World J Surg (2019) 43:3003–3012
https://doi.org/10.1007/s00268-019-05139-z

SCIENTIFIC REVIEW

Quality Assessment and Risk of Bias of Systematic Reviews of Prophylactic Mesh for Parastomal Hernia Prevention Using AMSTAR and ROBIS Tools

Josep M. García-Alamino[1] · Manuel López-Cano[2] · Leonard Kroese[3] · Frederik Helgstrand[4] · Filip Muysoms[5]

Published online: 22 August 2019

Abstract

Background Systematic reviews play a crucial role in clinical decision making and resource allocation and are expected to be unbiased and consistent. The aim of this study is a review of systematic reviews on the use of prophylactic mesh to prevent parastomal hernia (PH) formation using ROBIS and AMSTAR tools to assess the risk of bias and methodological quality.

Methods We included systematic reviews with or without meta-analysis of which the objective was to assess the use of a prophylactic mesh to prevent PH. A systematic search of the literature in five databases from inception until December 2017 was conducted. For each systematic review, methodologic quality and risk of bias were assessed using the AMSTAR and ROBIS tools, respectively. We estimated the inter-rater reliability for individual domains and for the overall methodological quality and risk of bias using Fleiss' *k*.

Results We identified 14 systematic reviews that met the inclusion criteria. Using the AMSTAR scale with a cutoff value, six reviews showed high methodologic quality and eight were of low quality. Using the ROBIS tool, the overall risk of bias was low in 50% of the reviews analyzed. In the remaining studies, the risk of bias was unclear.

Conclusions The global evidence in favor of the use of a prophylactic mesh for preventing PH is not uniform regarding quality and risk of bias. Surgeons cannot be equally confident in the results of all systematic reviews published on this topic.

Introduction

The decision to use a prophylactic mesh at end-colostomy construction to prevent parastomal hernia (PH) formation after surgery remains unsolved [1, 2]. The process of

Josep M. García-Alamino and Manuel López-Cano have contributed equally to this work.

✉ Manuel López-Cano
mlpezcano@gmail.com

[1] DPhil Programme in Evidence-Based Healthcare, University of Oxford, Oxford, UK

[2] Abdominal Wall Surgery Unit and General and Digestive Surgery Research Group, Department of General Surgery, Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autónoma de Barcelona, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain

[3] Department of Surgery, Erasmus University Medical Center, Rotterdam, The Netherlands

[4] Department of Gastrointestinal Surgery, Centre for Surgical Science, Zealand University Hospital, Køge, Denmark

[5] Department of Surgery, Maria Middelares Hospital, Ghent, Belgium

clinical decision making in surgical practice is based on three basic pillars: the level of evidence of studies published in literature, surgeon's personal experience, and patient's values and preferences after being provided with adequate information [3]. When evidence is strong (i.e., high quality), the validity behind patient advice and information on the benefits and harms of a particular procedure are higher [3]. Evidence-based medicine postulates that recommending procedures when there is high-quality evidence in favor of their support is the most rational recommendation to make [4].

Systematic reviews and meta-analysis of high-quality randomized controlled trials are placed at the top of the hierarchy of evidence [3, 4], and have become essential for developing clinical guidelines and decision aids to facilitate clinical decision making [3]. However, systematic flaws, limitations in the design or conduct of a review have the potential to bias results. To prevent this, explicit guidelines and recommended standards have been developed [5–7]. Also, several tools exist for undertaking critical appraisal of the quality of systematic reviews. Although none of them has become universally accepted, the AMSTAR tool is probably the most commonly accepted quality assessment tool for systematic reviews [8, 9]. The ROBIS is a recent and first rigorously developed tool designed specifically to assess the risk of bias of systematic reviews [10].

In this context, systematic reviews and meta-analysis of high methodologic quality (i.e., low risk of bias) play a key role in the decision and recommendation of placing a prophylactic mesh to prevent PH. The objective of this study was to present a review of systematic reviews focused on the prevention of PH with mesh, using both AMSTAR and ROBIS tools to assess the quality and risk of bias of the reviews.

Material and methods

Eligibility

Eligible studies were systematic reviews with and without meta-analysis aimed to assess the use of a prophylactic mesh for the prevention of PH formation in patients undergoing permanent end-colostomy. Characteristics of the systematic reviews should fulfill the following PICO approach. "Participants" were randomized controlled trials (RCTs) and observational studies carried out in adult patients in whom a prophylactic mesh was placed at the time of end-colostomy construction to prevent PH formation after surgery; "intervention" was mesh augmentation at the time of stoma formation including synthetic, biosynthetic as well as biologic meshes in all anatomical positions of the abdominal wall; "comparator" was defined as conventional permanent colostomy creation without mesh; and "outcome" was the incidence of PH after surgery and other adverse events directly related or unrelated to mesh insertion. Systematic reviews that included participants with other types of PH prevention different to mesh augmentation and those including studies mixing prevention of other types of hernias were excluded. Because the present study was not a systematic review, the study protocol was not registered in the PROSPERO database.

Search and selection methods

A systematic search of the literature was performed in MEDLINE/PubMed, SCOPUS, CINAHL, WOK (Web of Knowledge), and Cochrane database from inception until December 31, 2017. The search terms "prophylaxis," "prevention," and "parastomal hernia" were used separately or combined either in the title/abstract or all fields. We included reviews published in any language. The reference lists of all retrieved studies were cross-checked for additional reports. Two authors (M.L-C. and F.M.) independently screened all abstracts of the articles retrieved to identify eligible studies.

The methodologic quality of systematic reviews as outlined in the Standards for the Conduct of New Cochrane Interventions Reviews [5] and Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) [7] and the risk of bias (or internal validity) of each systematic review was assessed using both the AMSTAR scale [8, 9] and the ROBIS [10] tool, respectively.

The AMSTAR scale (A Measurement Tool to Assess Systematic Reviews) consists of 11 items, each of which is categorized into a standardized set of four possible responses: "yes," "no," "cannot answer," "not applicable." The items relate to a priori design, study selection and data extraction, the literature search, gray literature, the list of included and excluded studies, study characteristics, critical appraisal, formulation of conclusions, the combination of study results, publication bias, and conflicts of interest. If all criteria are met for an individual item (e.g., "yes"), it received a score of 1 and the sum of all "yes "responses indicates the total score out of 11. Although there is a debate on whether or not overall AMSTAR scores should be calculated [9, 11, 12], we used the tool as guidelines teams and assigned high quality to a review for which 8 to 11 items were judged fulfilled (answered "yes") and low quality to a review for which 3 or fewer items were judged fulfilled (answered "yes"). This threshold is largely used as described by other authors [13]. The aim of the ROBIS (Risk of Bias in Systematic Reviews) tool is to evaluate the risk of bias present within a systematic review

[10]. This tool assesses the level of bias across four specific domains: study eligibility criteria, identification and selection of studies, data collection, and synthesis and findings. Each domain has signaling questions and a judgment of concerns about risk of bias of the domain ("low," "high," "unclear"). In the final phase, the reviewer makes a judgment about the overall risk of bias [13]. ROBIS has a wider application and is intended for assessing effectiveness, diagnostic test accuracy, prognosis, and etiology. AMSTAR and ROBIS differ in their theoretical construct (methodological quality vs. risk of bias), and not all the domains of one tool can be matched with the domains in the other. However, some authors have defined up to eight domains as fully overlapping [13].

Data extraction and analysis

Four raters (FM, FH, LK, ML-C) independently assessed each review using both tools. In case of disagreement, a fifth rater (JG-A) participated in the discussion to reach consensus. Raters read the full article of each selected systematic review and applied the two tools. Raters were advised to completely read and follow the guidance available for AMSTAR scale [8, 9] and ROBIS tool [10]. The raters were free to read and make the assessments of the systematic reviews in the order of their preference. The length of time allowed to complete the assessments was 4 weeks. Also, before starting the reliability test, raters independently familiarized themselves with the two tools by reviewing the documents related to AMSTAR (https://amstar.ca/) and ROBIS (https://www.bristol.ac.uk/population-health-sciences/projects/robis/). We estimated inter-rater reliability (the degree of agreement between two or more raters in their appraisals) for individual domains and for the overall methodological quality and risk of bias using both the AMSTAR and ROBIS tools. The reliability was calculated using Fleiss' kappa (k) statistics for multiple raters. Agreement was classified as suggested by Landis and Koch [14]: Values less than zero indicate no agreement, 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 almost perfect. Statistical analyses were performed using several packages of the R software.

Results

Search results

A total of 36 systematic reviews were identified. After screening the titles and abstracts, 18 reviews [2, 15–31] were judged as potentially eligible for full-text analysis. Finally, 14 systematic reviews [2, 16–20, 22–24, 27–31] met the inclusion criteria and were included in the study. The flowchart of the study selection process is shown in Fig. 1.

Characteristics of the systematic reviews

Table 1 summarizes the main characteristics of the studies. All of them were published between 2008 and 2017 and originated from nine countries. A total of four systematic reviews included both RCTs and observational studies [16, 17, 22, 31]. Two systematic reviews [16, 22] did not synthesize data quantitatively, and meta-analysis was not performed. The risk of bias assessment for the individual RCTs was reported for all the systematic reviews except in two reviews [16, 22].

Methodologic Quality and Risk of Bias.

Table 2 shows the results of the AMSTAR scale, both as total scores and quality of each systematic review. The inter-rater agreement was poor for total scores of the instrument ($\kappa = 0.04$) using quantitative data, however, improved and was fair when dichotomous data (high vs low quality) of the reviews were considered ($\kappa = 0.30$).

Table 3 and Fig. 2 show the agreement on individual domains and overall judgments for ROBIS tool. With regard to the domain of "study eligibility criteria," the score for the risk of bias was low in 11 systematic reviews, high in one, and unclear in two; the inter-rater agreement was slight ($\kappa = 0.11$). In the domain of "identification and selection of studies," the risk of bias was low in eight reviews, high in one and unclear in the remaining five; the inter-rater agreement was slight ($\kappa = 0.19$). In the domain of "data collection and study appraisal," the risk of bias was low in 11 systematic reviews, high in two, and unclear in one; the inter-rater agreement was moderate ($\kappa = 0.55$). In the domain of "synthesis and findings," the risk of bias was low in four systematic reviews, high in five, and unclear in five; the inter-rater agreement was slight ($\kappa = 0.18$). The overall risk of bias for all systematic reviews was low in seven of them, high in one, and unclear in the remaining six. The global rating agreement for the four domains was fair ($k = 0.34$).

Discussion

The novelty of this study is the use of both AMSTAR and ROBIS tools in the surgical setting. Based on our literature search, 14 reviews on the effectiveness of placement of a prophylactic mesh at end-colostomy construction to prevent PH were published by December 31, 2017, and included in the analysis. Threshold of the AMSTAR scale for quality assessment allowed categorizing six systematic revisions of high quality and the remaining eight of low

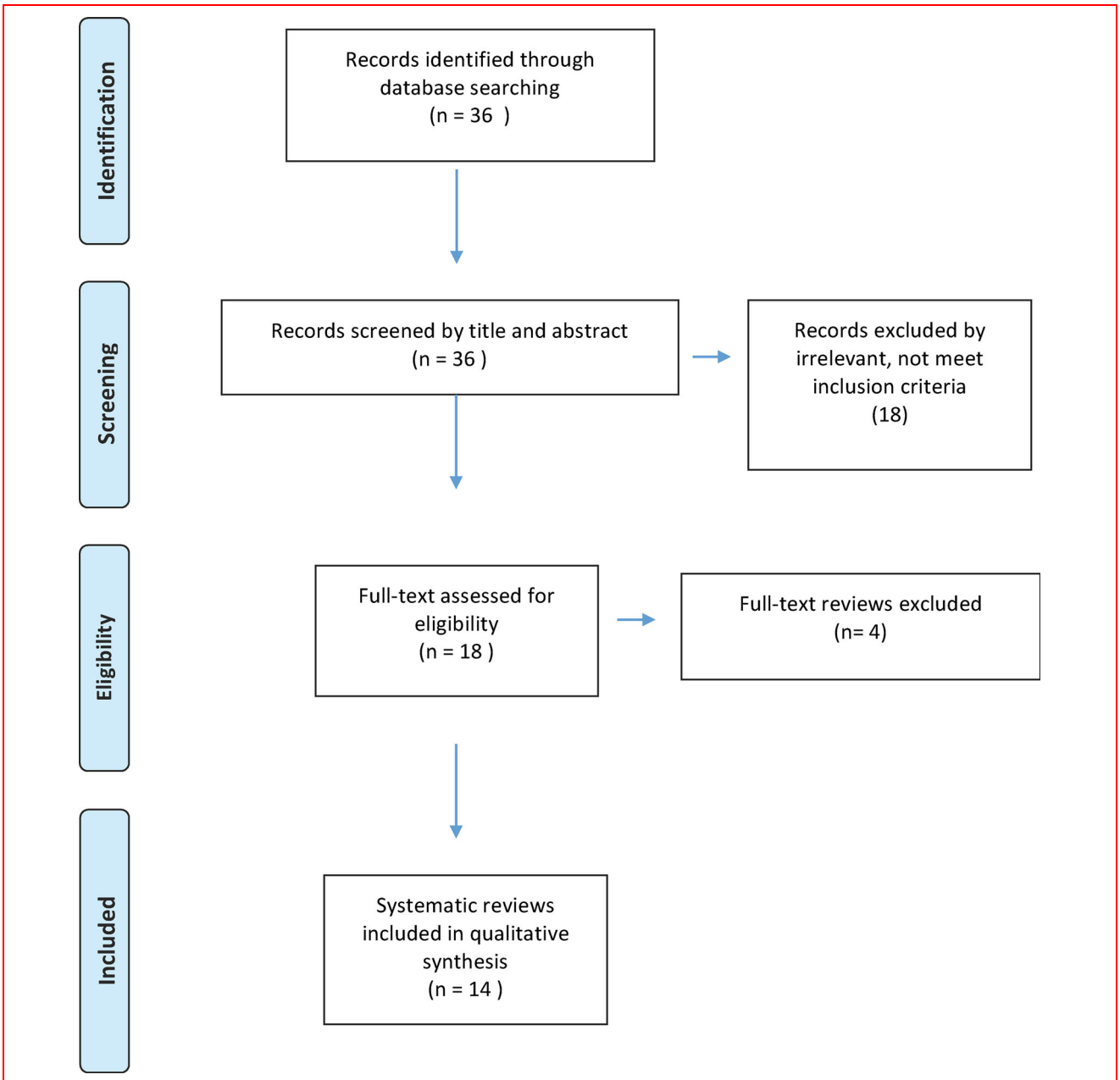

Fig. 1 Flowchart of the systematic reviews screened and included in the study

quality. With the use of ROBIS tool, the overall risk of bias was low in 50% of the reviews. However, in almost the remaining 50%, the risk of bias was unclear. With regard to individual dimensions of ROBIS, "synthesis and findings" showed the highest risk of bias. This higher risk may be related to the similarity of the research questions, the between-study heterogeneity, and the robustness of the findings through meta-analyses.

From a methodologic point of view, it has been questioned whether correlation between judgments on similar domains common to both AMSTAR and ROBIS can be compared. However, in some studies in which this comparison was performed, correlations ranged from slight to moderate agreement [32, 33]. The possibility of comparing the correlation between the two tools is improved by establishing a threshold, as we have done in our study. Also, the use of qualitative data rather than a specific numerical value contributes to improve inter-rater agreement.

Half of systematic reviews assessed with ROBIS showed a low risk of bias, which provides confidence to the use of a prophylactic mesh to prevent PH formation, and

Table 1 Characteristics and results of the 14 systematic reviews included in the study

Author, year, country	Studies included	Patients no	Type of surgery	Intervention mesh	Comparator	Primary outcome	Secondary outcome	Follow-up months mean	Diagnosis type	Risk of bias assessment	Quantitative synthesis	Sensitivity analysis	Safety adverse events
Helgstrand et al. [16] Denmark	1 RCT 4 Observational	112	Open colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	Complications	22.4	Not reported	Not reported	No	No	Yes
Tam et al. [17] China	3 RCTs 5 Observational	237	Open colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	Mesh placement and type of complications (infection, stoma necrosis)	28.9	Physical examination CT, US	Yes (for RCTs)	Yes	Not reported	Yes
Wijeyekoon et al. [18] UK	3 RCTs	129	Open colostomy	Synthetic and biologic	No mesh	PH incidence	PH with end and loop Stomas, PH repair Mortality, morbidity	27.3	Physical examination CT	Yes	Yes	Yes	Yes
Shabbir et al. [19] UK	3 RCTs	129	Open colostomy	Synthetic and biologic	No mesh	PH incidence	Morbidity	Not clearly stated	Physical examination CT	Yes	Yes	Not reported	Yes
Sajid et al. [20] UK	3 RCTs	129	Open colostomy	Synthetic and biologic	No mesh	PH incidence	Complications (hematoma, seroma, occlusion, infection)	33.6	Physical examination CT	Yes (Jadad, Cochrane)	Yes	Not reported	Yes
Fortelny et al. [22] Austria	6 Systematic Reviews, 2 RCTs, 2 case–control, 1 technical report	182 (?)	Open colostomy and ileostomy	Biologic	No mesh	PH incidence	Complications	Unclear	Unclear	Not reported	No	No	Yes
Wang et al. [23] China	6 RCTs (1 duplicated the same series with fewer participants)	309	Open and laparoscopic colostomy	Synthetic	No mesh	PH incidence	Morbidity PH-related reoperation	21.9	Physical examination US. CT	Yes (Cochrane GRADE)	Yes	Yes	Yes
Zhu et al. [27] China	8 RCTs	522	Open and laparoscopic colostomy	Synthetic	No mesh	PH Incidence	PH repair infection	25.2	Physical examination US, CT	Yes (Cochrane)	Yes	Yes	Yes
Cross et al. [30]N. Zealand	10 RCTs	649	Open and laparoscopic colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	PH repair, parastomal infection, necrosis, stenosis	12 months to 5 years	Physical examination US, CT	Yes (Cochrane)	Yes	Yes	Yes

Table 1 continued

Author, year, country	Studies included	Patients no	Type of surgery	Intervention mesh	Comparator	Primary outcome	Secondary outcome	Follow-up months mean	Diagnosis type	Risk of bias assessment	Quantitative synthesis	Sensitivity analysis	Safety adverse events
Patel et al. [29] Canada	9 RCTs	590	Open and laparoscopic colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	PH repair complications stoma complications	23.4	Physical Examination US. CT	Yes (Cochrane, GRADE)	Yes	Yes	Yes
Chapman et al. [28] UK	7 RCTs	464	Open and laparoscopic colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	PH repair complications	Not clearly stated	Physical examination CT, US	Yes (Cochrane)	Yes	Yes	Yes
López-Cano et al. [2] Spain	7 RCTs	451	Open and laparoscopic colostomy	Synthetic	No mesh	PH incidence	Wound infection, hematoma, seroma	26.02	Physical examination CT	Yes (SIGN, GRADE)	Yes	Yes	Yes
Cornille et al. [24] UK	8 RCTs	430	Open and laparoscopic colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	Infection cellulitis seroma	26.5	Physical examination CT, US	Yes (Cochrane)	Yes	Yes	Yes
Pianka et al. [31] Germany	8 RCTs 3 Observational	755	Open and laparoscopic colostomy and ileostomy	Synthetic and biologic	No mesh	PH incidence	Stoma necrosis mesh-related reintervention stoma fistula stoma site infection perioperative mortality	22.2	Physical examination CT, US	Yes (Cochrane)	Yes	Yes	Yes

RCT randomized controlled trial, *PH* parastomal hernia, *CT* computed tomography, *US* ultrasound

Table 2 Results of the AMSTAR evaluation

Systematic review	Rater #1		Rater #2		Rater #3		Rater #4	
	Total score	Quality	Total score	Quality	Total score	Quality	Total score	Quality
1. Helgstrand et al. [16]	0/11	Low	3/11	Low	0/11	Low	1/11	Low
2. Tam et al. [17]	7/11	Low	6/11	Low	2/11	Low	6/11	Low
3. Wijeyekoon et al. [18]	9/11	High	9/11	High	5/11	Low	8/11	High
4. Shabbir et al. [19]	6/11	Low	6/11	Low	4/11	Low	5/11	Low
5. Sajid [20]	5/11	Low	6/11	Low	5/11	Low	7/11	Low
6. Fortelny et al. [22]	3/11	Low	3/11	Low	2/11	Low	0/11	Low
7. Wang et al. [23]	4/11	Low	8/11	high	6/11	Low	9/11	High
8. Zhu et al. [27]	7/11	Low	9/11	High	4/11	Low	7/11	Low
9. Cross et al. [30]	10/11	High	10/11	High	7/11	Low	6/11	Low
10. Patel et al. [29]	7/11	Low	10/11	High	5/11	Low	8/11	High
11. Chapman et al. [28]	10/11	High	10/11	High	6/11	Low	9/11	High
12. López-Cano et al. [2]	10/11	High	10/11	High	7/11	Low	9/11	High
13. Cornille et al. [24]	9/11	High	10/11	High	6/11	Low	7/11	Low
14. Pianka et al. [31]	8/11	High	10/11	High	8/11	High	8/11	High

AMSTAR scale consisted of 11 items, for each of which there are four possible responses: "yes," "no," "cannot answer," "not applicable." The sum of all "yes" responses indicates the total score out of 11. High quality is assigned when 8 to 11 items were judged fulfilled (answered "yes") and low quality when 3 or fewer items were judged fulfilled (answered "yes")

Table 3 Tabular presentation for the ROBIS results

Systematic review	Phase 2				Phase 3
	1. Study eligibility criteria	2. Identification and selection of studies	3. Data collection and study appraisal	4. Synthesis and findings	Risk of bias in the review
1. Helgstrand et al. [16]	?	?	☹	☹	?
2. Tam et al. [17]	☺	?	?	☹	?
3. Wijeyekoon et al. [18]	☺	?	☺	?	?
4. Shabbir et al. [19]	☺	?	☺	☹	?
5. Sajid et al. [20]	☺	☺	☺	☹	☺
6. Fortelny et al. [22]	☹	☹	☹	☹	☹
7. Wang et al. [23]	☺	☺	☺	☺	☺
8. Zhu et al. [27]	☺	?	☺	?	?
9. Cross et al. [30]	☺	☺	☺	?	☺
10. Patel et al. [29]	?	☺	☺	?	?
11. Chapman et al. [28]	☺	☺	☺	☺	☺
12. López-Cano et al. [2]	☺	☺	☺	☺	☺
13. Cornille et al. [24]	☺	☺	☺	?	☺
14. Pianka et al. [31]	☺	☺	☺	☺	☺

☺ low risk, ☹ high risk, ? unclear

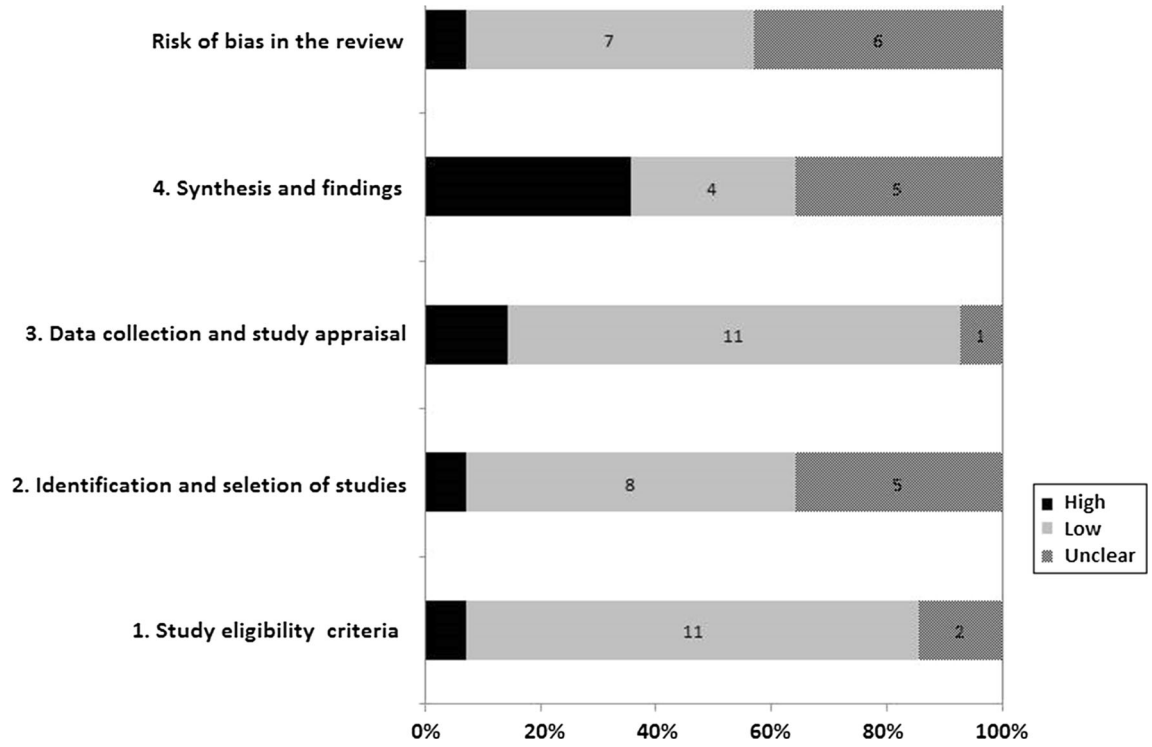

Fig. 2 Overall risk of bias for the ROBIS domains

the incorporation of this type of intervention in clinical practice. In another sense, it is important to consider that in almost the other half of the reviews, the risk of bias was unclear, which means that systematic reviews on the use of prophylactic mesh to prevent PH published so far have different quality and risk of bias. Therefore, we cannot have the same confidence in the results of all systematic reviews published on this topic. This aspect reinforces the need to carry out a critical evaluation of each systematic review before its results are incorporated into daily practice or its results become part of the recommendations of clinical practice guidelines. Using the AMSTAR scale, a high percentage of reviews were of high quality, which reflects that adequacy of the reviews to consensuated standards (AMSTAR) was appropriate in reporting results and, therefore, it is assumed that studies were also performed correctly.

It is difficult to extrapolate how systematic reviews influence the use of a prophylactic mesh for the prevention of PH. When the quality of the systematic reviews is "solid," the estimate about harms or benefits of a procedure will be better calibrated [2], and this appears to be the best context to perform a procedure [3]. However, there are few data in the literature on the use of a prophylactic mesh for the prevention of a PH and it seems that its use is low. Parkinson et al. [34] conducted an e-mail survey of the corresponding authors of 200 original articles published in two journals, Colorectal Disease and Diseases of the Colon and Rectum, with a response rate of 55%. Of the 111 respondents, 65 (58.6%) were colorectal specialists with at least 10 years in practice. Overall, only 17% of respondents said that they had used or had observed the use of prophylactic mesh during the last elective permanent colostomy in which they had participated. Surgeons' apprehension at using a preventive device based on a fear of infection and other mesh-related complications or the lack of convincing efficacy data has been also reported [6]. There is no a "unique model" in decision making, and what some surgeons may find useful in a given context may be useless for others. Some authors have called this "normative pluralism" [3, 35]. However, when making decisions regarding the prophylactic mesh, it should be considered that not all the systematic reviews analyzed in the present study have the same quality and risk of bias, so that low-quality reviews or with a high risk of bias should be excluded as reliable sources for the establishment of recommendations on the use of mesh to prevent PH in clinical practice.

The usability and applicability of both AMSTAR and ROBIS tools have not been formally evaluated in this study, although we consider (informally) that they meet both criteria to be incorporated as an assessment tool

administered by clinicians and it would be advisable to use both tools to perform a comprehensive assessment of a systematic review. However, recent studies have shown that the time to administer ROBIS is substantially longer than for AMSTAR [13, 36].

Another interesting aspect is the existence of a greater number of systematic reviews than RCTs—individual studies—on the topic of interest (14 systematic reviews are assessed and between 7 and 10 RCTs are included depending on the review). This situation has already been analyzed, criticized by other authors and should reflect on the clinical and ethical implications of the waste of resources, because a large majority of produced systematic and meta-analyses may be unnecessary and misleading [37].

One of the strengths of the study is the inclusion of all published systematic reviews on this topic up to date of analysis, which has made it possible to observe the different quality among all existing reviews. Of note, that a further systematic review [38] showing a reduction in the incidence of PH in patients who had a prophylactic synthetic mesh was published when the analysis had been completed. However, given the objective of our study and the large number of systematic reviews assessed, the fact of not including this last systematic review would probably have a little influence on the present findings. A further strength is the use of two validated tools to assess the quality (AMSTAR) and the risk of bias (ROBIS). This provides a more comprehensive assessment of the systematic reviews. However, one of the limitations inherent in the comparison of both tools is that they are not directly comparable because, although they are related, they do not determine exactly the same. Logically, the items in one tool and another are different and therefore are not directly comparable [8–10]. Also, the study was focused especially on evaluating the reliability of these tools, but the usability and applicability of AMSTAR versus ROBIS have not been formally evaluated in a descriptive or comparative way. The different profiles of the evaluators as well as the inexperience with this type of tools should also be mentioned. The slight inter-rater agreement between raters for ROBIS tool has been previously observed in a similar study [33]. However, the inter-rater agreement for the four domains was fair. One aspect that should also be considered is the use of a threshold in the AMSTAR scale because it allows comparability with ROBIS and also improves the inter-rater agreement in the administration of AMSTAR to the systematic reviews. In our work when a threshold was used (dichotomous data—high *vs* low quality), the inter-rater agreement improves.

In the case of future similar studies, our recommendation is to use both tools to evaluate both the quality of the review and the risk of bias. ROBIS is a sophisticated instrument that focuses specifically on the risk of bias introduced by the conduct of the review. It covers most types of research questions, including diagnosis, prognosis, and etiology. In contrast, AMSTAR, which is one of the most widely used instruments, is assessing the quality only of systematic reviews of RCTs about interventions [39]. Recently, AMSTAR 2 has been published, which is a major revision of the original 11-item AMSTAR instrument. The main modifications include simplified response categories, a more detailed consideration of risk of bias with included studies, and how this was handled by review authors in summarizing and interpreting the results of their reviews [40]. It would be relevant to compare the agreement between AMSTAR, AMSTAR 2, and ROBIS as has been proposed in a study protocol [41].

In summary, using the ROBIS tool, only half of the 14 systematic reviews of mesh to prevent PH had a low risk of bias, and using the AMSTAR tool, only six systematic reviews were of high methodologic quality. But, the inter-rater agreement was slight. Accordingly, the global evidence in favor of the use of a prophylactic mesh for preventing PH is not uniform regarding quality and risk of bias. Surgeons cannot be equally confident in the results of all systematic reviews published on this topic. This aspect reinforces the need to carry out a critical evaluation of each systematic review before its results are incorporated into daily practice or its results become part of the recommendations of clinical practice guidelines. AMSTAR and/or ROBIS can be suitable tools for this purpose.

Acknowledgements The authors are grateful to Marta Pulido, MD, for editing the manuscript and editorial assistance.

Funding No direct or indirect financial support was received.

Compliance with ethical standards

Conflict of interest ML-C reports personal fees (consultancy, speaker) from B. Braun and Bard Davol outside the submitted work. FM reports grants and personal fees (consultancy, speaker, grants) from Medtronic, CMR Surgical, Intuitive Surgical, Dynamesh, Bard Davol outside the submitted work. JMG-A, LK, FH declare no competing interests.

References

1. Antoniou SA, Agresta F, García Alamino JM et al (2018) European Hernia Society guidelines on prevention and treatment of parastomal hernias. Hernia 22(1):183–198
2. López-Cano M, Brandsma HT, Bury K et al (2017) Prophylactic mesh to prevent parastomal hernia after end colostomy: a meta-analysis and trial sequential analysis. Hernia 21:177–189
3. Djulbegovic B, Guyatt GH (2017) Progress in evidence-based medicine: a quarter century on. Lancet 390:415–423

4. Djulbegovic B, Elqayam S, Dale W (2018) Rational decision making in medicine: implications for overuse and underuse. J Eval Clin Pract 24:655–665
5. Higgins J, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological expectations of cochrane intervention reviews (MECIR). Standards for the conduct and reporting of new cochrane intervention reviews, reporting of protocols and the planning, conduct and reporting of updates. Version 1.07 (last update November 2018). https://community.cochrane.org/mecir-manual. Accessed 8 Dec 2018
6. Eden J, Levit L, Berg AO, Morton S (eds.) (2011) Finding what works in health care: standards for systematic reviews. The National Academies Press, Washington, D.C
7. Liberati A, Altman DG, Tetzlaff J et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Plos Med 6(7):e1000100
8. Shea BJ, Grimshaw JM, Wells GA et al (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 7:10
9. Shea BJ, Hamel C, Wells GA et al (2009) AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 62:1013–1020
10. Whiting P, Savović J, Higgings JP et al (2016) ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 69:225–234
11. Burda BU, Holmer HK, Norris SL (2016) Limitations of a measurement tool to assess systematic reviews (AMSTAR) and suggestions for improvement. Syst Rev 5:58
12. Pieper D, Buechter RB, Li L et al (2015) Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. J Clin Epidemiol 68:574–583
13. Banzi R, Cinquini M, Gonzalez-Lorenzo M et al (2018) Quality assessment versus risk of bias in systematic reviews: AMSTAR and ROBIS had similar reliability but differed in their construct and applicability. J Clin Epidemiol 99:24–32
14. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
15. Israelsson LA (2005) Preventing and treating parastomal hernia. World J Surg 29:1086–1089https://doi.org/10.1007/s00268-005-7973-z
16. Helgstrand F, Gögenur I, Rosenberg J (2008) Prevention of parastomal hernia by the placement of a mesh at the primary operation. Hernia 12:577–582
17. Tam KW, Wei PL, Kuo LJ et al (2010) Systematic review of the use of a mesh to prevent parastomal hernia. World J Surg 34:2723–2729https://doi.org/10.1007/s00268-010-0739-2
18. Wijeyekoon SP, Gurusamy K, El-Gendy K et al (2010) Prevention of parastomal herniation with biologic/composite prosthetic mesh: a systematic review and meta-analysis of randomized controlled trials. J Am Coll Surg 211:637–645
19. Shabbir J, Chaudhary BN, Dawson R (2012) A systematic review on the use of prophylactic mesh during primary stoma formation to prevent parastomal hernia formation. Colorectal Dis 14:931–936
20. Sajid MS, Kalra L, Hutson K et al (2012) Parastomal hernia as a consequence of colorectal cancer resections can prophylactically be controlled by mesh insertion at the time of primary surgery: a literature based systematic review of published trials. Minerva Chir 67:289–296
21. Hotouras A, Murphy J, Thaha M et al (2013) The persistent challenge of parastomal herniation: a review of the literature and future developments. Colorectal Dis 15:e202–214
22. Fortelny RH, Hofmann A, May C et al (2015) Prevention of a parastomal hernia by biological mesh reinforcement. Front Surg 2:53
23. Wang S, Wang W, Zhu B et al (2016) Efficacy of prophylactic mesh in end-colostomy construction: a systematic review and meta-analysis of randomized controlled trials. World J Surg 40:2528–2536https://doi.org/10.1007/s00268-016-3576-0
24. Cornille JB, Pathak S, Daniels IR et al (2017) Prophylactic mesh use during primary stoma formation to prevent parastomal hernia. Ann R Coll Surg Engl 99:2–11
25. Muysoms FE, Dietz UA (2016) Prophylactic meshes in the abdominal wall. German Version Chirurg 87:751–761
26. Muysoms FE, Dietz UA (2017) Prophylactic meshes in the abdominal wall. Chirurg 88(Suppl 1):34–41
27. Zhu J, Pu Y, Yang X, et al. (2016) Prophylactic mesh application during colostomy to prevent parastomal hernia: a meta-analysis. Gastroenterol Res Pract 2016: Article ID 1694265
28. Chapman SJ, Wood B, Drake TM et al (2017) Systematic review and meta-analysis of prophylactic mesh during primary stoma formation to prevent parastomal hernia. Dis Colon Rectum 60:107–115
29. Patel SV, Zhang L, Chadi SA et al (2017) Prophylactic mesh to prevent parastomal hernia: a meta-analysis of randomized controlled studies. Tech Coloproctol 21:5–13
30. Cross AJ, Buchwald PL, Frizelle FA et al (2017) Meta-analysis of prophylactic mesh to prevent parastomal hernia. Br J Surg 104:179–186 https://doi.org/10.1002/bjs.10402
31. Pianka F, Probst P, Keller A-V et al (2017) Prophylactic mesh placement for the PREvention of paraSTOmal hernias: the PRESTO systematic review and meta-analysis. PLoS ONE 12:e0171548
32. Gómez-García F, Ruano J, Gay-Mimbrera J et al (2017) Most systematic reviews of high methodological quality on psoriasis interventions are classified as high risk of bias using ROBIS tool. J Clin Epidemiol 92:79–88
33. Perry R, Leach V, Davies P et al (2017) An overview of systematic reviews of complementary and alternative therapies for fibromyalgia using both AMSTAR and ROBIS as quality assessment tools. Syst Rev 6:97
34. Parkinson F, Dafydd L, Singh R et al (2014) Preventing parastomal herniation in 2014 and beyond. Colorectal Dis 16:390
35. Djulbegovic B, Elqayam S (2017) Many faces of rationality: implications of the great rationality debate for clinical decision-making. J Eval Clin Pract 23:915–922
36. Bühn S, Mathes T, Prengel P et al (2017) The risk of bias in systematic reviews tool showed fair reliability and good construct validity. J Clin Epidemiol 91:121–128
37. Ioannidis JP (2016) The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 94:485–514
38. Jones HG, Rees M, Aboumarzouk OM, et al. (2018) Prosthetic mesh placement for the prevention of parastomal herniation. Cochrane Database Syst Rev 7:CD008905
39. Shea BJ, Bouter LM, Peterson J et al (2007) External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS ONE 2:e1350
40. Shea BJ, Reeves BC, Wells G et al (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 358:j4008
41. Gates A, Gates M, Duarte G et al (2018) Evaluation of the reliability, usability, and applicability of AMSTAR, AMSTAR 2, and ROBIS: protocol for a descriptive analytic study. Syst Rev 7:85