COACTION
PUBLISHING

Tools used to assess medical students competence in procedural skills at the end of a primary medical degree: a systematic review

Marie C. Morris*, Tom K. Gallagher and Paul F. Ridgway

Department of Surgery, Trinity Centre for Health Sciences, Tallaght Hospital, Dublin, Ireland

The objective was to systematically review the literature to identify and grade tools used for the end point assessment of procedural skills (e.g., phlebotomy, IV cannulation, suturing) competence in medical students prior to certification. The authors searched eight bibliographic databases electronically – ERIC, Medline, CINAHL, EMBASE, Psychinfo, PsychLIT, EBM Reviews and the Cochrane databases. Two reviewers independently reviewed the literature to identify procedural assessment tools used specifically for assessing medical students within the PRISMA framework, the inclusion/exclusion criteria and search period. Papers on OSATS and DOPS were excluded as they focused on post-registration assessment and clinical rather than simulated competence. Of 659 abstracted articles 56 identified procedural assessment tools. Only 11 specifically assessed medical students. The final 11 studies consisted of 1 randomised controlled trial, 4 comparative and 6 descriptive studies yielding 12 heterogeneous procedural assessment tools for analysis. Seven tools addressed four discrete pre-certification skills, basic suture (3), airway management (2), nasogastric tube insertion (1) and intravenous cannulation (1). One tool used a generic assessment of procedural skills. Two tools focused on postgraduate laparoscopic skills and one on osteopathic students and thus were not included in this review. The levels of evidence are low with regard to reliability – $\kappa = 0.65$–0.71 and minimum validity is achieved – face and content. In conclusion, there are no tools designed specifically to assess competence of procedural skills in a final certification examination. There is a need to develop standardised tools with proven reliability and validity for assessment of procedural skills competence at the end of medical training. Medicine graduates must have comparable levels of procedural skills acquisition entering the clinical workforce irrespective of the country of training.

Keywords: *competence*; *competence assessment*; *assessment tools*; *clinical skills*; *surgical skills*; *technical skills*; *procedural skills*; *medical students*; *student physicians*; *medical trainees*; *final medical examination*

Introduction

Assessment and certification of doctors competence is a priority in health care provision today. Certifying bodies in developed countries require medical students to demonstrate proficient and appropriate use of procedural skills pre-registration (1). However, traditionally, the focus has been on knowledge-based competencies to the detriment of procedural tools validation. Identification of those who are not competent is paramount to patient safety. The aspiration today is to train to competence not time. The challenge remains to improve assessments to ultimately allow standards to be set and tested. It is essential that objective assessments of competence of procedural skills have reliability, validity and homogeneity (2). The current status of bespoke assessments of procedural skills at the point of certification is unknown. Institution-specific checklists are utilised routinely in pre-qualification procedural skills assessment. This leads to inequity of assessment and competence before qualification.

Objective Structured Clinical Examinations are the most common methods of assessing procedural skills (3). Routinely, procedural checklists are utilised in this assessment format. The assertion of competence in a simulated setting does not, however, guarantee the transferability of the skill to the workplace (unpublished data MM, PFR, 2012). The ongoing challenge to Medical Educators is in 'teaching for transfer' – facilitating students to be consistent in competent performance

irrespective of whether the setting is simulated or an emotionally charged clinical area. The intent of this review and research question was therefore 'Is there any evidence base to the procedural skills assessment tools routinely used as competency markers at the end of current medical training programmes?' This paper specifically seeks to systematically review the number and quality of current assessment tools measuring end point competence in procedural skills abilities of medical students prior to certification.

Methodology

Search strategy
This review was conducted utilising the electronic databases Journals Ovid for the most recent decade January 2000–October 2010. Medline, Embase, CINAHL, PschcINFO, ERIC, and the Cochrane Library were also searched. The initial search terms 'clinical skills assessment' yielded a total of 659 articles. These terms were further refined to 'assessment tools in clinical skills', 'technical skills', 'procedural skills' and 'surgical skills'. Further search terms 'in medical students', 'student physicians', 'medical trainees' and 'final examinations' were added. In Pubmed these refined search terms yielded 56 articles. In Eric, 43 articles were identified and in PsychINFO, 5 articles (Fig. 1). EMBASE, EBM reviews and Cochrane database revealed no further articles. Duplicate results were removed leaving 56 articles for review. In eligible articles where the assessment tool was not published the author was contacted directly to avail of the assessment tool (MM).
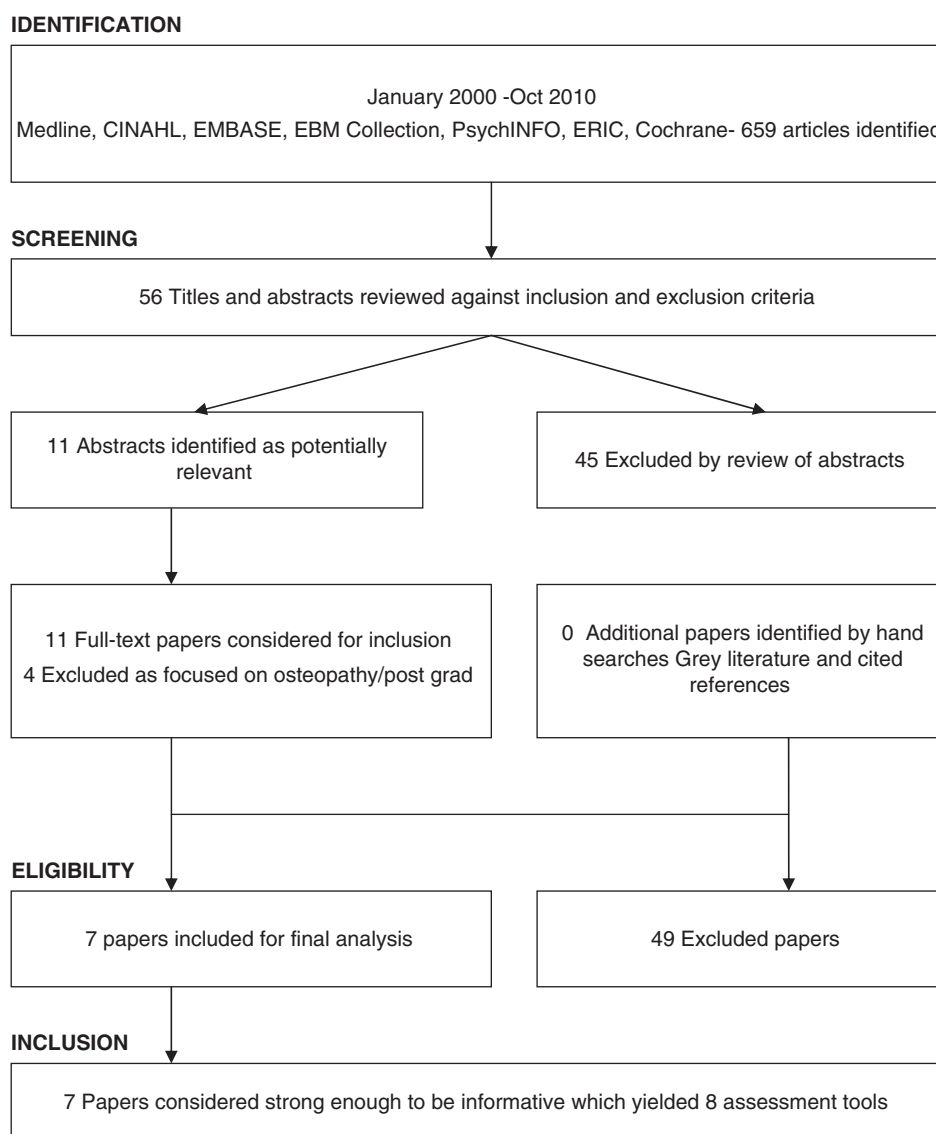
IDENTIFICATION

January 2000 -Oct 2010
Medline, CINAHL, EMBASE, EBM Collection, PsychINFO, ERIC, Cochrane- 659 articles identified

SCREENING

56 Titles and abstracts reviewed against inclusion and exclusion criteria

11 Abstracts identified as potentially relevant

45 Excluded by review of abstracts

11 Full-text papers considered for inclusion
4 Excluded as focused on osteopathy/post grad

0 Additional papers identified by hand searches Grey literature and cited references

ELIGIBILITY

7 papers included for final analysis

49 Excluded papers

INCLUSION

7 Papers considered strong enough to be informative which yielded 8 assessment tools

Fig. 1. Flowchart of search and selection strategy – PRISMA (21).

Inclusion criteria

The inclusion criteria covered all English language papers that included tools for assessing competence in procedural skills in medical students at the end of their primary education. This included papers where students were undergraduate and/or postgraduate on entering their medical training. Papers published in the last decade and papers on end point summative assessment by faculty on procedural skills abilities were also included. Workplace/clinically based assessments were included provided the assessment was summative and undertaken by faculty or senior clinical staff. Papers with evidence from other health care professions were included only if medical students were in the study sample as medical schools do not routinely utilise assessment tools from Allied Health Professionals preferring to devise institution specific checklists.

Exclusion criteria

Papers with the following content were excluded as they did not address the research question. Continuous assessment, formative assessment, portfolios, case or logbooks. Self-assessment or patient assessment. History taking or physical examination of patients. Papers on procedural skills of doctors in practice rather than pre-practice, e.g., laparoscopy. The systematic reviews of Miller and Archer (4), and Kogan (5) were excluded as Miller and Archer focused on 'workplace based assessment in postgraduates' and Kogan reports that 'tools used in simulated settings or assessing surgical procedural skills were excluded' from their review.

Term definitions

For the purpose of this paper the listed term definitions are accepted (Table 1).

Data extraction/synthesis

Validation of the selection process followed with two independent reviewers utilising a version of a standard Best Evidence in Medical Education – BEME Guide Coding Sheet (6), (Appendix 2). All papers were reviewed independently and assigned a score from 1 to 5 with regard to the strength of the research within the paper. Where there was a difference between reviewers grading of a paper a decisive final independent review occurred (PR).

Results

From the original 659 articles identified a refinement in search terms revealed 56 eligible articles. These abstracted articles were reviewed and 45 were deemed inappropriate against the inclusion and exclusion criteria. The remaining 11 heterogeneous studies were deemed suitable for review (Fig. 1). Of the 11 studies, two addressed postgraduate skills and were discarded (7, 8) and one study

Table 1. Term definitions

Competence: Defined for the purpose of this paper as 'Being functionally adequate, i.e., having sufficient knowledge, judgement, skill or strength for a particular duty' (3).

Assessment: Defined for the purpose of this paper as 'Any method that is used to better understand the current knowledge and skills that a student possesses' (1).

Procedural skills: Defined for the purpose of this study as 'skills requiring manual dexterity and health related knowledge which are aimed at the care of a single patient' (18).

Summative assessment: Assessments generally formative or summative. This paper is focused on summative assessment. In summative assessment findings and recommendations are designed to accumulate all relevant assessments for a go/no-go decision to be recommended for certification (4).

Validity: The concept of test validity is to what extent an assessment actually measures what it is intended to measure and permits appropriate generalisations about student's skills and abilities (5).

Reliability: Refers to consistency of a test over time over different cases (inter-case) and different examiners (inter-rater). We term measurements of reliability coefficients above 0.65 as moderate and above 0.85 as strong (5).

focused on osteopathic students (9) and was excluded. The remaining 11 studies which met the studies specific focus consisted of 1 randomised controlled trial (RCT), 4 comparative and 6 descriptive studies. These studies yielded 12 procedural assessment tools for analysis (Table 2). This review revealed two methods of procedural skills assessment in current pre registration practice namely item specific checklists and global rating scales.

The level of evidence resultant from these studies is low (6). The one RCT identified (10) is of moderate quality – scoring 3/5 on Jadad's RCT Grading Instrument (11). The remaining study designs were of lesser evidence grades insofar as they were comparative (four studies) or descriptive studies (three studies). The reliability and validity of the tools identified is limited. Three tools assess basic suture skills (12–14) of these studies one (12) reports a κ of 0.65 and no reliability is reported in the other two papers. Minimal validity is reported in all three papers, i.e., face and content validity and all papers fail to attempt to establish construct validity of the tools. The ultimate heterogeneity of these three tools precludes a Meta analysis study.

Two papers present tools for assessing airway management (12, 15) both of which report a κ of 0.71 and 0.93 respectively. Face and content validity are established through expert review, no construct validity is reported. Engum's (16) tool on assessing IV cannulation reports establishment of 'good inter-rater reliability (IRR) with three clinical nursing staff', however, no correlation

Table 2. Summary of tools and study design methodologies

Author	Country	Procedure	Study type	Medical students (total sample)	Tool type/items/scoring	Reliability IRR/validity	Paper strength BEME – Veloski (2006)
Liddell et al. (13)	Australia	Basic suture	Comparative	92+93 – 3rd+5th years	Global 4 themes 13 items weighted (13)	*Validity – content/face – EV* Reports IRR – No stats- no blinding	4 –Results clear
Wang et al. (14)	USA	Suturing	Comparative	8 (23) inc Dermatology Residents 4th years	Checklist 12 items weighted (20)	Single assessor – *content/face validity – EV*, IRR – N/A	3 – Probably can accept results
O' Connor et al. (12)	Canada	Suturing/ intubation	Descriptive tool development	88 (88) 5th years	Checklist 21 items weighted (21)	*Validity – content/face/IRR-* 0.65 suture/ 0.71 intubation	4 – Results clear
Kovacs et al. (15)	Canada	Airway	RCT	66 (84) inc OT's/dentists 1st years Power study conducted***	Checklist 24 items weighted (52)	*IRR* – correlation co-efficient 0.93(Bullock)*- not re-tested *validity content/face* – expert validation (EV)	4 – Results clear
McKinley et al. (18)	UK	Procedural skills-generic	Descriptive tool development	46(42)inc nurses	Global (4 themes) 42 items – 12 procedural (not yet weighted)	*Content/face validity/*-EV Predicted reliability of 0.79 Generalisability – 0.76	4 – Results clear
Engum et al. (16)	USA	IV cannula	Comparative	93 (163)inc nurses 3rd years	Checklist 21 items weighted (29)	Reports IRR (no statistics) *Validity – face/content – EV*	4 – Results clear
Morse et al. (17)	Canada	NG tube	Descriptive	4 (32) inc nurses Medical year not stated	Checklist 4 items on tube prep (yes/no)	Nil reported on reliability or validity of tool	3 – Probably can accept results

co-efficient is documented. The validity of this checklist was also not reported.

Morse's (17) study on NG tube insertion focused specifically on interpersonal skills. The tool includes four items on tube preparation and does not comprehensively assess procedural skills in nasogastric tube insertion. No data is provided on the reliability or validity of this tool or the competence of the participants. Of the total sample of 32 participants only four were medical students thus this tool has no proven feasibility in procedural skills assessment in medical students. The one global assessment tool reports a κ of 0.79 (18). It appears from the literature that the use of global scoring tools by expert practitioners is as reliable an assessment method as item specific checklists; moreover, they are considerably less labour intensive to devise (Tables 3 and 4).

Discussion

An evidence base to the pre-qualification evaluation of medical students for procedural based competencies is limited. Our review found limited literature with regard to the assessment of pre-practice medical student with predominately heterogeneous assessment tools being

Table 3. Task specific check list

Positives	Negatives
Examines all components of a skills	Time consuming to devise high volumes of content specific tools
Useful where the examiner is a novice in skills examination	Issues with reliability and validity
Inexperienced examiners can use this with little instruction	Often site specific not standardised
	Used mostly in low stakes exams for formative assessment or in the early clinical years

Table 4. Global rating scales

Positives	Negatives
IRR 0.85 (6)	Cost-need expert examiners to assess candidates
Construct validity reported with surgical skills abilities (6)	
Examiners broader abilities with regard to safety, professionalism, etc.	
Can be used with certainty in high stakes exams- certification/ re-certification	

utilised. The literature reveals two methods of procedural skills assessment in current pre registration practice namely item specific checklists (six tools) and global rating scales (two tools). Both methods of competence assessment are used in examinations from the first to the final year of the medical training programme thus none are specifically designed for an end point competence assessment. Of the two methods global assessments appear the preferred method as they have κ = 0.79 (18) than checklists and have been used with certainty in high stakes exams in the post-graduation setting (19). Moreover, these tools measure wider issues such as professionalism which is an essential component of a final pre qualification assessment tool.

The three papers presenting tools for assessing suturing skills (12–14) which have common themes, i.e., safe handling of instruments, suture insertion, knot tying and suture removal. Wang (14) used a single observer which may have inherent bias albeit consistent and the sample size of eight medical students is a limiting factor in this study. Wang's (14) tool did, however, include data on proficiency with the task, i.e., amount of suture material used, bending/blunting of needles and time taken to complete task making this tool more amenable to the future establishment of construct validity. Minimal validity is reported in all three papers, i.e., face and content validity and all papers fail to attempt to establish construct validity of the tools. The ultimate heterogeneity of these tools preclude a meta analysis study, however, the content of the three tools could be merged and used in a panel of competence markers for surgical skills ability.

Two papers present tools for assessing airway management (12, 15) both of which report good IRR (0.71 and 0.93 respectively). Face and content validity are established through expert review only. Although face and content validity are important, construct modelling allows the task to be utilised in high stakes assessment. Kovac's (15) study reports high κ reliability of 0.93 when used previously with expert observers, however, adjustments were made to the checklist since this study and the K of the modified checklist is not reported. Two points were awarded to most micro-skills in this tool but it is not clear when/whether to award one mark which may have a confounding effect on the marks awarded. O'Connor (12), however, addresses this potential study design weakness by providing a booklet of instruction for each item on his checklist. These tools could be merged and added to a panel of surrogate markers for assessing airway management.

Engum's (16) tool on assessing IV cannulation reports establishment of 'good IRR with three clinical nursing staff'. No correlation co-efficients were documented. The validity of this checklist was also not reported. Students performed real rather than simulated cannulation

procedures hence actual rather than simulated competence was assessed. This approach addresses the potential competence-performance gap. As IV cannulation is a core procedural skill required in clinical practice this tool could be re-tested to establish IRR and construct validity and added to a panel of accredited tools for intravenous cannulation skills. As intravenous cannulation is a fundamental procedural skill requiring competence from the beginning of clinical practice it is surprising that only one assessment tool is published on this procedure. Currently the academic ability of graduates of medicine is examined with rigour yet the same priority does not appear to be levelled at assuring competence pre-practice in requisite procedural skills.

Finally, Morse's (17) study on NG tube insertion focused specifically on interpersonal skills hence the tool only included four items with regard to tube preparation and thus does not a comprehensively assess skills in nasogastric tube insertion. Furthermore, no data was provided on reliability or validity of the tool or competence of participants.

In summary, reporting of reliability in these studies are rare and the validity of the tools identified is limited as all were developed locally with face and content validity reported through expert review. Construct validity was not established in any of the checklists or global assessment tools utilised in the pre-practice setting, yet Reznick (19) reports the importance of ensuring construct validity of a tool in surgical skills assessment where there is an expectation of increasing levels of competence with increasing years of training to identify and instigate remediation of students at the earliest opportunity. Procedural skills competence is a complex entity that is not easily objectively defined and is specific not generic (2).

According to medical governing bodies pre-practice competence in procedural/technical skills are a prerequisite to ensure patient safety yet there is no definitive criteria for assessing competence in the form of a reliable valid tool to fulfil this objective. Competence requires more than just technical skills abilities thus some argue (18) for assessment of broader issues of professionalism, communication, team working, values and ethics. However, validation of this assessment method is yet to be established. Assessment is a powerful driver for learning as 'learners respect what we inspect' (20) yet the procedural skills assessment tools currently in use have limitations. It appears from the literature that the use of global scoring tools by expert practitioners is as reliable an assessment method as item specific checklists; moreover, they are considerably less labour intensive to devise. Provided construct validity of a global assessment tool is assured in the undergraduate setting the tool could be used incrementally on a yearly basis from first to the final year of medical training. This would allow students to receive feedback on their level of competence be it novice, intermediate, advanced intermediate or expert. Furthermore, students could be advised as to whether they were meeting the expected level of competence for the stage of their training and in comparisons to their peers.

Finally, reports of specific procedural skills analysis provide minimal validity. No currently used tool in the pre-certification setting has established construct validity. From a theoretical perspective, reliable and valid tools are required for the assessment of procedural skills in medical education. This would ensure an internationally standardised approach of a minimal, albeit safe level of competence on entering medical practice. As medical certification boards in developed countries require proficiency in procedural skills pre-registration, much work is still needed to develop and validate new tools and ultimately devise standards of practice. This would ensure that the mobile work force has transferable skill sets irrespective of the country of training. They would thus be truly fit for practice. Furthermore, in a time of ever increasing transparency, medical training institutions may soon be required to demonstrate the rigour of their assessment processes to certifying bodies.

Conclusion

No criteria are yet offered for judging the quality of procedural skills competency in medical students at the most basic level of our entry graduates. It appears that single method, single testing of procedural skills is limited in ensuring clinical competence. Assessment needs to include many methods by many people at regular intervals across the primary training curriculum. Global assessment by Experts has been proven to be as reliable as item specific checklists. Assessment tools need to have established reliability and validity ensuring unambiguous identification of competence or incompetence in procedural skills. Fundamentally an international consensus statement by expert academic clinicians is required depicting the minimum level of competence required in procedural skills at the end point of medical training. In the interest of public safety achieving below this expert agreed level would be a barrier to certification. This would ensure that the student on completion of their primary medical degree programme is safe to perform procedural skills in the medical setting. The degree of competence could be further honed and improved during residency.

Conflict of interest and funding

Practice Points

There are eight published tools for assessing procedural skills in medical students.

Few of these tools report good IRR and all have the lowest level of validity- face and content.

There is a dearth of published work with regard to the establishment of construct validity in procedural skills assessment tools in the undergraduate setting.

There is a need for standardised assessment tools with proven reliability and validity to ensure all medicine graduates are at an internationally agreed acceptable level of competence on entering the clinical workforce.

References

1. Frank JR, ed. The CanMEDS 2005 physician competency framework. Better standards. Better physicians. Better care. Ottawa: The Royal College of Physicians and Surgeons of Canada; 2005. Available from: http://rcpsc.medical.org/canmeds/CanMEDS20705/index.php
2. Van der Vleuten C. Pitfalls in the pursuit of objectivity: issues of relaibility. Med Educ 1991; 25: 110–8.
3. Harden R. Assessment of clinical competence using an objective structured clinical examination (OSCE). Med Educ 1979; 13: 41–5.
4. Miller A, Archer J. Impact of workplace based assessment on doctors education and performance: a systematic review. BMJ 2010; 341: 5064.
5. Kogan J. Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. JAMA 2009; 302: 1316–26.
6. Harden R. BEME Guide 1: best evidence medical education. Med Teach 1991; 21: 553–62.
7. Wilhelm D, Ogan K, Roehrborn C, Cadeddu J, Peaerle M. Assessment of basic endoscopic performance using a virtual reality simulator. J Am Coll Surg 2002; 195: 675–81.
8. Andrales GL, Donnelly MB, Chu UB, Witzke DB, Hoskins JD, Mastrangelo MJ, et al. Determinants of competency judgements by experienced laporoscopic surgeons. Surg Endosc 2004; 18: 323–7.
9. Boulet JR, Gimpel JR, Dowling DJ, Finley M. Assessing the ability of medical students to perform osteopathic manipulative treatment techniques. J Am Osteopath Assoc 2004; 104: 203–11.
10. Kovacs G, Bulllock G, Ackroyd S, Cain E, Petrie D. A randomised controlled trial on the effect of educational interventions in promoting airway management skill maintenance. Ann Emerg Med 2000; 36: 301–9.
11. Jadad AR, Enkin MW. Randomised controlled trials, questions, answers and musings. London: BMJ Books, Blackwell Publishing; 2007, pp. 48–61.
12. O'Connor HM, McGraw RC. Clinical skills training:developing objective assessment instruments. Med Educ 1997; 31: 359–63.
13. Liddell MJ, Davidson SK, Taub H, Whitecross LE. Evaluation of procedural skills training in an undergraduate curriculum. Med Educ 2002; 36: 1035–41.
14. Wang T, Schwartz J, Karimipour D, Orringer J, Hamiliton T, Johnson T. An education theory-based method to teach a procedural skill. Arch Dermatol 2004; 140: 1357–61.
15. Kovacs G. A randomised controlled trial on the effect of educational interventions in promoting airway management skill maintenance. Ann Emerg Med 2000; 36: 301–9.
16. Engum SA, Jeffries P, Fisher L. Intravenous catheter training system: computer-based education versus traditional learning methods. Am J Surg 2003; 186: 67–74.
17. Morse JM, Penrod J, Kassab C, Dellasega C. Evaluating the efficiency and effectiveness of approaches to nasogastric tube insertion during trauma care. Am J Crit Care 2000; 9: 325–33.
18. McKinley RK, Strand J, Ward L, Gray T, Alun-Jones A, Miller H. Checklists for assessment and certification of clinical procedural skills omit essential competencies: a systematic review. Med Educ 2008; 42: 338–49.
19. Reznick R. An objective structured clinical examination for the licentiate of the Medical Council for Canada: from research to reality. Acad Med 1993; 68: S4–6.
20. Norcini J. Work place based assessment as an educational tool. AMEE Guide No. 31 2008; 29(9–10): 855–71.
21. Moher, D. Research methods and reporting – preferred-reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ; 339: 2535.

*Marie C. Morris
Department of Surgery
Trinity Centre for Health Sciences
Tallaght Hospital
Dublin 24, Ireland
Email: morrism4@tcd.ie

Appendix 1.
Systematic review guideline – PRISMA checklist (21)

Title: Identify the report as a systematic review, meta-analysis, or both.

Abstract

Structured summary: Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.

Introduction

Rationale: Describe the rationale for the review in the context of what is already known.

Objectives: Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).

Methods

Protocol: Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide Registration information including registration number.

Eligibility criteria: Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.

Information sources: Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.

Search: Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.

Study selection: State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, Included in the meta-analysis)

Data collection process: Describe method of data extraction from reports (e.g., piloted forms, Independently, in duplicate) and any processes for obtaining and confirming data from investigators.

Data items: List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.

Risk of bias in individual studies: Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.

Summary measures State the principal summary measures (e.g., risk ratio, difference in means).

Synthesis of results Describe the methods of handling data and combining results of studies done, including measures of consistency

Appendix 2.
BEME guide (6) – coding sheet

Citation Information
Author
Year
Reference/ID
Undergraduate entry/post graduate entry to medicine

INCLUSION CRITERIA	Yes	No
Is the article/book/report in the English language?		
Is it recent enough (<10 years) to be relevant to medical training today?		
Does it assess procedural based competence?		
Does the article/report include a tool for assessing procedural skills, i.e., excluding history/physical examination assessment?		
Does it focus on medical students?		
Does it focus specifically on assessment at the end of medical training/final procedural skills assessment pre registration?		
Does it focus specifically on summative assessment of procedural skills by Faculty or Senior Clinical Staff?		

Review Criteria

Research methods
- ○ Research Design
- ○ Data collected in the study

- Aims/Intended outcomes of assessment method
- The assessment method
  - ○ Description
  - ○ Description of control condition
  - ○ Location of the study
  - ○ Stage of the curriculum when assessment occurred
- The learners
  - ○ Number of intervention subjects
  - ○ Number of control subjects

Outcomes: Each outcome of the study, an evaluation of its methodological rigour.

Grading of strength of findings of the paper

Grade 1: No clear conclusions could be drawn; not significant
Grade 2: Results ambitious; but there appears to be a trend
Grade 3: Conclusions can probably be based on the results
Grade 4: Results are clear and very likely to be true
Grade 5: Results are unequivocal