Radiography 31 (2025) 102986

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Systematic Review

Evidence from other healthcare professions to support the standardisation of clinical assessment in pre-registration education for diagnostic radiography: a systematic review☆

E. Wyke [a], G. Coleman [a], D. Holdcroft [c], R. Tulloch [a], B. Whitehead [b], E. Hyde [a, d, *]

[a] University of Derby, Kedleston Road, Derby, DE22 1GB, United Kingdom
[b] Birmingham Newman University, Genners Ln, Bartley Green, Birmingham, B32 3NT, United Kingdom
[c] Keele University, Keele, Staffordshire, ST5 5BG, United Kingdom
[d] Society and College of Radiographers, 207 Providence Square, Mill Street, London, SE1 2EW, United Kingdom

A R T I C L E I N F O

Article history:
Received 6 January 2025
Received in revised form
8 April 2025
Accepted 7 May 2025
Available online xxx

Keywords:
Standardisation
Clinical assessment
Clinical competencies
Diagnostic radiography
Common assessment tool

A B S T R A C T

Introduction: This review aimed to gather, assess and synthesise the existing literature pertaining to the standardisation of clinical assessment tools in pre-registration education for competency-based healthcare professions, to inform the development of a standardised tool for clinical assessment in diagnostic radiography.

Methods: A systematic review based on searches in the British Education Index, CINAHL, MEDLINE, Scopus and Embase, and APA PsycINFO & Psych Articles was performed in June and July 2024. Handsearching and reference mining was also conducted. Included studies were assessed for quality using the QuADs tool. A summary table was used to extract data and results were narratively synthesised.

Results: A total of 19,327 articles were found through electronic database searches. 34 studies met the predetermined inclusion criteria and were included in the final review. Eighteen different standardised tools or frameworks were presented in the included papers. Studies reported on the process of standardising clinical and professional competencies, methods of standardising clinical assessment tools and the evaluation of standardised tools. There were also several papers which did not refer to specific tools or frameworks, but were reviews or opinion-based papers relevant to this review.

Conclusion: This review indicates that standardising clinical assessment for pre-registration diagnostic radiography students would align with what is happening in other healthcare professions globally. Although the quality of the included studies varied widely, there is a considerable amount to learn from the processes, methods and evaluations of previous standardisation attempts.

Implications for practice: This review brings valuable insight to the development of a national standardised clinical assessment tool for diagnostic radiography training. A standardised approach would support consistency and parity within and between placement settings in the UK, reducing the multiple tools some assessors are currently required to navigate. To support the design and creation of this clinical assessment tool, there is a need for a thorough consensus process to be carried out with key stakeholders, and rigorous ongoing evaluation and training.

Introduction

Background

In many countries globally, healthcare professional trainees must meet certain training standards to qualify in their profession.[1–3] However, there are differing philosophies on whether

☆ For the purpose of Open Access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

* Corresponding author. University of Derby, Kedleston Road, Derby, DE22 1GB, United Kingdom.

E-mail addresses: e.wyke@derby.ac.uk (E. Wyke), G.Coleman@derby.ac.uk (G. Coleman), donnah1963.dh@gmail.com (D. Holdcroft), r.tulloch2@derby.ac.uk (R. Tulloch), whit511@newman.ac.uk (B. Whitehead), EmmaH@sor.org (E. Hyde).

https://doi.org/10.1016/j.radi.2025.102986

healthcare education should be standardised to ensure a consistent and fair approach, or alternatively whether to value flexibility and diversity in training and assessment, embracing the individuality of local training contexts.[4] Benefits of standardisation include the potential to increase collaboration between institutions, assessors not having to learn how to use clinical assessment tools from multiple education institutions, increased reliability of assessment, and potential to increase clinical placement capacity.[5–8]

Clinical education is essential in training radiographers who can practice safely and competently, but there is heterogeneity of pre-registration radiography programmes within and between countries, presenting a challenge in terms of a shared recognition of qualifications and the movement of the radiographer workforce both nationally and internationally.[9,10]

In the United Kingdom (UK), Higher Education Institutions (HEIs) offering pre-registration diagnostic radiography programmes must ensure learners meet the Health Care and Professional Council (HCPC) Standards of Proficiency (SoPs) in order to become registered practitioners.[11] Completing clinical placements is a core aspect of pre-registration training, and in the UK, clinical sites often take learners from multiple HEIs, each of which have their own clinical assessment tools (CATs). This raises questions about the parity and consistency of these clinical assessments and means assessors must navigate multiple assessment tools. To address these issues, the Society of Radiographers (SoR) commissioned a project with a view to developing a standardised clinical assessment tool for pre-registration diagnostic radiography education in the UK.[12] The first step in this process was to conduct a review of attempts to standardised clinical assessment in healthcare professions globally. A systematic review method was chosen to assess, synthesise and learn from previous research in this area.[13]

Review questions

The aim of this review was to gather, assess and synthesise the existing literature pertaining to the standardisation of clinical assessment tools in pre-registration education for healthcare professions which require achievement of clinical proficiency to register to practice.[1,14,15]

The review aimed to answer the following research questions.

1) What evidence exists of standardisation of clinical practice pertaining to:
   a) clinical competencies
   b) clinical assessments.
2) What can be learnt from previous attempts to standardise clinical competencies and clinical assessment including methods used, key findings and evaluation processes?

Methods

The review was conducted by a research team of three diagnostic radiography academics from two HEIs (researchers 1, 2 and 3), and two research assistants with a health psychology background (researchers 4 and 5). The team was supported by a steering group consisting of key stakeholders (academics, current learners, clinical partners, researchers). Systematic review methodology was considered the most appropriate method to thoroughly identify, quality assess and summarise the findings of all studies relevant to the research aims.[16] The reporting of this systematic review was guided by the standards of the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) Statement.[17] The review protocol was not registered due to project time constraints.

Search strategy

A search strategy was developed by the research team in consultation with information specialists from the two HEIs to ensure specificity, sensibility, and replicability. The search strategy was adapted to suit each database and included MeSH terms and Boolean operators specific to each database (example in Appendix). In addition, abbreviations, truncated syntax, and asterisks were used to increase the sensitivity of the search, minimising the risk of missing relevant articles. Researchers 2 and 3 conducted hand-searching and checked reference lists of identified studies. The Society of Radiographers magazine, Insight (and its predecessor Imaging & Therapy Practice) were also searched by researchers 2 and 3 for relevant articles.

Inclusion and exclusion criteria

The search terms covered two domains: (i) construct of interest: standardisation of clinical assessment contents and methods, and (ii) population of interest: students of healthcare professions. Inclusion and exclusion criteria are detailed in Table 1.

Source of data

Electronic databases searched included: British Education Index, CINAHL, MEDLINE, Scopus, Embase, and APA PsycINFO & Psych Articles. Articles were considered between the years 2014 and 2024 to include contemporary literature on the topic of interest.[18] Reference management software (EndNote™) and Microsoft Excel™ were used to manage and document the screening process. Most duplicates were removed automatically by EndNote™, with remaining duplicates removed by hand.

Selection strategy and data extraction

Researchers 2, 3, 4 and 5 screened the titles and abstracts of the returned studies (approximately 5 % together, the remainder independently). Researchers 1, 4 and 5 then carried out a second screening of the full texts to determine inclusion or exclusion of studies. Discrepancies were discussed to reach consensus. Researchers 2, 3, 4 and 5 used a data extraction table in MS Excel ™ to extract data from relevant studies (approximately 5 % together, the remainder independently). The extracted data was developed by all researchers into themes relevant to the aims of this review.

Quality appraisal

Study quality was assessed using the Quality Assessment with Diverse studies tool (QuADs).[19] This tool was chosen because it has been demonstrated that it is easy to use and has strong inter-rater reliability and face and content validity, when applied to multi or mixed-methods research.[19] The tool assesses studies on the quality of their theoretical underpinning, how well they justify and report their aims, data collection, analytic methods, stakeholder engagement, and the strengths and limitations of their study. Each of the 13 criterion is given a score from 0 to 3, where 0 is the minimum score and 3 is the highest score. Scores are then added together to give an overall quality score.

Researcher 4 applied the QuADs criteria to all the studies, whilst researchers 1 and 3 split the studies for a second independent review. The researchers then met to discuss the independent ratings and reach consensus on scores.

Table 1
Inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria
• Studies pertaining to healthcare professions which require achievement of clinical competencies to register to practice. • Studies that involve pre-registration students or learners, including apprentices, BSc and MSc programmes. • Studies that focus on the standardisation of clinical competencies and clinical assessments • Studies which refer to clinical placement • Quantitative & qualitative studies, primary research or opinion pieces • Articles available in English	• Professions that to do not require clinical placement • Studies that do not focus on pre-registration students or learners • Studies that do not explore standardisation or that do not refer to competencies and assessments across clinical placement • Studies which focus on the entire curriculum • Review papers, commentary, meta-analyses or secondary research. • Non-human studies • Articles not available in English

Results

Summary of results

A total of 19,327 articles were returned through electronic database searches in June and July 2024, and 9723 duplicates were removed (See Fig. 1).

9479 articles were excluded by title and abstract screening. Full text screening of 126 articles left 39 articles deemed eligible for inclusion. Two more relevant articles were found through checking reference lists of eligible papers. A further 7 papers were removed during the quality assessment process as it became clear they did not meet the inclusion criteria. Thirty-four papers were included in the final review. Table 2 provides a summary of the study details and the QuADS results. The results from this table will be discussed in the following section.

Characteristics of included studies

Two studies covered multiple professions,[20,53] whilst the other studies each focused on one profession: nursing,[21–24,37–40,45–47,52] midwifery,[29,43,50,51] medicine,[25,31–36] physiotherapy,[30,41,42] occupational therapy,[26,27] anaesthesia & intensive care,[44] paramedicine,[49] pharmacy[48] and podiatry.[28] Six studies were international,[20,28,34,44,49,53] whilst other studies were conducted in Australia,[26,27,43,45–47,50–52] USA,[22,24,25,31–33,36,48] Canada,[30,41,42] Sweden,[38–40] Belgium,[29] Japan,[35] Saudi Arabia,[21] Thailand,[37] or the UK.[23]

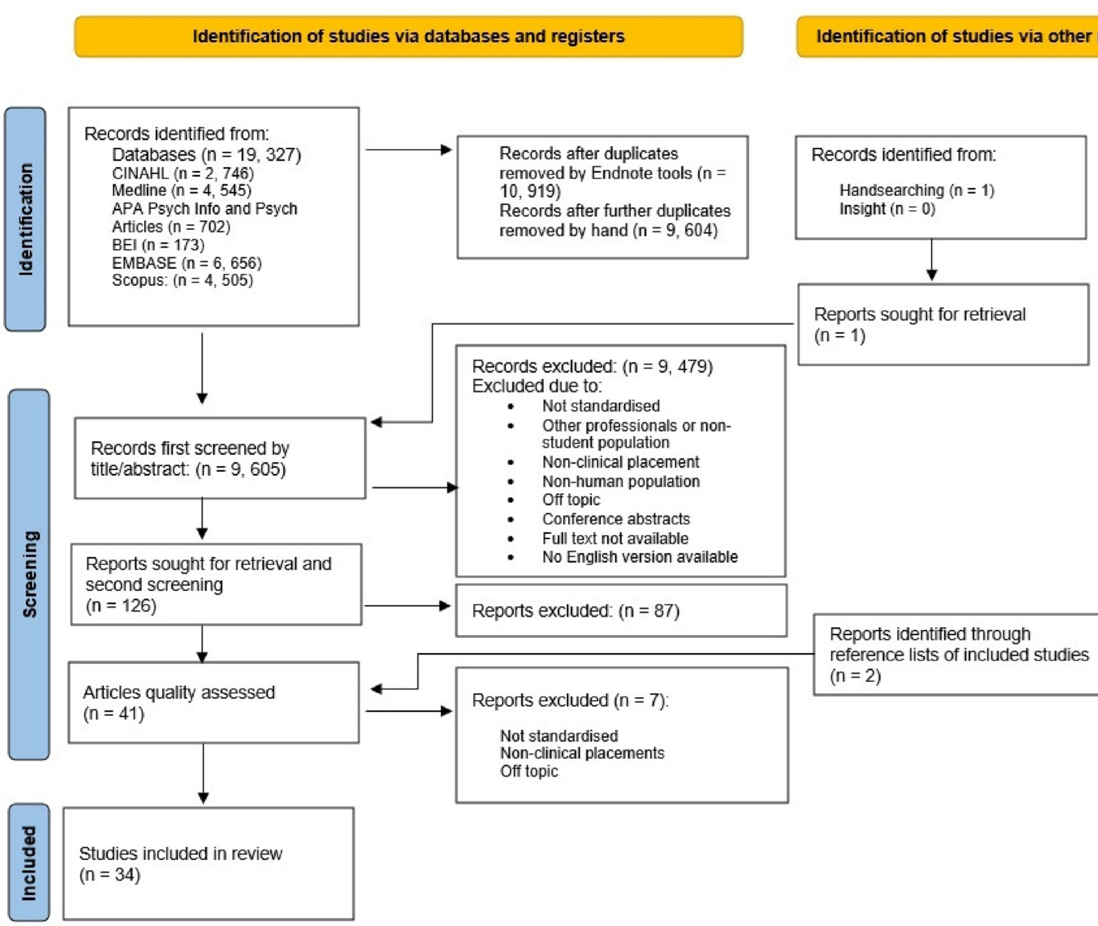

Figure 1. PRISMA diagram.

Table 2
Study summary details.

Author/Year	Study type	Country	Profession	Topic of Interest	Participant Recruitment	Sample	Key findings/Outcomes	QuADs Score
*Albarqouni*et al. *2018*[20]	Consensus statement based on a systematic review and delphi survey	Inter-national: 28 countries.	Multiple healthcare professions	Core competencies	Purposive and snowball sampling: Clinicians who had significant experience in teaching and/or practicing EBP	234 participants over a range of professions	From an earlier systematic review of 83 studies, 86 unique competencies were identified. In a delphi survey consensus was reached on 68 EBP core competencies. The final set of EBP core competencies were grouped into the main EBP domains. For each key competency, a description of the level of detail or delivery was identified. May guide to core competencies on an interprofessional level.	28
Attallah & Hasan. 2022[21]	Documentary analysis, framework drafting and delphi study	Saudi Arabia	Nursing (palliative care)	Competence	Experts from the fields of nursing practice and education were contacted as potential participants	10 experts with clinical and teaching experience	Following documentary analysis, drafting of a framework, and a delphi process, a consensus was reached on 6 domains to be included. These will provide a framework to guide implementing a competence-based assessment tool.	19
*Ayvazian*et al. *2021*[22]	Establishing speciality specific competence - repeated measures (within-groups design)	USA	Nursing	Competence	Not stated	85 NP residents	Reliability analysis of the NPRCA instrument demonstrates a high overall internal consistency when scored by the resident and faculty for both mental health and primary care specialties. Tool demonstrated consistency across departments so can be used as a basis for standardisation.	21
Baillie & Fish 2021[23]	Evaluation - focus groups	UK	Nursing	Standardised assessment tool that can be used on any pre-reg nursing programme	9 HEI's in London using PAN London tool. Information sheet distributed via email to mentors, students and academic staff. Targeted recruitment used.	Focus groups: Students ($n = 46$ in 6 groups), mentors ($n = 46$ in 7 groups) and academic staff ($n = 42$ in 6 groups).	Increased placement capacity and placement sharing for pre-reg nursing in London. increased collaboration between HEIs and placements, and HEI to HEI. Students liked the standardised CAT because it was clear they were all being assessed on the same thing - transparent and fair. Mentors liked it as it was easier for them than trying to understand different tools used by different HEIs.	22
Beauvais & Phillips 2020[24]	Evaluation - delphi study	USA	Nursing	Evaluate the validity of the ClinSimCAT	3-Round delphi study by invitation across USA. Certified nurse educators cascaded to contacts.	Nurse educators round 1: 42 participants Round 2: 13 responded Round 3: 12 responded	Agreement on a set of 20 competencies across eight domains (patient-centred care, teamwork and collaboration, evidence-based practice, quality improvement, safety, informatics, professionalism,	21

							and systems-based practice). A delphi study is a useful way for consensus on tools for assessing competency.	
*Bodamer*et al. *2015*[25]	Evaluation - prospective study design	USA	Medicine	Standardised simulation assessment	Limited detail on recruitment	207 thirds year medical students using simulators for assessment	Results provide evidence of the validity of the standardized simulation based examination (SSBE) as an extra assessment tool that uses a new approach for evaluating competency. 93 % agreed it was a fair exam.	28
*Caine*et al. *2020*[26]	Action research	Australia	Occupational therapy	Assessment of capability and performance whilst on placement. Assessment of existing tool with the aim to improve the tool	Occupational therapy educators and university contacts and OT students	213 started however only 113 completed	The majority (67 %) of respondents made at least one positive comment about the tool. The most frequent comments were that the tool was generally useful (27 %), comprehensive (14 %), easy to use (8 %), enabled learning (16 %), and assisted feedback provision (11 %). Criticisms included: Repetition (29 %), wording and understandability (27 %), length (11 %), scoring and rating scale (11 %), applicability to placement setting (11 %), and incorrect use or application of the tool by universities and/or practice educators (10 %). Some respondents requested less content; others suggested additional examples as appropriate to their clinical area.	22
*Caine*et al. *2022*[27]	Follow up from previous action research study	Australia	Occupational therapy	As above	As above	58 started and 48 completed	Participants found the tool relevant to a range of traditional and contemporary settings, easy to understand and an improvement over the previous version of the tool. Most participants found it more concise and less repetitive than the original. Participants particularly valued additions regarding culturally responsive practice, student health and well-being and reflective practice.	21
*Causby*et al. *2023*[28]	Modified delphi study - 3 rounds	Australia and New Zealand	Podiatry	Uses a delphi study with the aim of developing a consensus for a standardised assessment tool	Educators, clinicians and students and other stake holders across New Zealand and Australia	25 stakeholders participated and they reached a consensus on 55 questions after 3 reiterations	Consensus was reached on 55 statements relevant to the assessment of work integrated learning. This is the first step towards a standardised assessment tool.	28
*Embo*et al. *2017*[29]	Delphi study	Belgium	Midwifery	Article reporting the process of standardising an assessment tool,	90 educational and workplace experts was invited to participate in the	32 experts actively responded in the first delphi round (response rate of	Eleven competencies and fifty-six assessment criteria were found appropriate to assess midwifery students'	34

(*continued on next page*)

Table 2 (*continued*)

Author/Year	Study type	Country	Profession	Topic of Interest	Participant Recruitment	Sample	Key findings/Outcomes	QuADs Score
				focusing on clinical competencies	delphi study, considering a dropout of 70 %. Educational experts were selected from all current midwifery schools in flanders.	67 %). In the second and third delphi-round, 34 experts (71 %) and 28 experts (58 %) completed the survey.	competencies in the maternity ward. Three delphi rounds were needed to validate the competencies and assessment criteria.	
*Hall*et al. *2016*[30]	Anonymous online survey	Canada	Physiotherapy	Survey to evaluate whether new assessment tool required	Anonymous survey sent to all canadian physiotherapists registered with a regulatory college.	Total of 3148 physiotherapists responded (response rate 17 %). Of those who indicated relevance of student evaluation (2393).	70 % of respondents felt a new tool was required for evaluating student performance. Discussion on maximum time spent completing tool with median time 45 min Outcome on how training should be conducted to use the tool with 40 % preferring online training module. Discussion on how urgently a tool is needed that aligns with canadian practice standards. Consensus that the tool needed to be shorter and electronic.	26
*Hiller*et al. *2017*[31]	Consensus procedure (conference)	USA	Emergency medicine	Standardised assessment	Academic assembly meeting small group discussion then large group planning meeting	Sixty participants on day 1, 70 on day 2. Participants were key stakeholders in this area (not listed).	Both norm and criterion-based assessment should be incorporated in the tool. Agreement on how assessment would take place - single shift and 6-10 adequate for assessment. Grading should differ across level of learner.	14
*Hiller*et al. *2018*[32]	Evaluation - analysis of CAT scores to ascertain reliability and validity	USA	Emergency medicine	Standardised assessment	Clerkshop directors across country – not clear how recruited	EM medical students across 30 sites. 2963 NCAT-EM forms, representing 474 students and 538 assessors	Allows rigorous assessment of reliability and validity, improving the quality of clinical assessments. Benefits stakeholders. Allows comparison across students within different institutions.	15
*Hiller*et al. *2020*[33]	Evaluation of NCAT-EM tool with collection of deidentified data into separate server for analysis	USA	Emergency medicine	Assessment of standardised tool (NCAT-EM)	Convenience sampling via email to participate in study - all participants must be using the NCAT-EM tool for all shifts of medical students.	13 clinical sites participated in the study. Total of 6402 assessments across 748 students completed by 704 assessors. Three sites not included due to insufficient data.	Tool uses categories to score the student. Across the data all scores were used but the data from clinical domains was positively skewed. Also commented on skewing across institutions - positive assessment culture may have contributed to this. Rater training and calibration required - shared mental model for levels of proficiency and common system for applying them.	32
*Höhne*et al.*2022*[34]	Systematic literature review	Inter-national systematic review	Medicine	Systematic review of assessment methods in medical education	Literature search strategy	Medical students, residents, physicians	Self-assessment can be a method to evaluate (pre/post course) how skills have improved.	22

							OSCE testing skill development is widely accepted for assessing clinical competence. Direct observation of procedural skills (includes learning, supervision, rating and feedback - formative or summative.	
*Inaba*et al. *2023*[35]	Protocol for development of an assessment tool - modified delphi method	Japan	Medicine	Standardised assessment tool	Participants selected and invited via medical network system	Stage 1: 9 expert panels in ECG interpretation. Stage 2: Variety of professionals Proposed sample size 350.	Once conducted will provide an evidence-based assessment tool.	37
*Jung*et al. *2018*[36]	Consensus study and delphi process	USA	Emergency medicine (EM)	Standardised assessment tool	Participants at CORD academic assembly were invited to register for the consensus conference as well.	64 participated day 1, 76 on day 2. Modified delphi study (EM educators, clerkship directors, medical education directors, assistant program directors, programme directors, deans.	Both competency-based assessment and norm-based assessment preferred as assessment approaches. Generation of learner feedback important aspect of assessment tool. One tool should be used across all levels of learner. Professionalism is measured dichotomously with space for narrative comments.	31
*Laokhom-pruttajarn*et al. *2021*[37]	Mixed methods - document review, interviewing experts, focus groups, surveying opinions, a multi—attribute consensus process, and evaluation form (qualitative and quantitative)	Thailand	Nursing	Development of a standardised tool for professional competency assessment. Implementation and evaluation of tool	Not clear how participants were selected - survey conducted at training facilities for instructors and student nurses.	40 nursing students between year 1-year 4. 20 educators involved in training for evaluation of the tool. Consensus achieved with 10 experts.	When a tool is developed it should be tested in both and experimental phase and real-world model. Feasability, appropriateness, accuracy and usefulness all rated at a high standard. Importance of educating those key stakeholders who will be using the model for evaluation.	7
Löfmark & Mårtensson 2016[38]	Delphi study and auditing process	Sweden	Nursing	Validation of the assessment of clinical education (AssCE) tool	Workshop — clinical nurse lecturers working at 14 out of 25 universities.	Stage 1—40 clinical nurse lecturers. Consensus 70 % for delphi. Mean 9 years' experience. Stage 2—46 nurse lecturers 11 Swedish universities and 1 Norwegian. 18 nurses were present in stage 1 and delphi.	90 % response rate. Goals needed to be worded in two ways — one using behavioural terms and the other using everyday language. Clearly articulated behaviour statements that are connected to the learning outcomes are of great importance and promote consistency. “Inadequate, good achievement, very good achievement” were used as anchors. There needs to be a user-friendly validated tool and clear instructions how to use the tool with knowledgeable users to ensure a valid and reliable tool is created.	28
Löfmark & Thorell-Ekstrand 2014[39]	Evaluation: Cross-sectional evaluation study, descriptive design	Sweden	Nursing	Evaluation of the AssCE tool	Convenience sample. 4 groups of students. 1 first	192 students (38 1st year, 91 2nd year, 62 3rd year). 101 preceptors.	Nursing students and preceptors accepted revised version of AssCE. Participants wanted learning outcomes	17

(*continued on next page*)

Table 2 (*continued*)

Author/Year	Study type	Country	Profession	Topic of Interest	Participant Recruitment	Sample	Key findings/Outcomes	QuADs Score
	Based on questionnaire data				year, 2 s year and 1 third year group.	31 % qualified in the last 3 years. 32 % had more than 16 years' experience as a nurse. 64 % had no preceptor preparation education.	included in the syllabus. Combination of the AssCE and the LOs may increase preceptors' satisfaction. Education for preceptor role seem as essential for assessment situations. LOs that show demands of progression of the 3 years of study were useful for assessment process. Grading criteria supports objectivity and provides clear expectations. Students should contribute their own self-assessment before assessment discussions with preceptor. General content can be seen to be challenging.	
*Mårtensson*et al. *2020*[40]	Consensus group study	Sweden	Nursing	Evaluation and update of the AssCE-master tool for postgraduate nursing	Lecturers engaged in nursing education at 25 Swedish universities were invited to participate in workshops	Representation from 23 universities. Some attended all five workshops but not all. Numbers not clear.	Revision of the explanatory text for the tool was undertaken to provide more explanation of difference between good and very good grading. Tool involves student self-assessment to encourage reflection. National tool for consistency. Clarity of structure of assessment. limitation included if preceptor did not know how to use the tool for assessment	24
*Mori*et al. *2015*[41]	Phase 1: 3- Round delphi study; phase 2: Interviews	Canada	Physiotherapy	Standardised assessment tool that can be used in physiotherapy clinical education (measuring clinical performance behaviours via competencies)	Purposive sampling.	Delphi with 21 participants from members of the canadian national association for clinical education in physiotherapy (16) and canadian academic council interviews: 7 experts, 13 clinical instructors & 6 recent graduates.	Study produced a first draft of a new tool to assess PT students in Canada which will undergo further development and testing, including national consultation with stakeholders. Data from phase 2 will contribute to developing an online education module for CIs and students. Consensus (>80 %). Round 3 = 81 %. Graphic rating scale preferred but further refining needed on appearance. Feedback is both tick box and text. Importance of reaching consensus for terminology.	32
*Mori*et al. *2019*[42]	Modified delphi study 3 rounds	Canada	Physiotherapy	Placement evaluation form	Purposive sampling/snowball method to recruit physiotherapists, clinical education leads, centre coordinators, clinical instructors and students.	15 HEIs and 3 regional coordinators. 72 participants in total - clinical education 44.4 % of sample, clinical instructors 27.8 %, centre coordinators 22.2 % and students 19.4 %	89 % consensus at R3. Form used across Canada to enable students to share placement experience and feedback.	35

Morrowet al. 2016[43]	Evaluation of common clinical assessment tool: cross-sectional, web-based survey (quantitative and qualitative)	Australia	Midwifery	Development of a standardised tool (competency based) and subsequent adaption of tool with clinical changes/ developments	Enrolled students across 4 entry pathways across 8 HEIs and clinical educators and midwives.	27 % response rate (191 students out of 701) from 8 HEIs. 86 midwives.	94 % of student said CAT was suited to needs. CAT enabled objective assessment and fair and constructive feedback. Time constraints of assessors to complete CAT noted.	22
O'Connor & Doyle 2022[44]	Scoping review	Inter-national scoping review (Irish authors)	Anaesthesia and intensive care	Methods and practices in clinical placement (and theories relating to assessment methods)	N/A	N/A	Published studies use numerous assessment instruments, such as written, multiple choice and simulation. 7 studies used more than one CAT. Scoping review was inconclusive.	30
Ossenberget al. 2016[45]	Pilot study of an evaluation - 'think aloud' interviews and a post-test survey	Australia	Nursing	Clinical assessment tool to evaluate competence. (Ansat)	Convenience sampling for pilot	23 nursing staff who supervise and assess students from two HEIs.	Poor assessment of competency affects quality of care so valid and reliable assessment instruments need to be used. Expert groups, focus groups, statistics and interviews are needed to help review any CATs. Holistic approach to "real life professional competence".	34
Ossenberget al. 2020[46]	Modified delphi and a field test to determine psychometric properties of the revised ANSAT	Australia	Nursing	Revalidation of clinical assessment tool to evaluate competence (ANSAT)	Experts invited to take part. Registered nurses who assess from academic and clinical settings via online delphi study.	12 experts asked to participate who represented 5 of the 8 states/ territories in Australia. Potentially smaller groups in subsequent rounds but not clear. 7696 ANSAT forms completed across multiple years of study, to be included in factor analysis.	Tool is a valid measure of performance in nursing against the standards of proficiency. Student performance increases with longer placement duration. Training and support for assessors is key. Inclusion of feedback (with written examples) supports grade/pass/ fail.	25
Ossenberget al. 2023[47]	Evaluation: Quasi-experimental, non-equivalent control group design	Australia	Nursing	Practice feedback intervention for clinical placement	Final year nursing students across 3 timepoints	Two healthcare facilities in SE Australia for one HEI. 139 students in control, 70 in intervention.	Students in the intervention group had higher performance scores than controls. Students who have an active role in feedback are recommended. Clear, transparent feedback from the start of placement is advised.	34
O'Sullivan & Sy. 2018[48]	Qualitative study - semi-structured interviews	USA	Pharmacy	Identify common components of inpatient general medicine advanced pharmacy practice experiences, determine which components were important to standardize, and	Purposive sampling used to identify students and preceptors at each of 12 inpatient general medicine sites in Seattle area. Candidates contacted by a staff member, who was	Two students and one or more preceptors at each of 12 inpatient general medicine sites.	Key components to be standardized across sites were a structured orientation, written site-specific student guidelines and performance expectations, regular and meaningful student—preceptor interactions, identified performance competency standards, clear delineation of student patient	31

(*continued on next page*)

Table 2 (*continued*)

Author/Year	Study type	Country	Profession	Topic of Interest	Participant Recruitment	Sample	Key findings/Outcomes	QuADs Score
				distinguish a benchmark of each component that would indicate standardization.	not part of the research team.		care load and responsibilities, and daily interactions with patients and other health care providers.	
*Smith*et al. *2020*[49]	Development and evaluation. Survey with quantitative and qualitative questions.	Australia and New Zealand	Paramedicine	Developing and evaluating a standardised competency assessment tool for clinical placement	Phase 1: Flyers, emails, social media. Participants who were interested emailed the research team and were then sent a link to the survey.	Academics and on-road paramedics. Ninety-five (95) expressions of interest to participate in the study were received. Of those, $n = 38$ returned the completed the survey (36 % response rate). Across Australia, 3 from New Zealand.	81.5 % agreed (A) or strongly agreed (SA) the APCAT was user friendly. 71 % a or SA there was clarity around expectations of students' performance. 68 % a or SA that instructions about grading system were clear. 69 % a or SA they appreciated the practice examples to guide student assessment. 70.27 % a or SA that the year level descriptors reflected the expectations for practice level at that year. 65.79 % a or SA that the resource manual provided adequate guidance for assessors (though qualitative data suggested it may be hard to get assessors to read the guidance). 70.27 % a or SA on the inclusion of a formative assessment 89.47 % a or SA that inclusion of competency standards as criteria was important. 4 themes from the qualitative data: assessment tools, grading (concern about subjectivity, training needed, competent/not competent too binary, formative more valuable than summative), competency standards for assessment, and suitability and utility. Highlights the importance of mapping to national competency standards.	29
*Sweet*et al. *2018*[50]	Mixed method approach documenting development (focus groups) and statistical testing of the psychometric properties and evaluation to determine usability and acceptability of the tool.	Australia	Midwifery	Develop and validate a tool (AMSAT) to assess competence of midwifery students in practice-based settings	Ethics obtained. Purposive sampling.	Phase 1: 24 people from four university programs. Four iterative focus groups. Broad representation of stakeholders. Phase 2: Data collection within five university programs.	Twenty-seven midwife assessors provided 118 AMSAT forms and an unknown number of assessors provide an additional 32 forms from three clinical venues. A total of 150 completed AMSAT forms and 83 completed evaluation forms were collected. Initial draft version created, reviewed by an academic and a recent graduate then expertly reviewed by panel through 4 iterative focus groups. Each iteration resulted in refinement of the wording.	34

							Discussion around anchors or labels being an important component of an assessment tool. Bondy mentioned but not holistic enough. Global rating scale included. Also included a structured feedback table to promote self-assessment and development of a learning plan was included. Pendleton's rules approach to giving feedback.	
*Sweet*et al. *2020*[51]	Revision and validation of previous research on the AMSAT. Mixed methods: Workshops, statistical analysis of AMSAT data, and survey (quantitative and qualitative)	Australia	Midwifery	Revision and validation of a standardised assessment tool (AMSAT)	Phase 1: Prospective users of the tool from each institution were identified by local research team members and invited to a workshop. Phase 2: Clinical facilitators & preceptors from 6 universities emailed by members of the research team and invited to participate.	Phase 1: 66 participants (4 workshops), including the research team. Phase 2: A total of 255 AMSATs and 108 survey forms were collected from six australian university midwifery programs. In the AMSAT forms, there were 6305 completed data points and 70 missing data points (empty fields).	Demonstrates that the re-developed AMSAT is a valid and acceptable tool to assess midwifery students' performance to meet the NMBA midwife standards for Practice. Analysis of the 255 completed AMSAT showed the tool to be sensitive to changes in performance across formative and summative assessments. Feedback from the assessor evaluation of the AMSAT describes the tool's usefulness and adaptability. Assessor feedback about the statements was mostly positive. The tool is easily adaptable to personalized electronic devices while the integrity of the standards statements is maintained. Most feedback from assessors concerned the areas available for comment and reflection. It is recommended that feedback boxes be modified to suit local need and consider space for student and assessor comment. Changes such as these can easily be localised to suit the specific educational institution or placement facility without impacting on the validity of the tool.	36
*Takashima*et al. *2019*[52]	Opinion piece	Australia	Nursing	Possibility and benefits of using a standardised assessment tool to benchmark students' performance.	N/A	N/A	A standardised, systematic assessment process (using behavioural cues) enables benchmarking of nursing students' performance. These comparisons may facilitate and guide higher education providers and students to meet the prerequisite nursing standards successfully and consistently before graduation	N/A (tool not suitable to assess article)

(*continued on next page*)

Table 2 (*continued*)

Author/Year	Study type	Country	Profession	Topic of Interest	Participant Recruitment	Sample	Key findings/Outcomes	QuADs Score
*Witt*et al. *2022*[53]	Systematic review and a consensus procedure.	Inter-national (asia, Australia, europe, north America)	Integrative/Multi-professional (oncology)	Defining core competencies	Systematic review: Detailed search strategy. Consensus procedure: Task force of 12 experts who identified further experts to invite.	Systematic review: 21 studies Consensus procedure: 25 experts from 7 different healthcare professions and 3 international regions completed the initial web-based survey Second online survey: 28 experts from 7 different healthcare professions 57 participants.	28 competencies identified from the literature review, summarized and categorised. Several of the competencies were then split into single competencies. A total of 40 single competencies were developed.18 other competencies known to be core competencies for medical oncology were added. 58 competencies to be considered for the consensus procedure. The majority ($n = 20$, 80 %) of the participants agreed on suggested competency categories. Other competencies discussed in detail and either rephrased, merged, or deleted. A total of 38 competencies. Second survey: Participants prioritised competencies according to the importance for their profession and gave feedback on the new or modified competencies. All competencies met threshold of importance, so the task force decided to examine the feedback from the different professions and to retain all competencies that were of high priority for at least 80 % of the participants of each profession. All other competencies were again discussed by the task force group and either rephrased ($n = 1$) or omitted ($n = 1$). The final list with 37 competencies had full agreement among task force members that all were relevant for all participating healthcare professions.	26

Eighteen different standardised tools or frameworks were presented in the included papers, with some papers referring to the same tools or frameworks (Table 3). One of these was a protocol with the actual tool yet to be developed.[35] In addition, there were several papers which did not refer to specific tools or frameworks, but were studies, reviews or opinion-based papers relevant to the topic.[22,28,34,44,48,52]

Study quality

The studies varied widely in their scores using the QuADS criteria.[19] Most studies performed strongly in reporting of their research aims, study design and data collection. Approximately two thirds were strong in reporting the theoretical underpinning and analysis of their research. Only half of the studies were good or very good at reporting stakeholder involvement and study strengths and limitations. One study was an opinion piece which meant it was unsuitable for quality appraisal using the QuADs tool, but it was included as the authors felt it added valuable learning to the review.[52]

Key findings

The studies were reviewed, and their findings synthesised to address the review aims. The following section, supported by the results in Table 2, identifies the key findings under two main themes: (1) the standardisation processes and evaluation of standardisation tools; and (2) learning from attempts to standardise clinical competencies and assessments.

Standardisation processes and evaluation of standardised tools

Standardisation methods. Methods used in the standardisation process included comparing existing assessment tools, mapping professional standards onto existing assessment documents, interviewing experts, and surveying stakeholder opinion through focus groups and surveys.[21,26,35,37,41,43,48–51] Many authors used consensus methods to standardise clinical competencies and assessment tools, including non-specific consensus procedures,[31,37,40] and Delphi study methods.[20,21,28,29,35,36,38,41,46,53] Delphi study methods were also used to evaluate and update an assessment tool[24] and to develop a student evaluation form for an assessment tool.[42] Justification for using a consensus method included that it allowed for the involvement of a wide range of stakeholders (experts),[20,21,24,28,38,40–42,46,53] offered consideration of a realistic percentage of agreement,[20,21,28,38,40,41] to aid with validation of assessment tools,[24,40,41,46] and that consensus methods have been widely used in nursing and medicine where statistical analysis may be inappropriate[40].

It was highlighted that it is important to include a range of experience within the stakeholder groups such as those who are newly qualified, or newly able to assess learners, and those who have been in the profession and assessing for many years.[20,35,45,53] It must be noted that although conducting a consensus process in this way supports inclusion across the profession, the views of the participants may not represent the full range of views of the profession as a whole.[20,41,42] Furthermore, including a wide range of stakeholders assumes competence and expertise relevant to the profession from those participants, which could introduce elements of bias.[35,41]

Evaluation of standardised tools. The use of focus groups to evaluate the clinical assessment tool was an approach taken by both Baillie and Fish[23] and Caine et al.[27] This method allowed discussion of the content and usability of the clinical assessment tool across a variety of users including learners, academics and clinical staff.[23,27] Mori et al.[41] utilised interviews with clinical instructors and recent

Table 3
Existing tools and frameworks from review papers.

Tool/framework	Profession	Country	Papers (included in systematic review)
Nurse practitioner residency competency assessment (NPRCA)	Nursing	USA	Ayvazian et al. (2021)[22]
Competency framework for nursing	Pre-registration nursing	Saudi Arabia	Attallah (2022)[21]
Evidence based practice core competencies	Multi-professional	International	Albarquoni (2018)[20]
Clin sim common assessment tool (clin sim CAT)	Pre-registration nursing	USA	Beauvis & Phillips (2020)[24]
Student practice evaluation form - revised edition (SPEF-R)	Occupational therapy	Australia	Caine (2020)[26] and Caine (2022)[27]
Validated competency-based assessment instrument	Undergraduate midwifery	Belgium	Embo (2017)[29]
National clinical assessment tool for medical students in emergency medicine (NCAT-EM)	Medical: Emergency department	USA	Hiller (2017,[31] 2018,[32] 2020[33]) and Jung (2018)[36]
Assessment of clinical education (AssCE)-master tool	Post-registration nursing	Sweden	Löfmark & Mårtensson (2016)[37] Löfmark & Thorell-Ekstrand (2014)[38] and Mårtensson (2020)[39]
Pan London Practice assessment document	Nursing	UK	Baillie and Fish (2021)[23]
Assessment of clinical performance (ACP)	Undergraduate physiotherapy	Canada	Mori (2015[41] and 2019[42])
Australian nursing standards assessment tool (ANSAT)	Pre-registration nursing	Australia	Ossenberg (2016[45], 2020,[46] 2023[47]) and Takashima (2019)[52]
Australasian paramedic competency assessment tool (APCAT)	Paramedic	Australia	Smith (2020)[49]
Midwifery common assessment tool	Midwifery	Australia	Morrow et al. (2016)[43]
Australian midwifery standards assessment tool	Undergraduate midwifery	Australia	Sweet (2018[50] and 2020[51])
Assessment tool for competency of ECG interpretation (protocol)	Medicine (ECG interpretation)	Japan	Inaba et al. (2023).[35]
Professional competency evaluation model	Nursing	Thailand	Laokhompruttajarn et al. (2021)[37]
Standardized simulation- based examination (SSBE)	Medicine	USA	Bodamer et al. (2015)[25]
Education competencies for integrative oncology	Oncology (inter-professional)	International	Witt et al. (2022)[53]

graduates to explore the use and understanding of the Assessment of Clinical Performance (ACP). This provided qualitative data on its use and how the participants viewed rating scales for the competencies.

Caine et al.[26] utilised a survey tool to explore the views of university staff, clinical staff and learners on the use of the Student Practice Evaluation Form - Revised Edition (SPEF-R). The survey gathered quantitative data with free text comments at the end for further opinions. Results showed conflicting feedback on the acceptability of the tool and viewpoints which did not have a consensus.

A survey approach was also taken by Smith et al.[49] when evaluating the use of the Australasian Paramedic Competency Assessment Tool (APCAT) tool for paramedic training and competency. Closed and open-ended questions were utilised to evaluate the use of the tool, suggesting that this approach may provide a good level of detail on a tool's effectiveness and design. Conversely, Caine et al.[27] used action research to evaluate the revision to the SPEF-R tool. This included piloting of the standardised tool, survey post pilot and then focus groups and interviews prior to final revisions and publication of the tool.

Hiller et al.[32] used an alternative method to evaluate the National Clinical Assessment Tool for Medical Students in Emergency Medicine (NCAT-EM) tool through a central systemised approach. Collated data across multiple sites allowed comparison of data to allow evaluation of the use and effectiveness of this tool. This method identified areas for development around consistency, differences in curriculum and assessor expectations across institutions. It should be noted that this was a large-scale study which looked at the data of 748 students with 704 assessors over one academic year. The scale of this is much greater than would potentially be possible within a UK diagnostic radiography system.

Reviewed studies used a variety of qualitative and quantitative methods to test the validity and reliability of standardised assessment tools. Caine et al.[27] piloted their standardised assessment tool during their action research. To ensure validity of the tool, the participants also completed the original assessment tool. Findings on both tools were compared and found to be approximately the same. A small number of participants scored higher on the new tool, however this was linked to greater understanding of the use of the tool.

Whilst Mårtensson et al.[40] used a consensus process to evaluate the validity of the AssCE tool, the consideration of validity was subjective in its evaluation. Respondents were asked whether they found the tool valid. This may not be the most robust approach for ensuring validity of an assessment tool. A positive aspect of this tool was the impact that learner self-evaluation had alongside the assessment process. Learners could reflect on their abilities and skills prior to the supervisor completing the assessment form which supported more meaningful discussions.

A further approach to ensure validity of the standardised tool was described by Ossenberg et al.[46] De-identified data from participating placement sites was collated and analysed via SPSS software. With comparison of data across sites and within individual sites, this method of reviewing the validity of the tool allowed for a clear construct of the use and validity across a range of settings. Hiller et al.[33] used a centralised review of de-identified data to compare the use and validity of the NCAT-EM tool. Sweet and colleagues took a similar approach, comparing data on the use of the Australian Midwifery Standards Assessment Tool (AMSAT) tool from users.[50,51] Within both studies, the validity and reliability of the tool was evaluated with a high internal consistency identified demonstrating the reliability of the standardised assessment tool as a reliable method for assessment.[50,51]

Learning from standardising Clinical Competencies and assessments

Clinical Competencies. Authors highlighted that there were inconsistencies in which core clinical competencies were being taught and assessed at different HEIs for the same profession. Standardisation of clinical and professional competencies was proposed to ensure integrity and consistency of competence in newly qualified healthcare professionals.[20–22,24,27–33,35,40,52] Seven of the included papers outlined their process of developing a set of standardised clinical competencies.[20–22,28,29,38,52]

Ossenberg et al.[45] referred to their previous work developing a set of behavioural cues that described evidence of professional practice competence designed to provide assessors with examples of clear descriptions of behaviours that would indicate competence.[54] Ayvazian and colleagues[8] based the development of their Nurse Practitioner Residency Competency Assessment (NPRCA) on Benner's Stages of Clinical Competence framework and Entrustable Professional Activities model.[55,56] For learners on clinical placement at each level of proficiency Benner's model describes the performance characteristics developed at each stage. The definition of each level is based upon experience, how much guidance or supportive cues are needed to perform an activity, and growth in clinical knowledge, skills, judgment, and intuition.[55] The Entrustable Professional Activity model is based on how much entrustment and supervision a learner needs to perform a task at a specific level of competence. The processes and outcomes of these activities must be observable and measurable. Entrustment decisions require competency in multiple domains and learners' need to have an awareness of their own ability and limitations.[56]

Clinical assessments. Many of the papers in the review focussed on the need for standardised methods of assessing competence in learners.[20,22,23,25,29–41,50–52] Disparities in the way competence was being assessed across courses included differences in use of formative/summative assessment, grading or pass/fail of learners, use of rubrics, and continuous or one day assessments. Inconsistencies in the way assessors used existing assessment tools was also highlighted as an issue and authors expressed the importance of developing standardised, valid, reliable and easy to use tools for assessing clinical competency.

Smith et al.[49] highlighted that common competency standards had been developed by Paramedics Australasia in 2011, but that they had yet to be put into practice because of a lack of a standardised assessment approach.[57] Sweet et al.[50] referenced a literature review conducted by McAlister et al.[58] which identified four areas that should be carefully considered when designing professional competency assessments: measurement processes, quality measurement, assessment and learning, and assessment content.

Albarqouni et al.[20] stated that the levels of detail needed when completing assessment feedback must be sufficient to explain how competent a learner may be for any given assessment component. Jung et al.[36] supported this and suggested that there should be a rating scale, with a combined scale and narrative format. Mårtensson et al.[40] agreed with this approach but added that there should be clear guidance on the different anchor points when assessing and giving feedback to ensure that there are standardised responses that assessors are able to give alongside any additional comments. They also suggested that there should be behaviour cues included in the assessment tool which would help assessors improve the clarity of their feedback given to learners. Sweet et al.[50] implemented behavioural cues to their feedback to reduce ambiguity when assessing competency. For example, they felt that the term “poorly” was not an accurate descriptor of a learner's performance and removed it from their clinical assessment tool.

In their later study, Ossenberg et al.[47] found there were time constraints for assessors to provide feedback. Hall et al.[30] found that 75 % of assessors preferred to complete assessment forms online, with an average time of 45 min to complete each one. Although Hall's study is considerably older than that of Ossenberg, this highlights that time restraint has long been a factor in completing assessments and should be considered when designing a new clinical assessment tool.

Furthermore, Ossenberg et al.[47] found that learners reported a positive approach to feedback when it was co-created with their assessor via dialogue as well as in written format. They highlighted that feedback should be clear, concise and transparent to all parties, and should allow for an open exchange of inclusive ideas which starts at the beginning of the learning relationship and continues throughout the academic journey. This was also highlighted in other papers.[40,48–50]

Marking rubrics. Of the 10 studies that explored the wording for a marking rubric or standardised feedback, there were a range of anchor points used and the wording across these studies varied. These are summarised in Table 4.

Causby et al.[28] highlighted that for professions which require an assessment of competency, there should be a clear pass/fail grade to determine a learner's baseline level of competency. There was not a stipulation as to whether this grade was to be a percentage, a mark over mid-point on a Likert scale, or denoted by a marking rubric.

Hall et al.[30] found that there were discrepancies in how Physiotherapists they surveyed felt feedback should be formatted; 58 % of participants felt that only the extreme anchors of a scale should be defined, whereas 40 % felt that all anchors should be defined. These findings suggested that for each profession, there should be a consensus process to ensure that the anchor points, relevant scales and their wording are defined and utilised with the needs of that profession and their associated standards in mind.

Several papers mentioned the Bondy scale, an evaluation rating developed in 1983 by Kathleen Bondy to assess the clinical competency of nursing students.[59] Morrow et al.[43] remarked that feedback from participants on the use of the Bondy Scale was mixed, with some highlighting that the scale is difficult to interpret and that the anchor points were not clearly delineated and would be better placed for self-assessment. This finding was supported by Sweet et al.[50] who felt the Bondy Scale was a measure of assistance learners required, rather than a measure of competence. They, alongside Ossenberg et al.[45] supported the use of a global rating scale which contributes to the validity of competency assessments.

Self-assessment. Self-assessment was shown to support learners to evaluate their thoughts and experiences to identify coping strategies or correct technique to further their understanding and improve their clinical skills.[23,35,39] Undertaking self-assessment allows learners to score their own level of competency, with a view to motivating them to improve any inconsistencies between what they believe they can do, and what they are assessed as being competent in. The process of self-assessment may also contribute to enhancing learners' reflection skills, a skill learners will need as lifelong learners and a component of the Standards of Proficiency for Diagnostic Radiography.[11,39] Sweet et al.[50] highlighted the need for learners to complete self-assessments in line with Pendleton's rules[60] before feedback conversations to ensure learners strengths and weaknesses are explored through appropriate feedback discussions. When using the Assessment of Clinical Education (AssCE) tool, learners are expected to lead the discussions in formative halfway and end point meetings with clinical educators.[40]

Table 4
Feedback statements and anchor points from included studies.

Study	Scale	Additional information
Ayvazian et al. (2021)[22]	1 = critical deficiencies (an unsafe practice), resident demonstrates critical deficiencies in knowledge, skills, and attitude to safely perform the task 2 = basic, resident performs the task under the direct guidance of faculty (greater than 75 % of the time) 3 = novice, resident performs the task under extensive guidance of the faculty (between 50 and 75 % of the time) 4 = intermediate, resident performs the task under frequent guidance of the faculty (between 25 and 50 % of the time) 5 = advanced beginner, resident performs the task under occasional guidance of the faculty (less than 25 % of the time) 6 = competent, resident performs the task safely and effectively using own judgment without supporting cues.	Level 3 or higher is the expected skill level of entry into residency. Upon program completion, residents are expected to demonstrate proficiency at a minimum of level 5 in all competency areas
Hiller et al. (2020)[33] & jung et al. (2018)[36]	"Fully entrustable", "outstanding"; "mostly entrustable"; "pre-entrustable" "unable to assess"	
Laokhomp-ruttajarn et al. (2021)[37]	"Outstanding", "very good", "good", "moderate", "weak (should improve)"	
Ossenberg et al. (2016)[45]	1 = minimum standards not met, 2 = minimum standards met, 3 = performs above minimum standards in some areas, 4 = performs above minimum standards in most areas.	Each item was accompanied by a list of performance indicators in the form of behavioural cues that described evidence of professional practice competence. A global rating scale (GRS) of overall performance (not adequate, adequate, good and excellent) was also completed by the educator.
Ossenberg et al. (2023)[47]	1 (expected behaviours and practices not performed) to 5 (expected behaviours and practices performed at an excellent standard)	An item score of 3 or above qualifies as attaining the required standard.
Smith et al. (2020)[49]	"Competent" vs. "not yet competent"	
Sweet et al. (2018)[40]	"Unsatisfactory", "limited", "satisfactory", "proficient" and "excellent"	

Assessor training. Ossenberg et al.[46] stated that assessor training is a vital component to any clinical assessment tool to ensure that terminology is clearly articulated and include explanations of the necessary elements. Ossenberg et al.[46] and Smith et al.[49] highlighted the time constraints when asking assessors to read a training manual and suggested a concise manual to providing relevant information such as authentic practice settings, explanations of rating criteria and Frequently Asked Questions. This should be coupled with educational sessions to ensure that assessors are comfortable and familiar with the clinical assessment tool so that there is a reduction in potential subjectivity, an increase in consistency and a proficient understanding of how practice standards are implemented.[40,41,49] Hiller et al.[33] suggest that significant differences in assessor scores highlight differences in assessment culture across academic and professional roles; further supporting the need for standardised assessment criteria and training.

Discussion

This paper has systematically reviewed the literature published between 2014 and July 2024 on standardisation of clinical assessment across healthcare professions. Studies reported on the process of standardising clinical and professional competencies, methods of standardising clinical assessment tools and the evaluation of standardised tools. This systematic review has highlighted that other healthcare professions similar to diagnostic radiography have also recognised the importance of standardising assessment of clinical competency. Previous work carried out at a local level by the research team has shown inconsistencies in Diagnostic Radiography clinical assessment documents between HEIs, particularly when it comes to clinical competencies and how they are assessed. The Standards of Proficiency for Radiographers are set out by the Health and Care Professionals Council[45] and the Education and Career Framework for the radiography workforce[47] provides guidance for education in the radiography profession. A standardised way of assessing clinical competency that is mapped closely to these documents is required.

Learning can be taken from studies included in this review about the need to carefully consider how best to assess competence and when and how to use different forms of feedback (including self-assessment and co-created dialogue between learner and assessor); using formative or summative assessments; grading or pass/fail of learners; using marking rubrics and use of continuous or one day assessments. Conclusions on these areas cannot be drawn from this systematic review regarding what is best for diagnostic radiography. The reviewed literature shows that consensus methods are commonly used when standardising clinical assessment tools in professions similar to diagnostic radiography. A thorough consensus process could be used to support the standardisation of a UK diagnostic radiography clinical assessment tool. The consensus process should take place with relevant stakeholders who will be using the tool such as learners, clinical staff and HEI representatives. This would allow for varied opinions to be obtained, with option for comment and amendment to statements in each of the consensus rounds. The review also highlighted other methods used in the standardisation process including comparing existing assessment tools, interviewing experts, and surveying stakeholder opinion (e.g. focus groups). These steps are being carried out by our research team at the same time as this systematic review as part of a scoping review.

When developing the standardised diagnostic radiography clinical assessment tool, careful consideration should be given to the development of clinical assessor training to facilitate consistent interpretation of the tool so that it is used in the way it is intended. Key things to think about include explaining ambiguous terms that could be interpreted in different ways, keeping things concise (bearing in mind assessor's time constraints), and producing a frequently asked questions section.

Both the clinical assessment tool and the assessor training will need to be piloted and continually evaluated. Many of the articles included in this systematic review reported on evaluating standardised clinical assessment tools. Comparison of an old assessment tool to new could provide additional evidence on validity and reliability of a standardised tool,[27] however when piloting a new standardised radiography assessment tool, the results may not be comparable across institutions due to current non-standardisation.

Strengths and limitations

Due to the broad nature of the research questions, and multiple words that could be used for the same concepts, the database searches returned many papers that were not relevant to the review. However, piloting of the search strategy showed that restricting the search further would result in relevant papers being missed. Supplementing the electronic database searches with handsearching and reference list searching allows confidence that the relevant research from the last ten years is included in this systematic review. Other standardisation processes carried prior to 2014 have not been included in this review.[61,62] However, many of our review papers refer to and developed their studies based upon these prior works.

Conclusion

Collectively the studies in this systematic review indicate that standardising clinical assessment for pre-registration diagnostic

Recommendations

- Conduct a thorough consensus process with key stakeholders from across the UK to reach agreement on the wording of clinical and professional competencies for a standardised clinical assessment tool.
- Develop and pilot a standardised clinical assessment tool informed by evidence from this systematic review and the outcomes from the consensus process.
- Develop and pilot standardised assessor training to support the rollout of the standardised clinical assessment tool.
- Using qualitative and quantitative research methods design an evaluation which will assess the usability, acceptability, validity and reliability of the standardised clinical assessment tool.

Figure 2. Recommendations for next steps of developing a diagnostic radiography common assessment tool.

radiography learners would be in line with what is happening in other healthcare professions. Although the quality of the included studies varied widely, there is a considerable amount to learn from the processes, methods and evaluations of previous standardisation attempts. Most studies were in the initial stages of developing and evaluating a standardised clinical assessment tool (some using the Delphi method to evaluate) with positive outcomes in increasing collaboration, developing a transparent and fair tool to be used across multiple education settings. Recommendations for the next steps in the development of a common assessment tool for pre-registration education in Diagnostic Radiography in the UK are summarised in Fig. 2.

Conflict of interest statement

None.

Acknowledgements

We would like to thank the following people for their valuable contributions during the development of the search strategy for this systematic review: Amelia Drake (University of Derby) and Emma Edwards (Keele University), Melanie Pope (University of Derby), Rob Asher (Nottingham University Hospitals NHS Trust), the project steering group, Steve Parton (Keele University Library), Steven Bembridge (University of Derby Library) and Jill Boland (University of Derby Library).

This work was commissioned by the Society and College of Radiographers, using funding from NHS England Workforce, Training & Education Directorate.

Appendix. Example Search Strategy (Medline)

	Standardisation and Competencies	
S1	(MM "clinical Competence+")	50,694
S2	TI (standard* OR common OR shared OR unified OR national*) OR AB (standard* OR common OR shared OR unified OR national*)	4,203,269
S3	S1 AND S2	10,935
S4	TI (competenc* OR clinical N3 hours OR "clinical placement*" OR "clinical assessment*" OR practice N3 hours OR "practice placement*" OR "practice assessment*" OR simulat* OR portfolio* OR eportfolio* OR "e portfolio*" OR "electronic portfolio*" OR "assessment tool*") OR AB (competenc* OR clinical N3 hours OR "clinical placement*" OR "clinical assessment*" OR practice N3 hours OR "practice placement*" OR "practice assessment*" OR simulat* OR portfolio* OR eportfolio* OR "e portfolio*" OR "electronic portfolio*" OR "assessment tool*")	937,300
S5	S2 AND S4	150,991
S6	**S3 OR S5**	**157,143**
	Radiography	
S7	(MM "diagnostic Imaging+")	974,422
S8	TI (student* OR undergraduate* OR masters OR msc OR "post grad*" OR apprentic* OR "pre registrat*" OR bachelor* OR "foundation degree" OR diploma) OR AB (student* OR undergraduate* OR masters OR msc OR "post grad*" OR apprentic* OR "pre registrat*" OR bachelor* OR "foundation degree" OR diploma)	506,650
S9	S7 AND S8	6998
S10	TI (imaging OR "diagnostic imaging" OR "medical imaging" OR "medical radiation science" OR "radiation therap*" OR radiograph* OR "radiologic technologist*" OR radiotherap* OR sonograph* OR "therapeutic imaging" OR ultrasound* OR "assistant practition*") OR AB (imaging OR "diagnostic imaging" OR "medical imaging" OR "medical radiation science" OR "radiation therap*" OR radiograph* OR "radiologic technologist*" OR radiotherap* OR sonograph* OR "therapeutic imaging" OR ultrasound* OR "assistant practition*")	1,831,663
S11	S8 AND S10	15,563
S12	S9 OR S11	18,784
S13	**S6 AND S12**	**525**
	Allied health students	
S14	(MM "allied health Personnel+")	37,065
S15	S8 AND S14	2833
S16	TI (("allied health" OR "non-medical healthcare" OR "non-medical health-care") N2 (profession* OR occupation* OR personnel OR worker* OR staff OR specialist* OR practitioner* OR workforce)) OR AB (("allied health" OR "non-medical healthcare" OR "non-medical health-care") N2 (profession* OR occupation* OR personnel OR worker* OR staff OR specialist* OR practitioner* OR workforce))	4651
S17	**S8 AND S16**	713
S18	TI ("art therap*" OR "drama therap*" OR "music therap*" OR podiat* OR chirop* OR nutritionist* OR dietitian* OR dietician* OR "dietetic services" OR "occupational therap*" OR "operating department practitioner" OR orthop* OR osteopath* OR paramedic* OR paramedicine OR physiotherap* OR "physical therap*" OR "rehabilitation therap*") OR AB ("art therap*" OR "drama therap*" OR "music therap*" OR podiat* OR chirop* OR nutritionist* OR dietitian* OR dietician* OR "dietetic services" OR "occupational therap*" OR "operating department practitioner" OR orthop* OR osteopath* OR paramedic* OR paramedicine OR physiotherap* OR "physical therap*" OR "rehabilitation therap*")	260,741
S19	S8 AND S18	11,983
S20	**S15 OR S17 OR S19**	14,880
S21	**S6 AND S20**	**859**
	Nursing	
S22	(MM "students, Nursing+")	25,463
S23	S8 AND S22	20,893
S24	TI (nurs* OR midwif* OR "adult nurs*" OR "children's nurs*" OR "learning disability nurs*" OR "mental health nurs*") OR AB (nurs* OR midwif* OR "adult nurs*" OR "children's nurs*" OR "learning disability nurs*" OR "mental health nurs*")	551,658

(*continued on next page*)

(*continued*)

	Standardisation and Competencies	
S25	**S8 AND S24**	53,102
S26	**S23 OR S25**	54,987
S27	**S6 AND S26**	**2955**
	Medicine	
S28	(MM "education, Medical+")	144,474
S29	S8 AND S28	35,735
S30	(TI ("medical education" OR "clinical education" OR medici* OR "physician associate" OR dentist*) OR AB ("medical education" OR "clinical education" OR medici* OR "physician associate" OR dentist*))	864,233
S31	**S8 AND S30**	59,185
S32	**S29 OR S31**	78,667
S33	**S6 AND S32**	**4801**
S34	S13 OR S21 OR S27 OR S33	8020
Dates	2014–2024	5329
Language	English	5153
NOT	**"Standardized patient*" OR "standardised patient*"**	**4540**

References

1. The Heath and Care Professions Council. *Regulating health and care professionals*. 2025. https://www.hcpc-uk.org/. [Accessed 21 March 2025].
2. Medical Radiation Practice Board (MRPD) Ahpra. *Professional capabilities of medical radiation practice*. 2020. https://www.medicalradiationpracticeboard.gov.au/Registration-Standards.aspx. [Accessed 21 March 2025].
3. Saks M. The regulation of healthcare professions and support workers in international context. *Hum Resour Health* 2021;**19**(1):74. https://doi.org/10.1186/s12960-021-00618-8.
4. Bates J, Schrewe B, Ellaway RH, Teunissen PW, Watling C. Embracing standardisation and contextualisation in medical education. *Med Educ* 2019;**53**(1):15–24. https://doi.org/10.1111/medu.13740.
5. Gulikers J, Biemans H, Mulder M. Developer, teacher, student and employer evaluations of competence-based assessment quality. *Stud Educ Eval* 2009;**35**(2–3):110–9.
6. Kilgour AJ. Assessment of competency in radiography students—a new approach. *Radiographer* 2011;**58**(3):32–7.
7. McNamara J. The challenge of assessing professional competence in work integrated learning. *Assess Eval High Educ* 2013;**38**(2):183–97.
8. Wilkinson E. Survey of clinical placements within pre-registration diagnostic radiography programmes in the UK and Ireland. *Radiography* 2023;**29**(1):247–54.
9. England A, Geers-van Gemeren S, Henner A, Kukkes T, Pronk-Larive D, Rainford L, et al. Clinical radiography education across Europe. *Radiography* 2017;**23**:S7–15.
10. McNulty JP, England A, Shanahan MC. International perspectives on radiography practice education. *Radiography* 2021;**27**(4):1044–51. https://doi.org/10.1016/j.radi.2021.04.004.
11. The Health & Care Professions Council. *Standards of proficiency – radiographers*. 2023. Available from: https://www.hcpc-uk.org/standards/standards-of-proficiency/radiographers/. [Accessed 27 August 2024].
12. Beardmore C. Innovation and growth: key updates and next steps. *Synergy* 2025. Jan 2025 Available at: Innovation and growth: key updates and next steps.
13. Mulrow CD. Rationale for systematic reviews. *Br Med J* 1994;**309**(6954):597–9. https://doi.org/10.1136/bmj.309.6954.597.
14. General Medical Council (GMC). *Good medical practice*. 2024. Available via: https://www.gmc-uk.org/professional-standards/the-professional-standards/good-medical-practice. [Accessed 21 March 2025].
15. Nursing and Midwifery Council (NMC). *Standards of proficiency for registered nurses*. 2018. Available via: https://www.nmc.org.uk/standards/standards-for-nurses/standards-of-proficiency-for-registered-nurses/. [Accessed 21 March 2025].
16. Cajal B, Jiménez R, Gervilla E, Montaño JJ. Doing a systematic review in health sciences. *Clin Salud* 2020;**31**(2):77–83. https://doi.org/10.5093/clysa2020a15.
17. Sarkis-Onofre R, Catalá-López F, Aromataris E, Lockwood C. How to properly use the PRISMA Statement. *Syst Rev* 2021;**10**:1–3. https://doi.org/10.1186/s13643-021-01671-z.
18. Meline T. Selecting studies for systemic review: inclusion and exclusion criteria. *Contemp Issue Commun Sci Disord* 2006;**33**(Spring):21–7. https://doi.org/10.1044/cicsd_33_S_21.
19. Harrison R, Jones B, Gardner P, Lawton R. Quality assessment with diverse studies (QuADS): an appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. *BMC Health Serv Res* 2021;**21**:144. https://doi.org/10.1186/s12913-021-06122-y.
20. Albarqouni L, Hoffmann T, Straus S, Olsen NR, Young T, Ilic D, et al. Core competencies in evidence-based practice for health professionals: consensus statement based on a systematic review and Delphi survey. *JAMA Netw Open* 2018;**1**(2):e180281. https://doi.org/10.1001/jamanetworkopen.2018.0281.
21. Attallah D, Hasan AA. Approach to developing a Core competency framework for student nurses in Saudi Arabia: results from Delphi technique. *Nurs Rep* 2022;**12**(1):29–38. https://doi.org/10.3390/nursrep12010004.
22. Ayvazian J, Muirhead L, Belizaire S, King SS, Campbell J, Klink K. Establishing competency-based measures for Department of Veterans Affairs post-graduate nurse practitioner residencies. *J Prof Nurs* 2021;**37**(5):962–70. https://doi.org/10.1016/j.profnurs.2021.08.001.
23. Baillie L, Fish J. An evaluation of a unified practice assessment document for student nurses: students', mentors' and academics' views and experiences. *J Pract Teach & Learn* 2020;**18**. https://doi.org/10.1921/jpts.v18i1-2.1617.
24. Beauvais AM, Phillips KE. Incorporating future of nursing competencies into a clinical and simulation assessment tool: validating the clinical simulation competency assessment tool. *Nurs Educ Perspect* 2020;**41**(5):280–4. https://doi.org/10.1097/01.NEP.0000000000000709.
25. Bodamer C, Feldman M, Kushinka J, Brock E, Dow A, Evans JA, et al. An internal medicine simulated practical examination for assessment of clinical competency in third-year medical students. *Simulat Healthc J Soc Med Simulat* 2015;**10**(6):345–51. https://doi.org/10.1097/sih.0000000000000108.
26. Caine AM, Copley J, Turpin M, Fleming J, Herd C. Development of the student practice evaluation form—revised (SPEF-R2): the first action research cycle. *Aust Occup Ther J* 2021;**68**(1):21–31. https://doi.org/10.1111/1440-1630.12702.
27. Caine AM, Herd C, Copley J, Turpin M, Fleming J. Development of the student practice evaluation form–revised (SPEF-R2): the second action research cycle. *Aust Occup Ther J* 2022;**69**(4):391–402. https://doi.org/10.1111/1440-1630.12702.
28. Causby RS, Dars S, Ho M, Walmsley S, Munteanu S, Banwell HA. Consensus-based statements for assessing clinical competency in podiatry-related work integrated learning. *J Foot Ankle Res* 2023;**16**(1):43. https://doi.org/10.1186/s13047-023-00639-7.
29. Embo M, Helsloot K, Michels N, Valcke M. A Delphi study to validate competency-based criteria to assess undergraduate midwifery students' competencies in the maternity ward. *Midwifery* 2017;**53**:1–8. https://doi.org/10.1016/j.midw.2017.07.005.
30. Hall M, Manns P, Poth C, Beaupre L. Examining the need for a new instrument to evaluate Canadian physiotherapy students during clinical education experiences. *Physiother Can* 2016;**68**(2):151–5. https://doi.org/10.3138/ptc.2014–89e.
31. Hiller K, Jung J, Lawson L, Riddell R, Franzen D. Multi-institutional implementation of the national clinical assessment tool in emergency medicine: data from the first year of use. *AEM Educ Train* 2021;**5**(2):e10496. https://doi.org/10.1002/aet2.10496.
32. Hiller KM, Franzen D, Lawson L, Manthey D, Fisher J, Haughey M, et al. Clinical assessment of medical students in the emergency department, a national consensus conference. *West J Emerg Med* 2017;**18**(1):82. https://doi.org/10.5811/westjem.2016.11.32686.
33. Hiller K, Jung J, Lawson L, Franzen D, Johnson T. 37EMF A multi-institutional online database for implementation and validation of the national clinical assessment tool for emergency medicine: a consensus-derived tool for clinical assessment of medical students in emergency medicine. *Ann Emerg Med* 2018;**72**(4):S18.
34. Höhne E, Recker F, Dietrich CF, Schäfer VS. Assessment methods in medical ultrasound education. *Front Med* 2022;**9**:871957. https://doi.org/10.3389/fmed.2022.871957.
35. Inaba S, Yamamoto K, Kaga T, Wannous M, Sakata M, Yamaguchi O, et al. Protocol for development of an assessment tool for competency of ECG interpretation: expert consensus by the RAND/UCLA appropriateness method and cross-sectional testing using multidimensional item response theory. *BMJ Open* 2023;**13**(5):e072097. https://doi.org/10.1136/bmjopen-2023-072097.
36. Jung J, Franzen D, Lawson L, Manthey D, Tews M, Dubosh N, et al. The national clinical assessment tool for medical students in the emergency department

(NCAT-EM). *West J Emerg Med* 2018;**19**(1):66. https://doi.org/10.5811/westjem.2017.10.34834.
37. Laokhompruttajarn J, Panya P, Suikraduang A. The development of a professional competency evaluation model of nursing students. *Int J Nurs Educ* 2021;**13**(2):32–9. https://doi.org/10.37506/ijone.v13i2.14622.
38. Löfmark A, Mårtensson G. Validation of the tool assessment of clinical education (AssCE): a study using Delphi method and clinical experts. *Nurse Educ Today* 2017;**50**:82–6. https://doi.org/10.1016/j.nedt.2016.12.009.
39. Löfmark A, Thorell-Ekstrand I. Nursing students' and preceptors' perceptions of using a revised assessment form in clinical nursing education. *Nurse Educ Pract* 2014;**14**(3):275–80. https://doi.org/10.1016/j.nepr.2013.08.015.
40. Mårtensson G, Lind V, Edin K, Hedberg P, Löfmark A. Development and validation of a clinical assessment tool for postgraduate nursing education: a consensus-group study. *Nurse Educ Pract* 2020;**44**:102741. https://doi.org/10.1016/j.nepr.2020.102741.
41. Mori B, Brooks D, Norman KE, Herold J, Beaton DE. Development of the Canadian Physiotherapy Assessment of Clinical Performance: a new tool to assess physiotherapy students' performance in clinical education. *Physiother Can* 2015;**67**(3):281–9. https://doi.org/10.3138/ptc.2014–29e.
42. Mori B, Quesnel M, Wojkowski S. Students' perspectives on their experience in clinical placements: using a modified Delphi methodology to engage physiotherapy stakeholders in revising the national form. *Physiother Can* 2019;**71**(4):368–81. https://doi.org/10.3138/ptc-2018-43.e.
43. Morrow J, Biggs L, Stelfox S, Phillips D, McKellar L, McLachlan H. Evaluating a standardised clinical assessment tool for pre-registration midwifery students: a cross-sectional survey of midwifery students and midwives in Australia. *Women Birth* 2016;**29**(1):85–92. https://doi.org/10.1016/j.wombi.2015.09.001.
44. O'Connor E, Doyle E. A scoping review of assessment methods following undergraduate clinical placements in anesthesia and intensive care medicine. *Front Med* 2022;**9**:871515. https://doi.org/10.3389/fmed.2022.871515.
45. Ossenberg C, Dalton M, Henderson A. Validation of the Australian nursing standards assessment tool (ANSAT): a pilot study. *Nurse Educ Today* 2016;**36**:23–30. https://doi.org/10.1016/j.nedt.2015.07.012.
46. Ossenberg C, Mitchell M, Henderson A. Adoption of new practice standards in nursing: revalidation of a tool to measure performance using the Australian registered nurse standards for practice. *Collegian* 2020;**27**(4):352–60. https://doi.org/10.1016/j.colegn.2019.10.005.
47. Ossenberg C, Mitchell M, Burmeister E, Henderson A. Measuring changes in nursing students' workplace performance following feedback encounters: a quasi-experimental study. *Nurse Educ Today* 2023;**121**:105683. https://doi.org/10.1016/j.nedt.2022.105683.
48. O'Sullivan TA, Sy E. A qualitative study identifying key components to standardize across inpatient general medicine advanced pharmacy practice experiences. *Am J Pharmaceut Educ* 2018;**82**(2):6215. https://doi.org/10.5688/ajpe6215.
49. Smith AC, Framp A, Andersen P. Assessing competence of undergraduate paramedic student practice: a preliminary evaluation of the Australasian Paramedic Competency Assessment Tool. *Aust J Paramed* 2020;**17**:1–8. https://doi.org/10.33151/ajp.17.804.
50. Sweet L, Bazargan M, McKellar L, Gray J, Henderson A. Validation of the Australian midwifery standards assessment tool (AMSAT): a tool to assess midwifery competence. *Women Birth* 2018;**31**(1):59–68. https://doi.org/10.1016/j.wombi.2017.06.017.
51. Sweet L, Fleet J, Bull A, Downer T, Fox D, Bowman R, et al. Development and validation of the Australian midwifery standards assessment tool (AMSAT) to the Australian midwife standards for practice 2018. *Women Birth* 2020;**33**(2):135–44. https://doi.org/10.1016/j.wombi.2019.08.004.
52. Takashima M, Burmeister E, Ossenberg C, Henderson A. Assessment of the clinical performance of nursing students in the workplace: exploring the role of benchmarking using the Australian Nursing Standards Assessment Tool (ANSAT). *Collegian* 2019;**26**(4):502–6. https://doi.org/10.1016/J.COLEGN.2019.01.005.
53. Witt CM, Balneaves LG, Carlson LE, Cohen M, Deng G, Fouladbakhsh JM, et al. Education competencies for integrative oncology—results of a systematic review and an international and interprofessional consensus procedure. *J Cancer Educ* 2022:1–9. https://doi.org/10.1007/s13187-020-01829-8.
54. Ossenberg C, Henderson A, Dalton M. Determining attainment of nursing standards: the use of behavioural cues to enhance clarity and transparency in student clinical assessment. *Nurse Educ Today* 2015;**35**(1):12–5. https://doi.org/10.1016/j.nedt.2014.09.002.
55. Benner P. *From novice to expert: excellence and power in clinical nursing practice*. 2001. p. 336. https://doi.org/10.1097/00000446-198412000-00025. Commemorative edition.
56. Ten Cate O. Nuts and bolts of entrustable professional activities. *J Grad Med Educ* 2013;**5**(1):157–8. https://doi.org/10.4300/JGME-D-12-00380.1.
57. Paramedics Australasia. *Australasian competency standards for paramedics*. 2011. https://nasemso.org/wp-content/uploads/PA_Australasian-Competency-Standards-for-paramedics-2011.pdf. [Accessed 27 August 2024].
58. McAllister S, Lincoln M, Ferguson A, McAllister L. Issues in developing valid assessments of speech pathology students' performance in the workplace. *Int J Lang Commun Disord* 2010;**45**(1):1–14. https://doi.org/10.3109/13682820902745461.
59. Bondy KN. Criterion-referenced definitions for rating scales in clinical evaluation. *J Nurs Educ* 1983;**22**(9):376–82. https://doi.org/10.3928/0148-4834-19831101-04.
60. Pendleton D, Schofield T, Tate P, Havelock P. *The new consultation: developing doctor-patient communication*. Oxford: OUP; 2003.
61. Dalton M, Keating J, Davidson M, Alexander H. Development of the APP (Assessment of Physiotherapy Practice) instrument. *NSW: Australian Learning and Teaching Council* 2009:163–p.
62. Dalton M, Davidson M, Keating J. The Assessment of Physiotherapy Practice (APP) is a valid measure of professional competence of physiotherapy students: a cross-sectional study with Rasch analysis. *J Physiother* 2011;**57**(4):239–46. https://doi.org/10.1016/S1836-9553(11)70054-6.