA), was used as an indirect functional criterion for reduced muscle mass in the GLIM phenotypic criteria.

Data analysis

A comprehensive comparative assessment of the nutritional screening tools was pursued by examining measures of concurrent, construct and prognostic validities[18,21]. Concurrent validity was examined using prevalence-independent measures of diagnostic accuracy (sensitivity, specificity and diagnostic odds ratio) for binary classification of malnutrition, employing SGA as the reference standard[18–20]. In keeping with common clinical practice[22], a dichotomous version of each screening tool was examined, with two levels: 'at risk' (combining categories for moderate and high risk) and 'low risk' of malnutrition. Similarly, SGA was classified into 'malnourished' and 'well nourished'. Confidence intervals for sensitivity and specificity were computed using Wilson's score method, whereas confidence intervals for diagnostic odds ratios based on a log-normal distribution[23,24]. A priori, concurrent validity was defined as adequate when both sensitivity and specificity were 80% or higher[18].

In addition, the overall and centre-specific areas under the receiver operating characteristic curve (AUC) were estimated using the original ordinal scale of each screening tool. To examine interhospital heterogeneity, a random effects model was applied, assuming a Normal between-hospital distribution for the logit transformation of the AUC[25]. A Wald-type 95% c.i. was calculated to indicate highly probable values of pooled AUC for each screening tool. A 95% prediction interval was calculated to indicate the heterogeneity expected in the AUC when a screening tool is applied to a new population. AUCs above 0.80 were considered acceptable discriminative ability.

Examination of construct validity was based on the convergent associations of each screening tool with variables it should be associated with, if the tool sufficiently identifies malnutrition[18,21]. Cross-sectional associations between malnutrition risk ratings from each tool and known medical, physical and functional correlates of malnutrition were looked for. The associations were quantified using odds ratios estimated by fixed-effects logistic regression and were converted into Cohen's d standardized effect sizes to aid with interpreting the results[26,27].

Prognostic validity for each malnutrition screening tool was evaluated by examining the strength of its association with hospital mortality rate and the occurrence of serious 30-day postoperative complications. ORs for each outcome were estimated by mixed-effects logistic regression adjusting for known preoperative prognostic factors[28], including sex, age, ASA class, ambulatory status, emergency admission, operative severity, sepsis 48 h before surgery, ascites 30 days before surgery and history of diabetes mellitus. A random intercept term was included to account for between-hospital effects[29]. Prognostic validity was considered substantial when OR > 2[18].

MUST and NRI could not be calculated for 35 (2%) patients, for whom weight loss could not be quantified. There were no missing data for other variables and no imputation method was applied. The analyses involving MUST and NRI were performed on a complete case basis (listwise deletion). Two-sided confidence intervals and P were reported, considering statistical significance when $P < 0.05$. Stata version 18 (Stata Corporation, College Station, Texas, USA) was used for all analyses.

Ethics and reporting

The review boards of participating hospitals approved the study. All patients provided written informed consent to participate. The study was preregistered on clinicaltrials.gov (ID: NCT05393752) and complies with the STARD guidelines for reporting diagnostic accuracy studies[30]. A STARD checklist is included in *Table S2*.

Results

Participants

A flowchart of the study is shown in *Fig. S1*. Demographic and clinical characteristics and postoperative outcomes of 1649 patients are shown in *Table 1*. Mean(s.d.) patient age was

Table 1 Baseline characteristics and postoperative outcomes according to malnutrition score

Preoperative variables	Subjective Global Assessment (SGA)			
	Well-nourished ($N = 1087$)	Moderately nourished ($N = 508$)	Severely malnourished ($N = 54$)	P
Age (years)	64.6 ± 13.9	69.5 ± 12.6	68.4 ± 15.7	<0.001
Female sex	440 (40%)	221 (44%)	28 (52%)	0.163
BMI category (kg/m²)				
<18.5	7 (1%)	25 (5%)	14 (26%)	<0.001
[18.5, 25)	330 (30%)	254 (50%)	31 (57%)	
[25, 30)	499 (46%)	173 (34%)	7 (13%)	
≥30	251 (23%)	56 (11%)	2 (4%)	
ASA class				
I, without systemic disease	271 (25%)	62 (12%)	5 (9%)	<0.001
II, mild systemic disease	523 (48%)	245 (48%)	13 (24%)	
III, severe not life-threatening disease	253 (23%)	162 (32%)	23 (43%)	
IV, severe & life-threatening disease	40 (4%)	37 (7%)	13 (24%)	
V, moribund patient	0 (0%)	2 (0%)	0 (0%)	
Ambulatory status				
Independent	978 (90%)	268 (53%)	9 (17%)	<0.001
Partially dependent	71 (7%)	169 (33%)	16 (30%)	
Fully dependent	38 (3%)	71 (14%)	29 (54%)	
Psychological distress or acute disease	110 (10%)	120 (24%)	25 (46%)	<0.001
Emergency admission	240 (22%)	222 (44%)	43 (80%)	<0.001
Sepsis	99 (9%)	126 (25%)	25 (46%)	<0.001
Ascites	13 (1%)	47 (9%)	23 (43%)	<0.001
Heart disease	359 (33%)	208 (42%)	26 (48%)	0.001
Albumin (g/dl)	4.1 ± 0.6	3.6 ± 0.7	3.1 ± 0.8	<0.001
WBC (number/mm³)	8587 ± 5083	9690 ± 6232	11 732 ± 6337	<0.001
Timed up-and-go test (s)	23 ± 305	43 ± 447	53 ± 83	0.492
Difficulty walking	137 (13%)	233 (46%)	42 (78%)	<0.001
Difficulty climbing stairs	343 (32%)	353 (69%)	51 (94%)	<0.001
Exhaustion	131 (12%)	256 (50%)	45 (83%)	<0.001
Hand grip strength (kg)	28.0 ± 12.3	20.5 ± 13.0	12.2 ± 10.2	<0.001
Surgical specialty*				
Upper GI	232 (21%)	93 (18%)	8 (15%)	0.633
Lower GI	632 (58%)	301 (59%)	38 (70%)	
Hepatopancreaticobiliary	146 (14%)	73 (15%)	5 (9%)	
Other	77 (7%)	41 (8%)	3 (6%)	
Cancer diagnosis	747 (69%)	336 (66%)	40 (74%)	0.023
Type of operation*				
Major	551 (51%)	225 (44%)	22 (41%)	0.030
Major+	536 (49%)	283 (56%)	32 (59%)	
Postoperative outcomes				
Length of stay (days)	10.8 ± 14.1	16.1 ± 32.7	22.0 ± 21.0	<0.001
In-hospital mortality rate	51 (5%)	52 (10%)	28 (52%)	<0.001
30-day serious complication	167 (15%)	123 (24%)	33 (61%)	<0.001

Data are mean ± standard deviation or frequency (column %). P was calculated using Pearson's chi-squared test for differences in proportions and linear regression for differences in means. GI, gastrointestinal tract; WBC, white blood cells. *See also *Appendix S1* for operations included in the study.

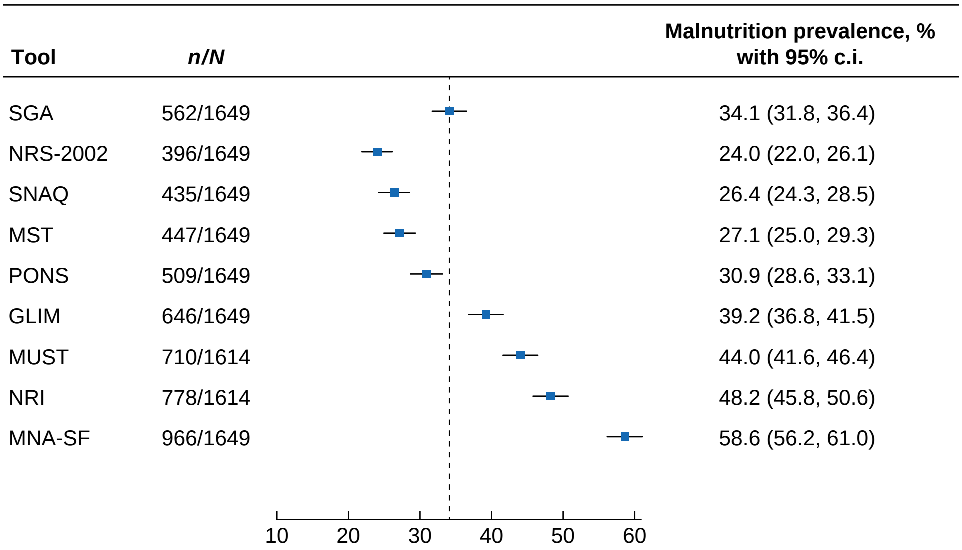


Fig. 1 Forest plot of the variation of malnutrition risk assessments by nutritional screening tools

Note: *n*/*N* is the ratio of the number of patients who were assessed 'at risk of malnutrition' or 'malnourished' to the total number of patients examined.

Table 2 Indicators of concurrent criterion validity

Tool	*N*	Sensitivity, %	Specificity, %	Diagnostic OR
NRS-2002	1649	52.0 (47.7–56.2)	90.4 (88.5–92.1)	10.2 (7.9–13.3)
SNAQ	1649	57.7 (53.4–61.8)	89.8 (87.8–91.5)	12.0 (9.3–15.5)
MST	1649	58.2 (54.0–62.3)	89.0 (86.9–90.8)	11.2 (8.7–14.4)
PONS	1649	66.5 (62.5–70.4)	87.6 (85.5–89.5)	14.0 (10.9–18.0)
GLIM	1649	76.0 (72.2–79.5)	79.9 (77.3–82.2)	12.5 (9.8–16.0)
NRI	1614	76.6 (72.8–80.1)	66.1 (63.2–69.0)	6.4 (5.0–8.1)
MUST	1614	83.6 (80.2–86.6)	76.0 (73.4–78.6)	16.1 (12.4–21.1)
MNA-SF	1649	95.2 (93.1–96.8)	60.3 (57.4–63.3)	30.2 (20.2–45.1)

Data are point estimates with respective 95% c.i. reported in parentheses. The binary form of each screening tool was used to calculate the diagnostic accuracy measure using the Subjective Global Assessment (SGA) as a reference.

66.3 ± 13.8 (range 18–95) years and 58% of the patients were male. Most operations involved the lower gastrointestinal tract (59%) followed by the upper gastrointestinal tract (20%). Of operations, 52% were classified as major and 48% as major+, 31% were emergent, and 68% involved a diagnosis of cancer. The mean(s.d.) length of hospital stay was 12.6(21.7) days, 323 of 1649 (20%) patients had a major complication within 30 days of the operation, and 103 patients (6.2%) died in the hospital.

Nutritional assessment

Mean(s.d.) body weight was 75.8(15.3) kg, mean(s.d.) BMI was 26.5(5.0) kg/m^2 and mean(s.d.) hand grip strength was 25.2(13.1) kg. Using the SGA, 562 of 1649 (34%) patients were defined as malnourished, of whom 54 were severely malnourished. SGA score was associated with most physical, functional and medical characteristics of the study population (construct validity; *Table 1*). Malnutrition risk assessment was highly variable using the different screening tools, ranging from a 24% risk of malnutrition (NRS-2002) to 59% (MNA-SF) (*Fig. 1*).

Concurrent validity

Diagnostic accuracy also varied widely between the screening tools. When considered as binary classifiers, NRS-2002, SNAQ, MST and PONS lacked sensitivity, NRI and MNA-SF lacked specificity, whereas GLIM and MUST appeared better balanced between sensitivity and specificity for detecting SGA-defined malnutrition (*Table 2*). The highest diagnostic ORs were observed for MNA-SF (OR 30.2; 95% c.i. 20.2 to 45.1) and MUST (OR 16.1; 95% c.i. 12.4 to 21.1). However, despite its high sensitivity, MNA-SF had low specificity. By contrast, MUST was the only tool that met the minimal standard for acceptable sensitivity and specificity close to 80%.

When the tools were applied with their originally proposed ordinal scales, high AUC values were obtained for MNA-SF (0.83; 95% c.i. 0.81 to 0.85), GLIM (0.81; 95% c.i. 0.76 to 0.85) and MUST (0.79; 95% c.i. 0.77 to 0.82). However, GLIM exhibited with high heterogeneity across hospitals as seen from a wide prediction interval for the AUC in *Fig. 2*. NRI was homogeneous across hospitals but had relatively low AUC (0.74; 95% c.i. 0.71 to 0.77), whereas NRS-2002, SNAQ, MST, and PONS had low AUCs and high interhospital dispersion. Interhospital heterogeneity was minimal for MUST and MNA-SF. Forest plots of hospital-specific AUCs for each tool are shown in *Figs S2–S9*.

Construct validity

Tables S3–S12 show the association between nutritional screening tools and medical, physical, and functional characteristics. Similarly to SGA, medium-to-large effect sizes were observed for most screening assessments (NRS-2002, PONS, GLIM, MUST, NRI and MNA-SF) concerning physical and functional variables.

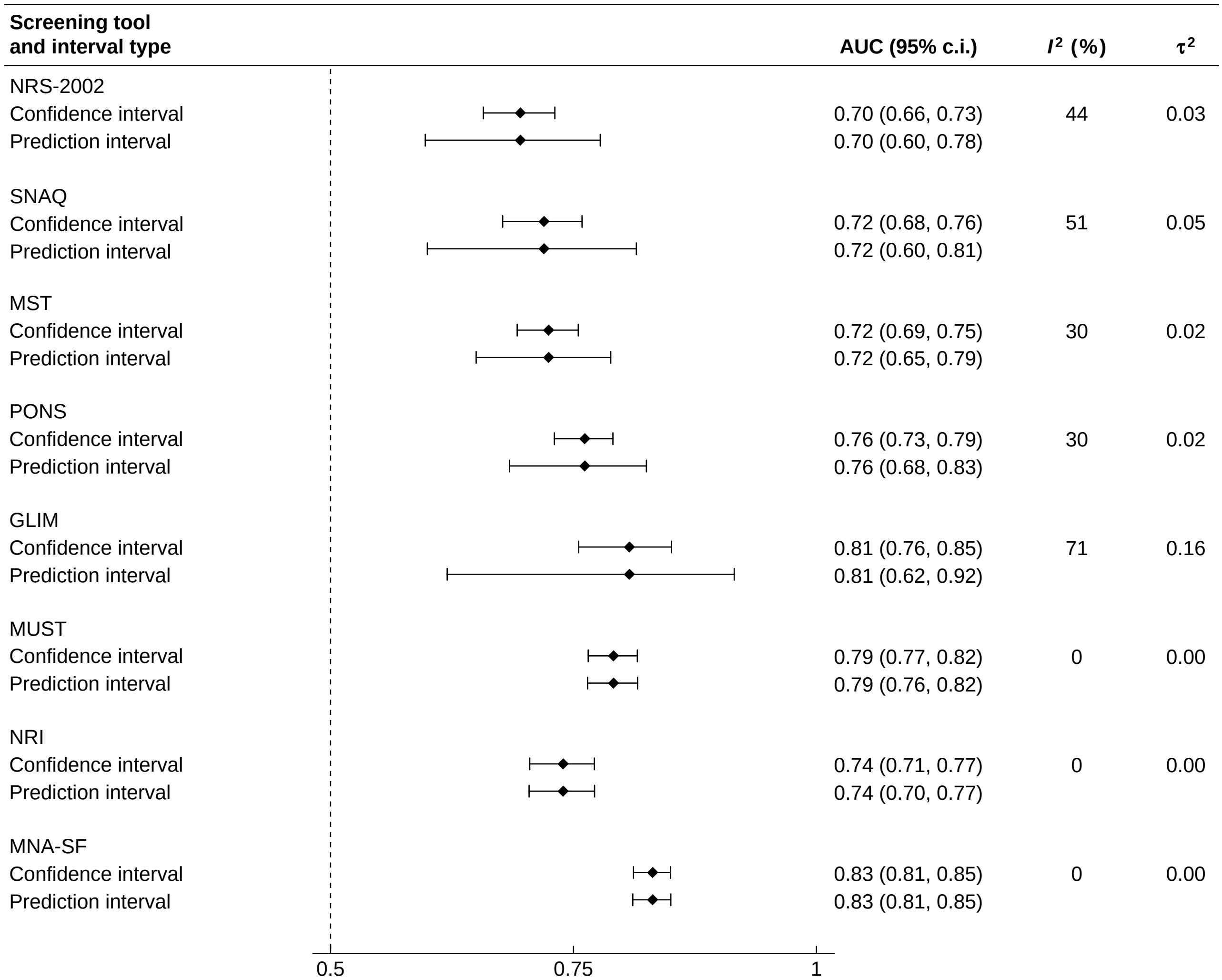


Fig. 2 Forest plot of the pooled area under the receiver operating characteristic curve (AUC) with 95% c.i. and prediction intervals from random-effects meta-analysis of hospital-specific data

Note: Each nutritional screening tool was assessed against the Subjective Global Assessment. The confidence intervals quantify the precision in estimating mean AUCs in this study whereas the prediction intervals quantify the expected dispersion of AUCs in future single-centre studies by accounting for between-hospitals heterogeneity.

Overall, most tools demonstrated good construct validity, but SNAQ and MST performed worst, having the weakest associations with known correlates of malnutrition.

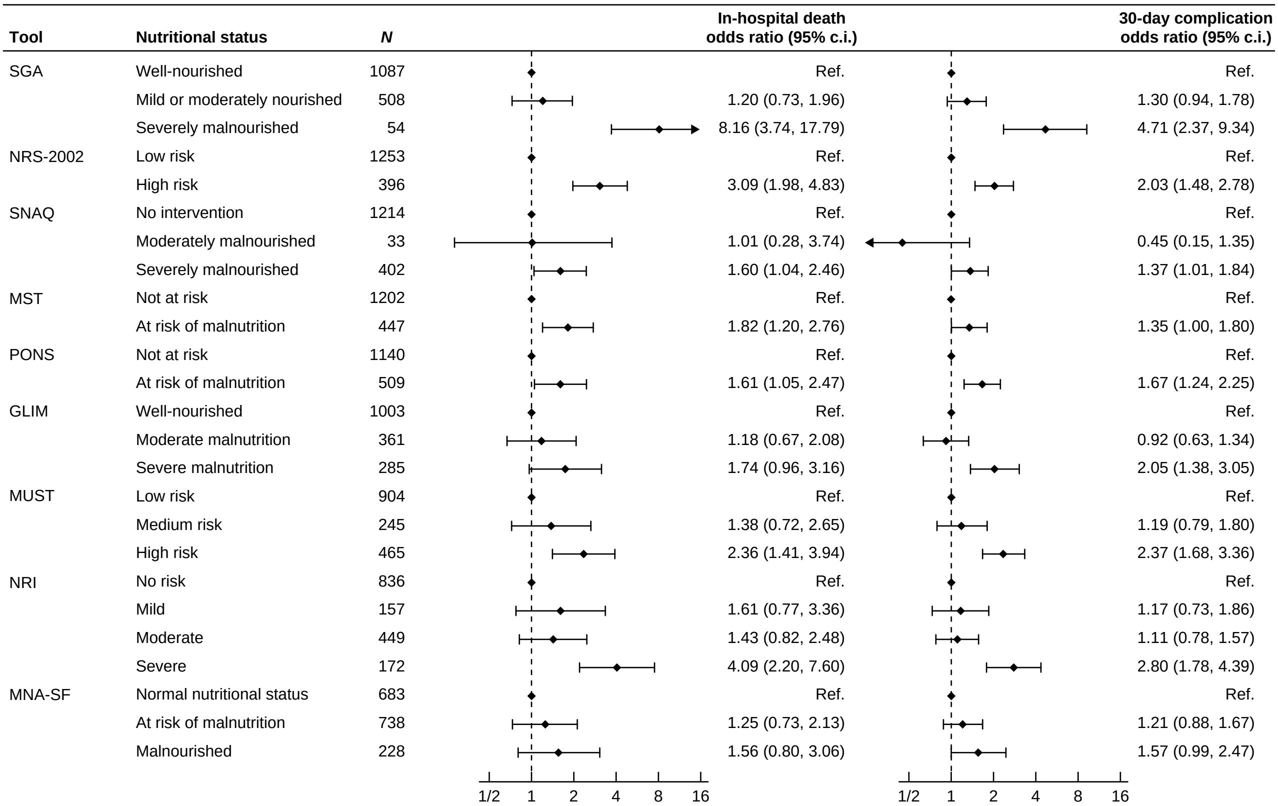
Prognostic validity

Multivariable regression analysis confirmed the prognostic role of SGA independently of major baseline prognostic factors. Patients classified as severely malnourished by SGA were at much higher odds for 30-day serious postoperative complications (OR 4.71; 95% c.i. 2.37 to 9.34) and in-hospital death (OR 8.16; 95% c.i. 3.74 to 17.79) compared to those classified ‘well nourished’ (*Fig. 3*). Prognostic validity was confirmed for NRS-2002, MUST and NRI in both clinical outcomes (ORs > 2). Although NRS-2002 and NRI categories for the highest risks of malnutrition had larger point estimates of ORs than MUST, the respective estimates were less precise as reflected by wide confidence intervals. MUST classified the largest number of patients (29%) in the highest risk category for malnutrition and patients had more than two-fold odds for serious complications (OR 2.36; 95% c.i. 1.41 to 3.94) and in-hospital death (OR 2.37; 95% c.i. 1.68 to 3.36) than patients classified ‘low risk’.

Discussion

This prospective multicentre study examined the validity of eight commonly used malnutrition screening tools in a large cohort of patients undergoing major gastrointestinal surgery. The MUST was the screening tool with the highest sensitivity and specificity, strongest association with postoperative outcomes and minimal between-centre heterogeneity. Therefore, MUST is recommended as the most valid nutritional screening tool for routine use in gastrointestinal surgery.

The present study also showed variability among the screening tools to identify patients at risk of malnutrition and a high between-centre heterogeneity of diagnostic accuracy. This suggests that the choice for a malnutrition screening tool may have a great impact on the management of the patient. Between-centre heterogeneity further implies limited comparability of nutritional data between different studies and even between hospitals using the same screening tool. Hence, policymakers may consider implementing MUST as the preferred tool in guidelines on screening and management of malnutrition. This may lead to improvements in perioperative care and surgical outcomes and enable comparability of data across healthcare centres.



Tool	Nutritional status	N	In-hospital death odds ratio (95% c.i.)	30-day complication odds ratio (95% c.i.)
SGA	Well-nourished	1087	Ref.	Ref.
	Mild or moderately nourished	508	1.20 (0.73, 1.96)	1.30 (0.94, 1.78)
	Severely malnourished	54	8.16 (3.74, 17.79)	4.71 (2.37, 9.34)
NRS-2002	Low risk	1253	Ref.	Ref.
	High risk	396	3.09 (1.98, 4.83)	2.03 (1.48, 2.78)
SNAQ	No intervention	1214	Ref.	Ref.
	Moderately malnourished	33	1.01 (0.28, 3.74)	0.45 (0.15, 1.35)
	Severely malnourished	402	1.60 (1.04, 2.46)	1.37 (1.01, 1.84)
MST	Not at risk	1202	Ref.	Ref.
	At risk of malnutrition	447	1.82 (1.20, 2.76)	1.35 (1.00, 1.80)
PONS	Not at risk	1140	Ref.	Ref.
	At risk of malnutrition	509	1.61 (1.05, 2.47)	1.67 (1.24, 2.25)
GLIM	Well-nourished	1003	Ref.	Ref.
	Moderate malnutrition	361	1.18 (0.67, 2.08)	0.92 (0.63, 1.34)
	Severe malnutrition	285	1.74 (0.96, 3.16)	2.05 (1.38, 3.05)
MUST	Low risk	904	Ref.	Ref.
	Medium risk	245	1.38 (0.72, 2.65)	1.19 (0.79, 1.80)
	High risk	465	2.36 (1.41, 3.94)	2.37 (1.68, 3.36)
NRI	No risk	836	Ref.	Ref.
	Mild	157	1.61 (0.77, 3.36)	1.17 (0.73, 1.86)
	Moderate	449	1.43 (0.82, 2.48)	1.11 (0.78, 1.57)
	Severe	172	4.09 (2.20, 7.60)	2.80 (1.78, 4.39)
MNA-SF	Normal nutritional status	683	Ref.	Ref.
	At risk of malnutrition	738	1.25 (0.73, 2.13)	1.21 (0.88, 1.67)
	Malnourished	228	1.56 (0.80, 3.06)	1.57 (0.99, 2.47)

Fig. 3 Prognostic validity of preoperative malnutrition risk assessments for predicting hospital mortality rate (left panel) and 30-day postoperative major complications (left panel)

Findings of the present study are consistent with a recent meta-analysis (six studies, 5695 patients) that also concluded that MUST had the highest test accuracy for screening preoperative malnutrition (sensitivity 86%, specificity 89%) using SGA as the reference[31]. Similarly high pooled diagnostic accuracy measures for MUST were estimated in another recent meta-analysis of 21 studies involving a combined mixed surgical and non-surgical cohort of 7237 hospitalized patients[22]. MUST outperformed MST, MNA-SF and NRS-2002 in concurrent criterion validity using either the SGA or the ESPEN criteria as a reference. However, methodological quality and results were inconsistent across the studies. The predictive validity for clinical outcomes was not assessed in either meta-analysis.

Other studies have conflicting results. NRS-2002 was recommended among MNA-SF, MUST, MST and NRI as the most appropriate screening tool in colorectal cancer patients[12]. Although the sensitivity and specificity of NRS-2002 was comparable with MUST and MNA-SF, only NRS-2002 was independently associated with postoperative complications. Another study could not recommend a particular screening tool given the low sensitivity and specificity[32]. An earlier systematic review of 83 studies published until 2012 concluded that none of the 28 nutrition screening tools and assessment methods was accurate and consistent in predicting nutritional status and clinical outcomes. However, MUST showed fair to good construct and predictive validity for the adult hospital population. The authors recommended that future research should focus on evaluating existing tools in the same population, which aligns with the present study[33].

The absence of a gold standard for the diagnosis of malnutrition is a limitation. Alternative reference standards include a full nutritional evaluation by a trained professional and assessment methods like SGA and GLIM. Assessments by dieticians and muscle mass measurements are time-consuming and costly and as such difficult to implement in hospitals with limited resources[2]. In the present study, hand grip strength was used to indirectly assess muscle mass to calculate GLIM, which might explain the inferior construct and predictive validity of this tool. The SGA was employed as a reference standard and is arguably the most widely used method in studies evaluating nutritional screening tools. SGA is valid, comprehensive and a good predictor of outcomes in different clinical settings[22]. However, some of the components in all screening tests in the present study are also included in the SGA. This may have raised the sensitivity, specificity and AUC estimates. Moreover, SGA relies on clinical judgment and lacks objective measurements. This may have led to heterogeneity of the results. To limit interrater variation, SGA was performed by a single surgical registrar in each study centre. Finally, the study was conducted on a Caucasian population, which might limit the generalizability of the findings.

The next step is to implement standardized malnutrition management strategies based on MUST screening tool and to assess its impact on long-term outcomes including quality of life and healthcare costs. The potential of combining several screening tools, such as GLIM or expert consultations, to improve perioperative patient management in clinical practice is another future area for research.

Collaborators

Alexandros Tsiavos, John Tsikritzakis, Georgios Papadakis, Konstantina Maria Raouzaiou, Stamatios Kokkinakis and

Aikaterini Rasouli (University Hospital of Heraklion, Crete, Greece); Kyriaki Konstantinou, Panayiotis Papatheodorou, Andia Stylianou and Despo Evripidou (University of Cyprus, Nicosia, Cyprus); Athanasios Syllaios, Markos Despotidis, Adam Mylonakis, Panagiotis Dorovinis, Nikolaos Machairas, Myrto Keramida and Stylianos Kykalos (Laikon Hospital, Athens, Greece); Maximos Frountzas, Dimosthenis Sotirakis and Elpida Tsekoura (Hippokrateio Hospital, Athens, Greece); Konstantinia Kofina (University Hospital of Heraklion, Crete, Greece); Eleni I. Effraimidou, Markos Markou, Rodopi Sotiropoulou, Michael Pitiakoudis Christos Tsalikidis and Anastasios Karayiannakis (Democritus University of Thrace, Thrace, Greece); Georgios Giannos (Evangelismos General Hospital, Athens, Greece); Savvas Tsaramanidis and Antonios Morsi-Yeroyannis (Ippokrateio General Hospital of Thessaloniki, Thessaloniki, Greece); Ifigenia Tzartzalou and Ioannis Gkionis (Venizeleio Hospital, Heraklion, Greece); Garyfallia-Apostolia Karali (University of Thessaly, Larissa, Greece); Dimitrios Kehagias and Georgios-Ioannis Verras (University of Patras, Patras, Greece); and Emmanuel Chrysos (University Hospital of Heraklion, Crete, Greece).

Funding

No external funding was received for this study.

Acknowledgements

None.

Disclosure

The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *BJS* online.

Data availability

Requests to access data should be submitted via the corresponding author. Access to anonymized data may be granted after a review of the data management and sharing plan.

Author contributions

Georgia Petra (Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing—original draft), Evangelos Kritsotakis (Conceptualization, Data curation, Formal analysis, Methodology, Software, Supervision, Visualization, Writing—original draft), Nikolaos Gouvas (Investigation, Writing—review & editing), Dimitrios Schizas (Investigation, Writing—review & editing), Konstantinos Toutouzas (Investigation, Writing—review & editing), Michael Karanikas (Investigation, Writing—review & editing), George Pappas-Gogos (Investigation, Writing—review & editing), Stylianidis Georgios (Investigation, Writing—review & editing), George Zacharioudakis (Investigation, Writing—review & editing), Aggelos Laliotis (Investigation, Writing—review & editing), Gregory Christodoulidis (Investigation, Writing—review & editing), Ioannis Kehagias (Investigation, Writing—review & editing), and Konstantinos Lasithiotakis (Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Supervision, Writing—original draft)

References

1. Soeters P, Bozzetti F, Cynober L, Forbes A, Shenkin A, Sobotka L. Defining malnutrition: a plea to rethink. *Clin Nutr* 2017;**36**: 896–901
2. Cederholm T, Jensen GL, Correia MITD, Gonzalez MC, Fukushima R, Higashiguchi T *et al.* GLIM criteria for the diagnosis of malnutrition—a consensus report from the global clinical nutrition community. *Clin Nutr* 2019;**38**:1–9
3. Skeie E, Koch AM, Harthug S, Fosse U, Sygnestveit K, Nilsen RM *et al.* A positive association between nutritional risk and the incidence of surgical site infections: a hospital-based register study. *PLoS One* 2018;**13**:e0197344
4. Venianaki M, Andreou A, Nikolouzakis TK, Chrysos E, Chalkiadakis G, Lasithiotakis K. Factors associated with malnutrition and its impact on postoperative outcomes in older patients. *J Clin Med* 2021;**10**:2550
5. Saunders J, Smith T. Malnutrition: causes and consequences. *Clin Med (Lond)* 2010;**10**:624–627
6. Kang CY, Halabi WJ, Chaudhry OO, Nguyen V, Pigazzi A, Carmichael JC *et al.* Risk factors for anastomotic leakage after anterior resection for rectal cancer. *JAMA Surg* 2013;**148**:65–71
7. Riad A, Knight SR, Ghosh D, Kingsley PA, Lapitan MC, Parreno-Sacdalan MD *et al.* Impact of malnutrition on early outcomes after cancer surgery: an international, multicentre, prospective cohort study. *Lancet Glob Health* 2023;**11**:e341–e349
8. Zhang B, Najarali Z, Ruo L, Alhusaini A, Solis N, Valencia M *et al.* Effect of perioperative nutritional supplementation on postoperative complications—systematic review and meta-analysis. *J Gastrointest Surg* 2019;**23**:1682–1693
9. Cadili L, van Dijk PAD, Grudzinski AL, Cape J, Kuhnen AH. The effect of preoperative oral nutritional supplementation on surgical site infections among adult patients undergoing elective surgery: a systematic review and meta-analysis. *Am J Surg* 2023;**226**:330–339
10. Skořepa P, Ford KL, Alsuwaylihi A, O'Connor D, Prado CM, Gomez D *et al.* The impact of prehabilitation on outcomes in frail and high-risk patients undergoing major abdominal surgery: a systematic review and meta-analysis. *Clin Nutr* 2024;**43**:629–648
11. Weimann A, Braga M, Carli F, Higashiguchi T, Hübner M, Klek S *et al.* ESPEN practical guideline: clinical nutrition in surgery. *Clin Nutr* 2021;**40**:4745–4761
12. Xie B, Sun Y, Sun J, Deng T, Jin B, Gao J. Applicability of five nutritional screening tools in Chinese patients undergoing colorectal cancer surgery: a cross-sectional study. *BMJ Open* 2022;**12**:e057765
13. Probst P, Haller S, Bruckner T, Ulrich A, Strobel O, Hackert T *et al.* Prospective trial to evaluate the prognostic value of different nutritional assessment scores in pancreatic surgery (NURIMAS pancreas). *Br J Surg* 2017;**104**:1053–1062
14. Copeland GP. The POSSUM system of surgical audit. *Arch Surg* 2002;**137**:15–19
15. Dindo D, Demartines N, Clavien P-A. Classification of surgical complications. *Ann Surg* 2004;**240**:205–213
16. Barker LA, Gout BS, Crowe TC. Hospital malnutrition: prevalence, identification and impact on patients and the healthcare system. *Int J Environ Res Public Health* 2011;**8**:514–527

17. Vierron E, Giraudeau B. Design effect in multicenter studies: gain or loss of power? *BMC Med Res Methodol* 2009;**9**:39
18. de van der Schueren MAE, Keller H, GLIM Consortium, Cederholm T, Barazzoni R, Compher C *et al.* Global Leadership Initiative on Malnutrition (GLIM): guidance on validation of the operational criteria for the diagnosis of protein-energy malnutrition in adults. *Clin Nutr.* 2020;**39**:2872–2880
19. Field LB, Hand RK. Differentiating malnutrition screening and assessment: a nutrition care process perspective. *J Acad Nutr Diet* 2015;**115**:824–828
20. Matarese LE, Charney P. Capturing the elusive diagnosis of malnutrition. *Nutr Clin Pract* 2017;**32**:11–14
21. Jones JM. Validity of nutritional screening and assessment tools. *Nutrition* 2004;**20**:312–317
22. Cortés-Aguilar R, Malih N, Abbate M, Fresneda S, Yañez A, Bennasar-Veny M. Validity of nutrition screening tools for risk of malnutrition among hospitalized adult patients: a systematic review and meta-analysis. *Clin Nutr* 2024;**43**:1094–1116
23. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;**17**: 857–872
24. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003;**56**:1129–1135
25. Snell KI, Ensor J, Debray TP, Moons KG, Riley RD. Meta-analysis of prediction model performance across multiple studies: which scale helps ensure between-study normality for the C-statistic and calibration measures? *Stat Methods Med Res* 2018;**27**: 3505–3522
26. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 2000;**19**:3127–3131
27. Cohen J. A power primer. *Psychol Bull* 1992;**112**:155–159
28. Lasithiotakis K, Kritsotakis EI, Kokkinakis S, Petra G, Paterakis K, Karali G-A *et al.* The Hellenic Emergency Laparotomy Study (HELAS): a prospective multicentre study on the outcomes of emergency laparotomy in Greece. *World J Surg* 2023;**47**:130–139
29. Jandhyala BS. Studies on the effects of felodipine on the vascular resistance and reactivity in pentobarbital anesthetized dogs. *Naunyn Schmiedebergs Arch Pharmacol* 1985;**329**:398–403
30. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L *et al.* STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; **6**:e012799
31. Cheung HHT, Joynt GM, Lee A. Diagnostic test accuracy of preoperative nutritional screening tools in adults for malnutrition: a systematic review and network meta-analysis. *Int J Surg* 2024;**110**:1090–1098
32. Thomas J, Kaambwa B, Delaney C, Miller M. An evaluation of the validity of nutrition screening and assessment tools in patients admitted to a vascular surgery unit. *Br J Nutr* 2019;**122**:689–697
33. van Bokhorst-de van der Schueren MAE, 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**:39–58