



# A scoping review of assessment methods of competence of general surgical trainees

Lauren Hackney\*, Stephen O'Neill, Mark O'Donnell, Roy Spence

Queens University Belfast, United Kingdom

## ARTICLE INFO

### Article history:

Received 24 November 2021

Accepted 19 January 2022

Available online 14 March 2022

### Keywords:

General surgery

Surgical training

Surgical registrar

Medical education

## ABSTRACT

**Background:** Only rigorous evaluation of competence will result in the production of safe surgeons that are able to provide the best care for patients. The development of competency-based assessment should ultimately be evidence driven.

**Objectives:** Explore the volume of existing evidence pertaining to the different objective assessment methods reported in the literature.

**Eligibility criteria:** Studies describing objective assessment of postgraduate general surgical trainees within the last 20 years.

**Sources of evidence:** PubMed, Ovid Medline and Web of Sciences.

**Charting methods:** A data chart proforma was designed and data were extracted into tables. Basic numerical analysis of extracted data and narrative synthesis of charted data.

**Results:** A total of 343 papers were reviewed. 26 were eligible for inclusion. 92% of articles were published from 2008 onwards. 50% have been published in the last five years. The articles originated from 6 different countries, predominantly the United Kingdom (42%), followed by the United States of America (38%). In addition, a small number were published from Canada (8%), Japan (4%), Germany (4%) and Australia (4%). UK publications were predominantly between 2008 and 2014 while the USA had a later predominance between 2015 and 2018. 42% were based on quantitative methodology, 27% had a qualitative approach while 31% had mixed analysis. There were sixteen assessment methods presented. The most common type of assessment was Objective Structured Assessments (27%), which included Objective Structured Assessment of Technical Skill (OSATS) (23%) and Objective Structured Assessment of Non-Technical Skill (4%). Procedure Based Assessment (PBA) (23%) and Entrustability Scales (23%) were also prevalent.

**Conclusions:** This scoping review has identified a range of different assessment methods. The assessment methods with a higher volume and level of supporting evidence were OSATS, PBAs and Entrustability Scales. There was a lower volume and level of supporting evidence found within this review for the remaining assessment methods.

Crown Copyright © 2022 Published by Elsevier Ltd on behalf of Royal College of Surgeons of Edinburgh (Scottish charity number SC005317) and Royal College of Surgeons in Ireland.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

Reliable assessment is a vital part of postgraduate surgical training. A competency based approach to medical education originating in the 1970s has come to the foreground in more recent years.<sup>1</sup> The UK General Medical Council (GMC) and

Surgical Royal Colleges combined with the American Board of Surgery (ABS) and the North American Accreditation Council for Graduate Medical Education (ACGME) have identified the need for competency-based assessment within training.<sup>2–4</sup>

In the UK the “Modernising Medical Careers” (MMC) introduced competency based assessment for postgraduate medical training progression to completion of training. This

\* Corresponding author.

E-mail address: [laurenhackney3@gmail.com](mailto:laurenhackney3@gmail.com) (L. Hackney).

<https://doi.org/10.1016/j.surge.2022.01.009>

1479-666X/Crown Copyright © 2022 Published by Elsevier Ltd on behalf of Royal College of Surgeons of Edinburgh (Scottish charity number SC005317) and Royal College of Surgeons in Ireland. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Table 1 – Summary of articles reviewed.**

Assessment tool	Definition	No. of articles	First author	Country	Year of publication	Aim	Type of research
OSATS: Objective Structured Assessment of Technical Skill	Objective skills assessment using bench model simulation made up of two components: <ul style="list-style-type: none"> <li>• checklist specific to procedure</li> <li>• global rating scale<sup>20</sup></li> </ul>	7	Szasz <sup>21</sup>	USA	2017	To compare ratings given by internal raters vs external rater (performance scores in operating room for laparoscopic cholecystectomy)	Prospective Observational Study
			Beard <sup>19</sup>	UK	2011	Compare user satisfaction, acceptability, reliability and validity of PBA, OSATS and NOTSS	Prospective Observational Study
			Hove <sup>18</sup>	UK	2010	Review of current evidence for objective assessment methods of technical surgical skills	Systematic Review (Non RCT)
			Sanfey <sup>21</sup>	UK	2014	Review on the assessment of technical, non-technical and professional skills	Literature Review
			Niitsu <sup>20</sup>	Japan	2013	Demonstrate validity and accuracy of OSATS for assessing surgical skills in the operating room	Retrospective Observational Study
			Beard <sup>22</sup>	UK	2008	Review of assessments of surgical skills of trainees in the UK	Expert Opinion
			Moorthy <sup>17</sup>	UK	2003	Explore the available assessment methods, establish validity and reliability and possibility of using methods on the basis of available evidence	Expert Opinion
EPA's: Entrustable Professional Activities	Clinical activities that surgical trainees perform with minimal, or no supervision, which allows a competency-based assessment of trainees and allows a decision on level of supervision <sup>10,26</sup>	6	Sanfey <sup>21</sup>	UK	2014	Review on the assessment of technical, non-technical and professional skills	Literature review
			Brasel <sup>26</sup>	USA	2019	Describe the development and initial pilot process of implementing EPAs for general surgical trainees.	Pilot Observational Study
			Steiman <sup>7</sup>	USA	2018	Assess competency in surgical residents using SEPA framework (Breast and Gallbladder caseload)	Pilot Observational Study
			Moore <sup>10</sup>	Australia	2017	Describe implementation of EPAs as a form of competency-based assessment for postgraduate surgical trainees	Curriculum Presentation
			Reckman <sup>27</sup>	USA	2016	Outline how entrustability scales may bridge the gap between assessment of supervisors and WBAs	Expert Opinion
Wagner <sup>4</sup>	USA	2018	Use of EPAs to assess surgical residents' development. Identify differences in perspectives of residents and faculty	Survey			

*(continued on next page)*

**Table 1 – (continued)**

Assessment tool	Definition	No. of articles	First author	Country	Year of publication	Aim	Type of research
PBA: Procedure-Based Assessment	Assessment of technical, operative and professional skills within a procedure or parts of a procedure. Uses two principles: <ul style="list-style-type: none"> <li>• Competencies within 5 domains</li> </ul> Global assessment divided into 8 levels of ability/competency <sup>2</sup>	6	Shalhoub <sup>3</sup>	UK	2014	Explore evidence underlying introduction and ongoing use of WBAs in surgical training	Literature Review
			Beard <sup>19</sup>	UK	2011	Compare user satisfaction, acceptability, reliability and validity of PBA, OSATS and NOTSS	Prospective Observational Study
			Marriot <sup>23</sup>	UK	2011	Evaluate the validity, reliability and acceptability of PBA in postgraduate surgical education	Prospective Observational Study
			Eardley <sup>24</sup>	UK	2013	Report initial experiences of WBA systems	Clinical Review: Expert Opinion
			Beard <sup>22</sup>	UK	2008	Review of assessments of surgical skills of trainees in the UK	Clinical Review: Expert Opinion
			Pereira <sup>11</sup>	UK	2009	Evaluate user satisfaction with ISCP	Survey
WBA: Workplace based assessment	Formative assessment tool aimed to provide feedback between trainers and trainees to support learning, encompass assessment and provide a reference point for competence. An umbrella term for a group of assessment <sup>2</sup>	4	Shalhoub <sup>3</sup>	UK	2014	Explore evidence underlying introduction and ongoing use of WBAs in surgical training	Literature Review
			Eardley <sup>24</sup>	UK	2013	Report initial experiences of WBA systems	Retrospective Observational Study
			Gaunt <sup>36</sup>	UK	2018	Explore trainees' behavior in seeking feedback in postgraduate surgical	Qualitative study (Focus Group)
			Pereira <sup>11</sup>	UK	2009	Evaluate user satisfaction with ISCP	Survey
Global Rating Scales	Scales created and validated to give a global score of a trainee's performance <sup>2,3</sup>	3	Shah <sup>8</sup>	USA	2018	Compare number of cases in logbook versus operative proficiency/competence	Prospective Observational Study
			Ponton-Carss <sup>25</sup>	Canada	2016	Evaluate validity of evidence for surgical skills examination and examine relationship between technical and non-technical skills	Prospective Observational Study
			Niitsu <sup>20</sup>	Japan	2013	Demonstrate validity and accuracy of OSATS for assessing surgical skills in the operating room	Retrospective Observational Study
DOPS: Direct Observation of Procedures	DOPS: Assessment of technical, operative and professional skills within a range of basic procedures, parts of procedures from diagnostic to interventional procedures <sup>2</sup>	3	Shalhoub <sup>3</sup>	UK	2014	Explore evidence underlying introduction and ongoing use of WBAs in surgical training	Literature Review
			Eardley <sup>24</sup>	UK	2013	Report initial experiences of WBA systems	Retrospective Observational Study
			Pereira <sup>11</sup>	UK	2009	Evaluate user satisfaction with ISCP	Survey

**Table 1 – (continued)**

Assessment tool	Definition	No. of articles	First author	Country	Year of publication	Aim	Type of research
CEX: Clinical Evaluation Exercises	CEX <sup>2</sup> : Observation of trainee interaction in a clinical encounter <sup>2</sup>	3	Shalhoub <sup>3</sup>	UK	2014	Explore evidence underlying introduction and ongoing use of WBAs in surgical training	Literature Review
			Eardley <sup>24</sup>	UK	2013	Report initial experiences of WBA systems	Retrospective Observational Study
			Pereira <sup>11</sup>	UK	2009	Evaluate user satisfaction with ISCP	Survey
CBD: Case Based Discussions	CBD: Structured, in depth discussion of a clinical case <sup>2</sup>	3	Shalhoub <sup>3</sup>	UK	2014	Explore evidence underlying introduction and ongoing use of WBAs in surgical training	Literature Review
			Eardley <sup>24</sup>	UK	2013	Report initial experiences of WBA systems	Retrospective Observational Study
			Pereira <sup>11</sup>	UK	2009	Evaluate user satisfaction with ISCP	Survey
NOTTS: Non-Technical Skills for Surgeons	Behavioural rating system developed to assess non-technical clinical skills; including behaviours such as situation awareness, communication and teamwork, decision-making and leadership <sup>19,39</sup>	2	Beard <sup>19</sup>	UK	2011	Compare user satisfaction, acceptability, reliability and validity of PBA, OSATS and NOTSS	Prospective Observational Study
			Beard <sup>22</sup>	UK	2008	Review of assessments of surgical skills of trainees in the UK	Expert Opinion
Zwisch Scale	Four stage model framework that allows trainees and trainers to be guided through interactions in theatre and subsequent assessment that results in a designated level of autonomy for procedures. 4 stages: (1) Show and Tell; (2) Smart Help; (3) Dumb Help; (4) No Help <sup>8,28</sup>	1	Shah <sup>8</sup>	USA	2018	Compare number of cases in logbook versus operative proficiency/competence	Prospective Observational Study
OPR: Operative Performance Rating	Rating scale of 1–5 that incorporates procedure specific items alongside general items to score a trainee during a specific procedure <sup>31</sup>	1	Williams <sup>31</sup>	USA	2015	Determine if single item overall OPR provides a valid and stable appraisal of residents' operative performance	Retrospective Observational Study
OSPRE: Objective Structured Performance-Related Examination	Directly observes structured assessment of both technical and non-technical skills in a simulated examination setting <sup>25</sup>	1	Ponton-Carss <sup>25</sup>	Canada	2016	Evaluate validity of evidence for surgical skills examination and examine relationship between technical and non-technical skills	Prospective Observational Study
MOMS: Multiple Objective Measures of Skill	Based around OSCE format used in undergraduate medicine. Simulated task-based examination to assess skills <sup>32</sup>	1	Mackay <sup>32</sup>	UK	2003	Validation of a 6-task skills examination for junior surgical trainees including feasibility of tools	Validation of Assessment tool
Omni	Novel assessment tool for surgical technique that uses objective measures and scoring algorithms to evaluate performance. Aimed to be used for repeated evaluation throughout training to monitor progression and competency <sup>37</sup>	1	Cox <sup>37</sup>	USA	2018	Description of 'Omni' assessment tool for assessment of performance in competency-based training	Validation of Assessment tool
Portfolio	Collection of documents collected to provide evidence <sup>2,35</sup>	1	Webb <sup>35</sup>	USA	2012	Evaluate trainee and faculty perspectives on satisfaction, compliance and educational value of portfolio	Qualitative (Survey/Focus Group)

(continued on next page)

**Table 1 – (continued)**

Assessment tool	Definition	No. of articles	First author	Country	Year of publication	Aim	Type of research
Other	Informal Assessment	3	Drossard <sup>28</sup>	Germany	2019	Provide overview of current programme for surgical training in Germany	Expert opinion
			Chen <sup>29</sup>	USA	2017	Explore approaches of surgical educators for assessment of surgical residents' readiness/ autonomy in the operating room.	Qualitative (Interviews)
			Salim <sup>30</sup>	Canada	2020	Explore how trust was constructed between surgeons and residents in the operating room	Qualitative (Interviews)

replaced a traditional apprenticeship or time based training system.<sup>3</sup> There not only has been a shift from time spent in training but also less emphasis on 'number of procedures' and a shift towards demonstrating competence.<sup>3</sup> Competency based assessment can be described as a framework for ensuring trainees are capable of doing the work required of them; the verification of the knowledge, skills and behavioural characteristics needed for independent practice.<sup>1,3,5</sup> This observable ability is a task or attribute that needs to be assessed and measured against a common standard.<sup>1,6</sup>

The importance of competency is clear, however, there is concern that surgical trainees are lacking competence towards the end of training.<sup>4,7,8</sup> The suggestion is that trainees nearing the end of their training, or following completion of fellowship appointments, may be deficient in the clinical and practical skills required of a consultant in independent practice.<sup>4,7,8</sup>

Within the UK, trainees will undergo the Fellowship of the Royal College of Surgeons (FRCS) exam in order to complete training where they are then eligible to become a consultant. The FRCS exam consists of two parts; examination through multiple choice questions (MCQs) and a clinical oral examination. In addition, evidence is acquired from the Intercollegiate Surgical Curriculum Programme (ISCP) with regards to the trainees workplace-based assessments to provide evidence of surgical skill-set and competence to manage surgical patients.<sup>9</sup> Therefore, the current published literature suggests that the use of assessments throughout training is not only vital for progression but also completion of training.

Only rigorous evaluation of competence will result in the production of safe consultant surgeons that are able to provide the best care for patients. There is a suggestion that workplace-based assessments place a burden on both trainees and trainers.<sup>10,11</sup> In addition trainees have expressed concerns about the detrimental effect of misuse of assessments. Both of these factors need to be considered when determining the validity and reliability of trainee competency based assessments.<sup>11</sup>

Current pressures on training opportunities and assessment have been aggravated by the pressures of the COVID-19 pandemic. The reduction in clinical exposure, less operating time, more consultant led cases, redeployment and service changes have already had a significant impact on the training curriculum within surgery.<sup>12–14</sup> The consideration of

introducing virtual assessments as well as increasing assessments via simulation need to be a serious consideration given the fact that there is no clear end point for the changes occurring due to Covid-19.<sup>12–14</sup>

## Methods

This scoping review aims to evaluate the volume of existing evidence surrounding objective assessment of postgraduate general surgical trainees. The scoping review was conducted and reported in line with the PRISMA extension for scoping reviews.<sup>15</sup>

### Protocol and registration

The scoping review was registered with the Open Sciences Framework: DOI 10.17605/OSF.IO/F7ZK9.

### Study design

A systematic scoping review methodology was used to address the research question according to a predefined protocol. Due to the broad and heterogeneous nature of the scoping review there was no evaluation of evidence quality.<sup>16</sup>

### Eligibility criteria

Studies describing the evidence surrounding medical education relating to objective assessment of general surgical trainees were eligible. Eligible papers had to be full text articles within the last 20 years. This would permit exploration of previous research methodologies as well as processes relating to the development, introduction and evolution of an evidence-based surgical curriculum. Articles from all countries were eligible for inclusion, but they had to be available in English language. Any articles relating to undergraduate assessment methods, other medical specialities or specialty specific publications (e.g. orthopaedics, ENT, Vascular) were excluded. The scoping review did not explore the evidence surrounding MCQs or the final UK examination (FRCS).

### Information sources

The search strategy included searches on PubMed, Ovid Medline and Web of Sciences databases as well as grey literature such as reports, protocols and guidelines. Two independent researchers (LH and SON) carried out searches with the same search criteria. Concept mapping was used (see [Appendix I](#)) to design the search strategy.

### Data charting

The data chart was designed in a two-step process. A data chart proforma was created by a single investigator (LH) who initially reviewed 5 articles. Following this the proforma was subsequently analysed to permit refinement of the data to produce the following final headings of the data chart: *Author; Year of Publication; Journal; Location (Country); Aim of Study; Methods of Study; Findings; Assessment methods explored; Conclusions; Limitations of study and Further research*. No assumptions were made when charting data. A second researcher (SON) reviewed the final data charts with narrative summaries. Mendeley reference manager was used to formulate a systematic approach of reviewing each article and entering the information into the data chart.

### Synthesis of results

Basic numeric analysis of the charted data was performed. Narrative synthesis was then performed.

### Assessment methodology definitions

A list of definitions can be found in [Table 1](#).

## Results

### Characteristics

A total of 343 papers were reviewed. 26 were eligible for inclusion ([Appendix II](#)). All articles were published between 2003 and 2020. There was an increase in the number of articles related to assessment within surgical training from 2008 onwards where 92% of articles were published from 2008 onwards and 50% of all articles within the last five years. The articles originated from 6 different countries. The majority of publications originated in the United Kingdom (42%), followed closely by the United States of America (38%). In addition, a small number were published in Canada (8%), Japan (4%), Germany (4%) and Australia (4%). UK publications were predominantly between 2008 and 2014 while the USA had a later predominance between 2015 and 2018. The studies included a mix of qualitative and quantitative methods. 42% were on a quantitative nature, 27% had a qualitative approach while 31% had mixed analysis.

There were sixteen different assessment methods presented. The most common type of assessment was Objective Structured Assessments (27%) including Objective Structured Assessment of Technical Skill (OSATS) (23%) and Objective Structured Assessment of Non-Technical Skill (4%). Procedure

Based Assessment (PBA) (23%) and Entrustability Scales (23%) were also prevalent. The remaining assessments included Global rating scales, Workplace-Based Assessments (WBA), Direct Observation of Procedures (DOPs), Clinical Evaluation Exercises (CEX), Case Based Discussion (CBD), Portfolio, Zwisch Scale, Multiple Objective Measures of Skill (MOMs), Non-Technical Skills for Surgeons (NOTTS), Operative Performance Rating (OPR), Omni, Portfolio and other informal assessment methods ([Table 1](#)).

## Discussion

There is overarching agreement within the literature that the use of assessment is vital for the training of a safe and competent surgeon, and in support of that there is a need for continued vigorous research into the different assessment methods to ensure they are valid and reliable.<sup>2–4,7,8,11,17,21,33</sup>

This scoping review identified sixteen different assessment methods for general surgical trainees.

### OSATS: Objective Structured Assessment of Technical Skill

This type of assessment method was the most frequently described in the literature. OSATS consists of a procedure specific checklist alongside a global rating scale and is used as an objective measurement of skills.<sup>17,18</sup> This type of assessment lends itself well to general surgical training as it aims to assess technical skills within a ‘procedure specific’ setting. Two literature reviews by Moorthy et al.<sup>17</sup> and Hove et al.<sup>18</sup> concluded that OSATS have been shown to be reliable, as well as having construct validity. However, OSATS have mainly been researched in simulated settings and currently lack evidence within an operating theatre setting.<sup>18</sup> Conversely, Beard et al.<sup>19</sup> performed a prospective observational comparison of workplace-based assessments where they reported that OSATS had a lower reliability when compared to PBA. In addition, OSATS did not demonstrate construct validity in their study and had an overall lower user satisfaction. The majority of later studies included in this scoping review reported that OSATS were valid and reliable. However, such studies predominantly assessed procedural skills and were lacking in evidence-based assessment of non-technical skills.<sup>17–20</sup> Limiting factors for successful incorporation of OSATS methodology included resource, time and cost issues.<sup>17–20</sup>

### PBA: Procedure-Based Assessments

Another workplace-based assessment with a higher proportion of supporting research found in this scoping review were PBAs. PBAs are used within surgical training to assess technical, operative and professional skills whilst a trainee performs a procedure or part of a procedure.<sup>2</sup> Within this review, PBAs were largely shown to have a high reliability and validity. However, the published evidence suggested that a large number of such assessments are warranted to produce this high level of reliability and validity. An important finding relating to PBAs is that they were shown to be procedure specific. Consequently, one cannot presume that the competency within one skill or procedure is then transferable

globally across other skill-sets of the trainee.<sup>19,21–23</sup> When considering their use in practice it is fundamental that for PBAs to hold onto their reliability and validity they need to be numerous and varied.

### **EPA's: Entrustable Professional Activities**

The scoping review has disclosed a modest quantity of papers supporting entrustability tools such as Surgical Entrustable Professional Activities (SEPA) and Entrustable Professional Activities (EPAs). These assessment methodologies are typically used to assess how ready a trainee is to perform a specific task. Although such processes are often incorporated into a supervisor's natural teaching style, EPA's were shown to be a useful adjunct to objectively measure specific task competence. EPA's have been advocated as a tool that connects knowledge, theory and examination with clinical practice. Further, they have been shown to positively influence the user's behaviour around learning and assessment, as it broadens the context of the assessment and helps them focus on the end goal or outcome rather than a grade or score.<sup>4,10,21,24–26</sup> Ultimately, they can then be used to judge a trainee's readiness for independent practice, which in the end is the ultimate goal of competency-based assessment.

### **Global rating scales**

Within this scoping review a small quantity of evidence was available for global rating scales within surgical training. Limitations of these types of assessment include evaluator bias and difficulties with standardisation. Within this research, it has been concluded that these scales are able to discriminate well between trainees' level of training and are thought to be a reasonable and practical way of assessing progressive competency. Although the evidence surrounding incorporation of this type of assessment did appear supportive, published article volume was small.<sup>4,8,20,21,27</sup> It is important to consider that often global rating scales are incorporated into other assessment methods and may not be as extensively researched individually for that reason.

### **Assessment of autonomy**

Shah et al.<sup>8</sup> reported the utilisation of the Zwisch scale to assess trainees' autonomy. This is a validated scale used in the operating room by some institutions in North America which utilises a framework that trainees and trainers can use in theatre and during assessments which results in an overall level of autonomy. This is a useful assessment tool that is specific to surgical training. As part of the Zwisch Scale, goal setting is used to enhance training and learning, and it has been shown to improve trainee satisfaction and performance.<sup>8,28</sup> In addition, the use of entrustable professional activities (EPAs) has also been cited in supporting improved autonomy.<sup>4,21</sup> Overall, the literature suggests that graduated autonomy needs to be established to allow for progressive clinical and operative competence and this ideally would be done in an objective measured way, to ensure the safe graduation of independent surgeons at the end of training.<sup>4,8,21,29,30</sup> Chen et al.<sup>29</sup> proposed a set of "Performance-

Impression-Characteristics-Knowledge- Situation" questions to be used as standard prior to any observation of skills within a theatre setting. They found that the introduction of this concept contributed to both trainee and trainer development and increased autonomy.

### **OPR: Operative Performance Ratings**

A single paper discussed OPRs. OPRs are a set of rating instruments that are procedure specific methods of observing and evaluating performance. Within this paper OPR was only validated for Laparoscopic Cholecystectomy. More extensive research is warranted before it could be recommended as an applicable for other procedures.<sup>31</sup>

### **OSPPE: Objective Structured Performance-Related Examination (OSPPE)**

OSPPE is a structured assessment similar to the well-established Objective Structured Clinical Examination (OSCE). It consists of a number of stations that each have specific task as well as a global rating score.<sup>24</sup> Although OSCE is a well-established form of evaluation within medical education, this review revealed a lack of supporting evidence for the use of OSPPE in general surgery trainee assessments. Ponton-Carss et al.<sup>25</sup> performed an experimental study looking at the use of OSPPE in technical and non-technical skills and reported that OSPPE had good internal consistency (statistical measure that reflects the reliability of a scoring system). Interstation reliability, however, was low and therefore a certain score in one station did not predict a similar score in another, reiterating the need for multiple assessments. Conclusions echoed the need for further research.<sup>25</sup>

### **MOMS: Multiple Objective Measures of Skill**

Mackay et al.<sup>32</sup> presented a paper aiming to validate the Multiple Objective Measures of Skill (MOMS) assessment tool. This is a relatively new approach to assessing trainees; whereby multiple skills are assessed in an examination style setting, similar to that of an OSCE, for the assessment of technical ability during a task. MOMS is aimed to be an assessment targeted towards practical specialities and therefore lends itself towards general surgical training. Although this study demonstrated good reliability, the authors only evaluated a small sample size. This scoping review identified no other supporting evidence regarding the use of MOMS in technical ability assessment.<sup>32</sup>

### **DOPS, CEX and CBD: Direct Observation of Procedures (DOPS), Clinical Evaluation Exercises (CEX) and Case Based Discussions (CBD)**

This scoping review identified three literature reviews evaluating these methodologies with the most recent publication from 2014. These assessment methods are the predominant tools used within ISCP for UK trainees. The guidance to which has not been updated since 2015, suggesting a considerable lack of recent evidence supporting these continued implementation of these assessment methods within surgery,

despite their ongoing utilisation for UK surgical trainee assessments.<sup>2</sup> DOPS are used to assess technical, operative and professional skills, with key steps being assessed throughout.<sup>2</sup> In addition it is not clear how the key steps or particular procedures have been chosen and may be based on expert opinion without being subjected to scrutiny.<sup>3</sup> CBD is a structured, in depth discussion of a clinical case.<sup>2</sup> Eardley et al.<sup>24</sup> found that there was no apparent relation between stage of training and rating given in either DOPS or CBD.<sup>24</sup> CEX involves observing and assessing a clinical encounter.<sup>2</sup> In Shalhoub et al.'s<sup>3</sup> review, CEX was found to be the only assessment with some validity. Pereira and Dean<sup>11</sup> described how CEX was shown to be reliable but only in strict exam conditions which is far from what is experienced in the clinical setting. All three literature reviews did not significantly support the incorporation of all three assessment methodologies in surgical assessments. Similar to previous assessment tools above, more research is warranted to fully determine if such assessment modalities are to remain beneficial to the general surgical trainee.<sup>3,11,24</sup> It is also important to consider that initial studies into such workplace-based assessments were carried out to assess them as they were initially designed, in the context of a learning tool rather than an assessment tool of competency.<sup>3</sup>

### **NOTSS: Non-Technical Skills for Surgeons**

Only two papers by Beard et al.<sup>19,22</sup> specifically evaluated non-technical skill assessment methods in surgical training. Although there appears to be an abundance of authors mandating the need for more research into behavioural assessment of surgical trainees, the current published literature in evaluating the non-technical skills of general surgery trainees is lacking.<sup>20–22,25,32</sup> More research is warranted in particular the impact of NOTSS assessment of general surgery trainees and how such assessment methodology could potentially complement the practical assessment methods that have already been established and studied.

### **Feedback**

Valued assessments require consistent and rigorous assessors who are able to provide appropriate and timely feedback. There is still a tendency for assessors to hesitate to provide very direct or robust assessments particular during the provision of negative feedback. Such differences in assessor's leniency and stringency have been previously reported and were included in this scoping review.<sup>33,34</sup> Several studies suggested a requirement to incorporate training of surgeons on how to perform trainee assessments and how feedback is needed for these assessors to minimise such variances in assessment processes, in addition to the consideration of using external raters. Other researchers have also documented that surgical trainers do not always have training into how to perform assessments or provide feedback and therefore, assessments may not be standardised or always useful.<sup>20,33–36</sup> Feedback seeking behaviours have also been criticised.<sup>36</sup> Motives for seeking feedback can determine how often, when and with whom trainees will seek feedback. Additionally, trainees who view the feedback as a process for

development and a learning opportunity are more likely to seek feedback, whereas those who see it as an assessment of their learning are less likely to pursue it.<sup>36</sup> The importance of timely feedback is also highlighted, as immediate feedback is when the most valuable learning happens.<sup>20</sup>

### **Other**

Cox et al.'s<sup>37</sup> pilot study evaluated the a new “Omni” assessment tool. They presented a task specific tool that assessed speed, accuracy and quality of candidate performance which were then scaled from 0 to 10. The “Omni” tool showed promise after demonstrating both construct validity and internal consistency when evaluating surgical technical performance. Webb and Merkle<sup>35</sup> evaluated the use of trainee portfolio in surgical training. Their qualitative study explored trainee and faculty perspectives on the value of portfolio and concluded that quality of assessments through portfolio are variable for many reasons and therefore the quality of evidence within portfolios may not be reliable.<sup>34</sup>

### **Limitations**

The purpose of a scoping review was to establish the extent, nature and characteristics of available assessment methodologies for general surgical training published in the literature. The limitations of a scoping review are recognised in that it is unable to make comparisons between the studies. However, accepting the heterogeneity of the research identified, this review has allowed for the study question to be answered.

### **Further research**

The articles reviewed here clearly highlight the importance of competency. However, the majority concluded that there is an ongoing need for further research into the area, in particular with regards to non-technical skills.<sup>4,8,10,11,20–22,25,31,38</sup>

---

## **Conclusion**

The assessment methods with a higher volume of supporting evidence were OSATS, PBAs and Global Rating Scales. There was a lower volume of supporting evidence found within this review for the remaining assessment methods. OSATS were reported to be reliable and that they are a useful tool in assessing competence; however, their use was mainly validated within a simulated setting. PBAs were shown to have high reliability; however, one consideration was that they were ‘procedure specific’ and so cannot be relied as a general assessment of skills covering all areas. Of note these assessments can be somewhat time consuming, when carried out correctly, within an already time pressured service. This pressure is intensified by recommendations that there needs to be numerous assessments to ensure reliability and validity within many of the assessment methods.

A particular area of interest, which was described throughout this review, was the assessment of non-technical skills. The importance of behavioural skills was clearly identified and the associated lack of in-depth research, was noted. This type of assessment would, arguably, be the most

important area to focus in future. Taking into consideration the impact of COVID-19 and the potential need for increased virtual assessments this may be any area of particular focus in order to ensure the behavioural skills of surgical trainees continues to be evaluated.

### Declaration of competing interest

There are no competing interests. In addition there was no external funding provided in the course of this research.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.surge.2022.01.009>.

### REFERENCES

- Frank J, Snell L, Cate O, Holmboe E, Carraccio C, Swing S, et al. Competency-based medical education: theory to practice. *Med Teach* 2010;**32**(8):638–45. <https://doi.org/10.3109/0142159X.2010.501190>.
- Intercollegiate Surgical Curriculum Programme. *Background*. [https://www.iscp.ac.uk/home/start\\_background.aspx](https://www.iscp.ac.uk/home/start_background.aspx). [Accessed 19 April 2021].
- Shalhoub J, Vesey A, Fitzgerald J. What evidence is there for the use of workplace-based assessment in surgical training? *J Surg Educ* 2014;**71**(6):906–15. <https://doi.org/10.1016/j.jsurg.2014.03.013>.
- Wagner J, Lewis C, Tillou A, Agopian V, Quach C, Donahue T, et al. Use of entrustable professional activities in the assessment of surgical resident competency. *JAMA Surg* 2018;**153**(4):335–43. <https://doi.org/10.1001/jamasurg.2017.4547>.
- Weller J, Viren N, San Diego R. Systematic review and narrative synthesis of competency-based medical education in anaesthesia. *Br J Anaesth* 2020;**124**(6):748–60. <https://doi.org/10.1016/j.bja.2019.10.025>.
- Holmboe E, Long D, Swing S, Long D, Frank J, for the International CBME Collaborators. The role of assessment in competency-based medical education. *Med Teach* 2010;**32**(8):676–82. <https://doi.org/10.3109/0142159X.2010.500704>.
- Steiman J, Sullivan S, Scarborough J, Wilke L, Pugh C, Bell R, et al. Measuring competence in surgical training through assessment of surgical entrustable professional activities. *J Surg Educ* 2018;**75**(6):1452–62. <https://doi.org/10.1016/j.jsurg.2018.05.004>.
- Shah D, Haisch C, Noland S. Case reporting, competence, and confidence: a discrepancy in the numbers. *J Surg Educ* 2018;**75**(2):304–12. <https://doi.org/10.1016/j.jsurg.2018.01.007>.
- Joint Committee on Intercollegiate Examinations. *Intercollegiate Specialty Examination in General Surgery Syllabus Blueprint*. <https://www.jcie.org.uk/content/content.aspx?ID=14>, 2016. [Accessed 19 April 2021].
- Moore D, Young C, Hong J. Implementing entrustable professional activities: the yellow brick road towards competency-based training? *ANZ J Surg* 2017;**87**(2):1001–5. <https://doi.org/10.1111/ans.14120>.
- Pereira E, Dean B. British surgeons' experiences of mandatory online workplace-based assessment. *J R Soc Med* 2009;**102**(7):287–93. <https://doi.org/10.1258/jrsm.2009.080398>.
- James H, Pattison G. Disruption to surgical training during Covid-19 in the United States, United Kingdom, Canada, and Australasia: a rapid review of impact and mitigation efforts. *J Surg Educ* 2020. <https://doi.org/10.1016/j.jsurg.2020.06.020>.
- White E, Shaughnessy M, Esposito A, Slade M, Korah M, Yoo P. Surgical education in the time of COVID: understanding the early response of surgical training programs to the novel coronavirus pandemic. *J Surg Educ* 2020. <https://doi.org/10.1016/j.jsurg.2020.07.036>.
- Okland T, Pepper J, Valdez T. How do we teach surgical residents in the COVID-19 era? *J Surg Educ* 2020;**77**(5):1005–7. <https://doi.org/10.1016/j.jsurg.2020.05.030>.
- Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;**169**(7):467–73. <https://doi.org/10.7326/M18-0850>.
- OCEDM Levels of Evidence Working Group. *The Oxford 2011 Levels of Evidence*. <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebml-levels-of-evidence>. [Accessed 19 April 2021].
- Moorthy K, Munz Y, Sarker S, Darzi A. Objective assessment of technical skills in surgery. *BMJ* 2003;**327**(7422):1032–7. <https://doi.org/10.1136/bmj.327.7422.1032>.
- Hove P, Tuijthof G, Verdaasdonk E, Stassen L, Dankelman J. Objective assessment of technical surgical skills. *BJS* 2010;**97**(7):972–87. <https://doi.org/10.1002/bjs.7115>.
- Beard J, Marriot J, Purdie H, Crossley J. Assessing the surgical skills of trainees in the operating theatre: a prospective observational study of the methodology. *Health Technol Assess* 2011;**15**(1). <https://doi.org/10.3310/hta15010>.
- Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Tomato J, Sugiyama Y, et al. Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. *Surg Today* 2013;**43**:271–5. <https://doi.org/10.1007/s00595-012-0313-7>.
- Sanfey H. Assessment of surgical training. *Surgeon* 2014;**12**(6):350–6. <https://doi.org/10.1016/j.surge.2014.07.001>.
- Beard J. Assessment of surgical skills of trainees in the UK. *Ann R Coll Surg Engl* 2008;**90**(4):282–5. <https://doi.org/10.1308/003588408X286017>.
- Marriot J, Purdie H, Crossley J, Beard J. Evaluation of procedure-based assessment for assessing trainees' skills in the operating theatre. *BJS* 2010;**98**(3):450–7. <https://doi.org/10.1002/bjs.7342>.
- Eardley I, Bussey M, Woodthorpe A, Munsch C, Breard J. Workplace-based assessment in surgical training: experiences from the Intercollegiate Surgical Curriculum Programme. *ANZ J Surg* 2013;**83**(2013):448–53. <https://doi.org/10.1111/ans.12187>.
- Ponton-Carss A, Kortbeek J, Ma I. Assessment of technical and nontechnical skills in surgical residents. *Am J Surg* 2016;**212**(5):1011–9. <https://doi.org/10.1016/j.amjsurg.2016.03.005>.
- Brasel K, Klingensmith M, Englander R, Grambau M, Buyske J, Sarosi G, et al. Entrustable professional activities in general surgery: development and implementation. *J Surg Educ* 2019;**76**(5):1174–86. <https://doi.org/10.1016/j.jsurg.2019.04.003>.
- Rekman J, Gofton W, Dudek N, Gofton T, Hamstra S. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med* 2016;**91**(2):186–90. <https://doi.org/10.1097/ACM.0000000000001045>.
- DaRosa D, Fryer J. A theory-based model for teaching and assessing residents in the operating room. *J Surg Educ* 2013;**70**(1):24–30. <https://doi.org/10.1016/j.jsurg.2012.07.007>.
- Chen X, Sullivan A, Alseidi A, Kwakye G, Smink D. Assessing residents' readiness for or autonomy: a qualitative descriptive

- study of expert surgical teachers' best practices. *J Surg Educ* 2017;**74**(6):15–21. <https://doi.org/10.1016/j.jsurg.2017.06.034>.
30. Salim S, Govaerts M, White J. The construction of surgical trust: how surgeons judge residents' readiness for operative independence. *Ann Surg* 2020;**271**(2):391–8. <https://doi.org/10.1097/SLA.0000000000003125>.
31. Williams R, Verhukst S, Mellinger J, Dunnington G. Is a single-item operative performance rating sufficient? *J Surg Educ* 2015;**72**(6):212–7. <https://doi.org/10.1016/j.jsurg.2015.05.002>.
32. Mackay S, Datta V, Chang A, Shah J, Kneebone R, Darzi A. Multiple objective measures of skill (MOMS): a new approach to the assessment of technical ability in surgical trainees. *Ann Surg* 2003;**238**(2):291–300. <https://doi.org/10.1097/01.sla.0000080829.29028.c4>.
33. Szasz P, Louridas M, Harris K, Grantcharov T. Strategies for increasing the feasibility of performance assessments during competency-based education: subjective and objective evaluations correlate in the operating room. *Am J Surg* 2017;**214**(2):365–72. <https://doi.org/10.1016/j.amjsurg.2016.07.017>.
34. Murphy M, Seneviratne R, Remers O, Davis M. 'Hawks' and 'doves': effect of feedback on grades awarded by supervisors of student selected components. *Med Teach* 2009;**31**(10):484–8. <https://doi.org/10.3109/01421590903258670>.
35. Webb T, Merkley T. An evaluation of the success of a surgical resident learning portfolio. *J Surg Educ* 2012;**69**(1):1–7. <https://doi.org/10.1016/j.jsurg.2011.06.008>. [Accessed 19 April 2021].
36. Gaunt A, Markham D, Pawlikowska T. Exploring the role of self-motives in postgraduate trainees' feedback-seeking behaviour in the clinical workplace: a multicentre study of workplace-based assessments from the United Kingdom. *Acad Med* 2018;**93**(10):1576–83. <https://doi.org/10.1097/ACM.0000000000002348>. [Accessed 19 April 2021].
37. Cox M, Risucci D, Gilmore B, Nag U, Turner M, Sprinkle S, et al. Validation of the Omni: a novel, multimodality, and longitudinal surgical skills assessment. *J Surg Educ* 2018;**75**(6):e218–28. <https://doi.org/10.1016/j.jsurg.2018.10.012>.
38. Drossard S. Structured surgical residency training in Germany: an overview of existing training programs in 10 surgical subspecialties. *Innov Surg Sci* 2019;**4**(1):15–24. <https://doi.org/10.1515/iss-2018-0033>.
39. Yule S, Paterson Brown S. Surgeons' non-technical skills. *Surg Clin North Am* 2012;**92**(1):37–50. <https://doi.org/10.1016/j.suc.2011.11.004>.