Assessment tools in pre-licensure interprofessional education: A systematic review, quality appraisal and narrative synthesis

Hailah Almoghirah\textsuperscript{1,2} | Hamde Nazar\textsuperscript{3} | Jan Illing\textsuperscript{4}

\textsuperscript{1}Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK
\textsuperscript{2}Clinical Pharmacy Department, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia
\textsuperscript{3}School of Pharmacy, Newcastle University, Newcastle upon Tyne, UK
\textsuperscript{4}Health Professions Education Centre, Royal College of Surgeons in Ireland, Dublin, Ireland

Correspondence
Hailah Almoghirah, Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK.
Email: h.a.a.almoghirah2@newcastle.ac.uk

Funding information
This research was supported by a scholarship from King Saud University, Riyadh, Saudi Arabia.

Abstract
Objectives: Interprofessional education (IPE) aims to provide students with the opportunity to develop and demonstrate the team working behaviours and skills that will lead to positive patient outcomes. This systematic review aims to identify and critically appraise the assessment tools used after a pre-licensure IPE intervention and provide guidance on which tool to use according to the focus of the intervention.

Methods: In July 2019, the following electronic databases were searched: MEDLINE, ERIC, CINHAL, EMBASE and NEXUS website. All studies involving pre-licensure health care students exposed to an IPE intervention and undertook an assessment measuring student knowledge, skills, behaviour, or change in organisational practice or a benefit to patients were included. Studies that used tools relying on self-assessment only were excluded. Constructive alignment of the IPE intervention with the assessment was evaluated and quality assessment of the studies and critical appraisal of the validity evidence for the tools was undertaken.

Results: From 9502 returned studies, 39 studies met the inclusion criteria and were analysed. These were rated as good in terms of methodological quality. Acquisition of knowledge was the most commonly assessed outcome, mainly with pre/post knowledge tests, followed by behaviour change, which was measured by a range of validated tools. Patient benefit was defined as change in clinical effectiveness, patient safety or patient satisfaction. Constructive alignment of the IPE aim with the assessment was limited due to issues with study reporting. Tools measuring behaviour change demonstrate mixed adherence to quality standards around reliability, validity and scales and scoring.

Conclusions: Various methods have been used to identify change following IPE; however, choosing the most appropriate tool to support and align with the aim of the IPE intervention is crucial. We have critically appraised the available tools and offered an indication of their quality. This has informed the production of a decision aid to support the selection of the appropriate IPE assessment tool depending on the purpose of the intervention. More studies using these tools in rigorous study designs are required to substantiate the evidence base.
1 | INTRODUCTION

A recent report by West and Cola,¹ commissioned by the UK General Medical Council (GMC), emphasised the significance of effective, supportive team working environments that benefits both patient outcomes and staff well-being. Research has shown that staff well-being significantly improves productivity, care quality, patient safety, patient satisfaction, financial performance and the sustainability of health services.¹ Effective team working, however, is not being achieved throughout the health care system for a variety of reasons, including, but not exclusively, overburdened services, lack of organisational support and leadership, and poor preparation for collaborative multi-professional working in pre-licensure education and subsequent training.²,³ Interprofessional education (IPE) that fosters the competencies underpinning effective team working has long been recognised as a strategy towards improving interprofessional working and collaboration.⁴ It occurs when ‘two or more professions learn with, from and about each other to improve collaboration and quality of healthcare’.⁵ It is widely believed that IPE improves the health care system and reduces medical errors, thus improving patient safety.⁶ Many professional regulators in the UK, such as the GMC, General Pharmaceutical Council and the Nursing and Midwifery Council, support IPE as an integral part of health professional programmes.⁷⁻⁹

Teaching and learning environments are a complex system involving educators, teaching and learning context, environment, learning activities and the outcome, where all components interact with each other towards a state of equilibrium.¹⁰ It therefore follows that many forms of conventional assessment will be inextricably linked to the specifics of that context. Keeping in mind the importance of constructive alignment, educators are challenged to devise teaching and learning IPE interventions with accompanying assessments that will reassure higher education institutions, health care organisations and the public that students are prepared for effective team working and collaborative clinical practice.¹¹ Clinical practice itself will also be highly context-specific (ie clinical environment, staff, patients) subject to unpredictable variability. As such, the design and delivery of effective IPE experiences with appropriate, fair and standardised assessments appears a formidable undertaking.

A critical review of the literature identified that most of the available studies about assessments for IPE focused on pre/post study designs and used self-assessment tools measuring student’s perceptions of their own performance, rather than measures of interprofessional team working.¹²⁻¹⁴ Tools that measure self-reported change have been reported as the most common assessment employed. However, self-assessment measuring a change in either student reaction or student perception is considered a weak evaluative approach and a more comprehensive assessment approach has been recommended.¹³⁻¹⁵ Lapkin et al¹⁶ critiqued studies for lack of strong methodological approaches; poor sampling techniques; a focus on self-report outcomes; and lack of blinding when applying the education intervention.

There is a myriad of IPE interventions and a congested landscape of published tools and scales and this means educators need to dedicate significant time both to identify the most effective IPE interventions and the best tools to measure the anticipated change.

This systematic review aims to answer the following questions:

1. What tools have been developed to measure student knowledge, skills, behaviour, competence or patient benefit after an IPE experience?
2. What are the strengths and limitations of the tools identified?
3. How can this evidence be used to inform the future design and delivery of IPE and its assessment?

To address these questions, we conducted this review to interrogate the constructive alignment of the curricular activities within a narrative synthesis, investigate and critically appraise existing assessment tools that have been employed to measure the impact of IPE across pre-licensure health care programmes, and develop a decision aid that can be used to streamline the efforts of an educator looking to measure change in their students after an IPE intervention.

2 | METHOD

The work was conducted following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) standards for quality of reporting systematic review.¹⁷ The search strategy was refined in consultation with the research team and an expert librarian.

2.1 | Study identification

In July 2019, five electronic databases were systematically searched with no date limits applied: MEDLINE, ERIC, CINHAL, EMBASE and NEXUS website. Medical subject heading (MeSH) and keywords were used with ‘AND’ and ‘OR’ to narrow or broaden the search depending on the search strategy for each database. Details of search strategies based on each database are included in the Appendix S1. In order to capture as many relevant studies, different keywords were used to ensure inclusivity of the range of health care students. Included keywords such as healthcare student*, healthcare education, undergraduate*, health education and education. Also, Ovid MEDLINE(R) and In-Process & Other Non- Indexed Citations have been chosen to maximise the search for any published, not published or not cited yet article. The electronic database searches were followed by hand citation searches and reference list review to identify additional studies. Then, grey literature from the open grey databases (www.opengrey.eu) and grey literature report (www.greylit.org) were also searched.

2.2 | Study eligibility

The population, intervention, control and outcomes (PICO) framework¹⁸ was used to specify the inclusion and exclusion criteria.
Studies were included if they investigated pre-licensure health care students (Population), such as: Medicine; Pharmacy; Nursing; Dentistry; Physical Therapy; Occupational Therapy; Speech and Hearing Therapy; Health Education, and Nutrition. The studies included were those where these students were assessed after an interprofessional activity. The Kirkpatrick model (Box 1) was used to classify the outcomes measured by the assessment tool employed. This model has been recommended by Reeves et al.\(^{11}\) as an evaluation tool for measuring outcomes of IPE interventions (see Box 1).

Patient benefit was defined according to the definition of the UK Department of Health, which states that ‘The definition of quality in health care, enshrined in law, includes three key aspects: patient safety, clinical effectiveness and patient experience. A high quality health service exhibits all three’ P.B.\(^{19}\) Studies were included if it measured one or more of these three patient benefits.

To ensure capture of all available tools, studies reporting on the development and validation of a new assessment tool were also included even if they were not used with an IPE intervention. All studies written in English and available as a full paper were included. Authors of studies were contacted if full papers were unavailable online. Studies were excluded if not written in English, focused on postgraduate health care students or on employees. Studies which only used self-reported assessment as the form of evaluation were excluded; this included any form of self-assessment measuring a change in either student reaction (Kirkpatrick Level 1) or student perception or attitude (Kirkpatrick Level 2a).

One author [HA] undertook the initial searching and screening of abstracts. A second author [HN] also screened a sample (n = 30) of the abstracts as a quality assurance measure and discussed the application of inclusion and exclusion criteria with the first author to ensure a standard approach. Another author was consulted where there was any doubt of eligibility [JI, HN]. The full texts were read and screened by at least two authors [HA and HN/JI]. Data abstraction was undertaken independently by two authors [HA, HN], and a third was consulted where discrepancies arose [JI].

### 2.3 | Data extraction and synthesis

A data extraction sheet for included studies captured: study citation (authors, title, journal, year); country; study setting, study design and IPE design details including: professions involved; duration; number of students; aim and description of intervention and IPE assessment; outcomes measured; limitations of the study; findings reported and any suggestions for future work. The data extraction sheet is available in the Appendix S1.

#### 2.3.1 | IPE outcome classification and constructive alignment

The outcomes measured after the IPE intervention were classified using the levels described in the Kirkpatrick/Barr model.\(^{12}\) Constructive alignment was assessed through review of the reported IPE aims against the classification of the assessment outcomes.

#### 2.3.2 | Quality appraisal of the IPE assessment tools

It was recognised at the outset that assessments of knowledge and patient benefit would be highly specific to the context of the IPE interventions, so recommendations to IPE educators relating to these forms of assessment would have to pertain to the complete intervention rather than just the isolated assessment. However, those assessments measuring behaviour and competency change (as there were none that assessed change in organisational practice) were not limited to a topic or clinical area, for example cardiology and diabetes.

These tools were critiqued using the two-part Quality Appraisal of Interprofessional Learning Scales (QuaILS); a standardised checklist specifically developed to evaluate the qualitative features (Part A) and psychometric properties (Part B) of IPE assessment tools.\(^{13}\) Two authors [HA and JI] independently abstracted data for three tools. Authors met to compare, address inconsistencies in approach and then standardise further analysis. Data extraction was then performed by one author and verified for each instrument by a second author (JI). The authors/developers of each tool were contacted by email to provide the opportunity to review the critical appraisal conducted for their tool. Modifications were made to the appraisal of their tool if supporting evidence was provided.

---

**Box 1** The Kirkpatrick/Barr model detailing the level of outcomes from interprofessional education\(^{12}\)

| 1. Learner’s reaction | Learners’ views on the learning experience and its interprofessional nature |
|----------------------|---------------------------------------------------------------------------------|
| 2a. Modification of attitudes/perceptions | Changes in reciprocal attitudes or perceptions between participant groups. Change in perception or attitude towards the value and/or use of learn approaches to caring for specific client group |
| 2b. Acquisition of knowledge/skills | Including knowledge and skills linked to interprofessional collaboration |
| 3. Behavioural change | Identifies individuals’ transfer of interprofessional learning to their practice setting and changed professional practice |
| 4a. Change in organisational practice | Wider changes in the organisation and delivery of care |
| 4b. Benefits to patients/clients | Improvement in health or well-being of patients/clients |
2.3.3 Development of a decision aid

A decision aid was created to guide educators through the necessary choices that need to be made during the design of an IPE intervention and associated assessment. This was informed from the findings from the preceding steps of analysis and focused on assessment subject (individual or team or both), the domains of the assessment, for example communication and the indicative tool.

2.4 Quality assessment

Included studies were appraised by one author [HA] and then discussed with the research team using the Joanna Briggs Institute Critical Appraisal of Evidence Effectiveness Tools, which was developed by the Joanna Briggs Institution based in the University of Adelaide.20 The Joanna Briggs Institute, like the Cochrane Collaboration, is an evidence-based organisation formed to develop guidelines and methodologies on the process of conducting systematic reviews. The suite of tools offers a range of checklists to assess trustworthiness, relevance and results of studies across different study designs.

Systematic narrative synthesis was adopted to present the results more widely due to the heterogeneous nature of the data included in the studies. Narrative synthesis is particularly salient to inform onward practice in the field being investigated, by providing information on effects but also implementation.21 No statistical analysis was appropriate to apply to the heterogeneous quantitative data.

There were no ethical issues identified in conducting this review; therefore, no applications for approval were sought.

3 RESULTS

From 9502 returned studies, 39 studies met the inclusion criteria and were included for analysis as illustrated in the PRISMA flow diagram (see Figure 1).

3.1 Study characteristics

Analysis of included articles demonstrated that the nursing profession was the most frequently represented (n = 27),22-48 followed by medicine (n = 20),24,26,27,29,30,32-34,36,38,40-44,46-50 and pharmacy (n = 14).22,24,26-28,29,31,41,43,44,50-53 A large number of studies were undertaken in the United States of America (n = 24),22,25,26-33,36-41,45,49-52,54-56 with a smaller number from Australia (n = 4),34,43,57,58 Canada (n = 3),53,59,60 and the United Kingdom (n = 2).42,47 (see Appendix S1). Acquisition of knowledge was the most commonly assessed outcome (n = 17),24,26-29,32,33,35-38,39,42-45,49,53,55,57 followed by behaviour change (n = 17),22,25,30-32,36,37,40,41,44,45,50,55,56,58-60 benefit to patients (n = 11)23,34,39,42-46,48,51,52,54 and acquisition of skills (n = 3).25,30,31 No studies were found focusing on organisational change as an outcome. In some studies, two outcomes were assessed such as acquisition of skills and change in behaviour (n = 4),25,30,31,59 acquisition of knowledge and change in behaviour (n = 4)32,44,45,55 and acquisition of knowledge and benefit to patient (n = 1).39 Some studies also reported self-assessment measures. The self-report element was excluded from the analysis as this fell outside the remit of this review.

There was significant heterogeneity in IPE interventions reported in the included studies. The majority of these studies delivered IPE activities in classroom or simulated settings (n = 28)22,23,25-33,35-38,40,41,43-45,49,50,53,54,56-59 rather than in clinical/health service settings or community setting (n = 10). IPE interventions varied in their length and episodic nature, where the majority presented one-off sessions (n = 19),22,23,25-29,31,32,35,38,40,41,49,51,53,54,56,57 a short series of activities (n = 8)36,39,42-44,46,47,55 and some reported interventions that were of a significant longitudinal nature (months to years) (n = 10).24,30,33,34,36,38,48,50,52,59 There were a range of teaching and learning approaches adopted, with many facilitating experiential learning from clinics, wards or providing services to patients and the public (n = 13).23,30,34,42-46,48,50,52,54,55 many employed active learning strategies such as simulation, team-based learning and interactive workshops (n = 19)24,26,28,29,32,33,36-41,43-45,49,56,57,59 with much fewer exclusively using didactic methods such as lectures and self-directed exercises (n = 3).27,52,59

In most cases, the aim of the IPE intervention was either reported or could be interpreted from the study description. However, in some, this information was challenging to deduce and therefore difficult to evaluate the constructive alignment to the assessment strategy (n = 9),26,28,32,34,40,43,44,48,55 When reporting was appropriate, the complementarity of the assessment to the aims was generally reasonable (n = 16),23,24,27,30,33,36,39,41,42,45,49,50,52,56,57,59 for example where the aim was to achieve knowledge gain, a knowledge test was used.27 In some cases, the alignment was poor as an inappropriate assessment tool was selected to address the stipulated aim (n = 8),29,35,38,46,47,51,53,54 for example the aim was to develop competencies and the assessment tool just tested knowledge.38

3.2 Acquisition of knowledge

The most common research design to measure acquisition of knowledge was a pre/post knowledge test (n = 13).27,29,33,35,38,39,43-45,53,55,57 where most reported improved knowledge after the IPE intervention (n = 11).24,27-29,32,35,38,39,44,53,55 Some studies assessed this gain in the context of team-based learning, using the team readiness test (tRAT) and the individual readiness test (iRAT) (n = 2).26,49,1 positive outcomes were identified. Another study used the University of Missouri-Kansas Screening and Brief Intervention Knowledge Assessment (UMKC-SBIRT KA) (n = 1).24 Of the studies that adopted a pre/post testing approach, only two also had comparator groups.27,28 However, in one such study, the comparator group existed as a control where no IPE intervention was introduced.27 A good example was provided by Racic et al, who illustrated that...
students attending an interprofessional course on diabetes scored significantly higher in a multiple choice knowledge test than those attending a uniprofessional course covering the same material.

3.3 | Acquisition of skills

Two studies aimed to measure the development of skills. One study compared student skills post-IPE and compared results to a previous cohort not exposed to the IPE intervention and the other study involved repeated measures to capture skill development over time alongside a series of IPE sessions. In both cases, there was an improvement in clinical skills and communication skills (using the Indiana University Individual Communication Rubric (IUICR) and the Indiana University Team Communication Rubric).

3.4 | Patient benefit

Different tools were used to measure patient experience such as a post-appointment survey, IPE experience evaluation (n = 2), a patient experience questionnaire (n = 5), and one study.
included a patient focus group. Patients were satisfied with care provided via the IPE intervention and findings show there was improved knowledge in disease management.

Identification of drug therapy problems (n = 1) was used to assess patient safety and facilitated optimisation of medication use.

Referral due to a health issue identified (n = 1) and managing chronic disease (n = 1) were used to assess clinical effectiveness. This led to patients receiving more specialist care thereby improving clinical outcomes.

3.5 | Behaviour and competency development

The terms ‘behaviour’ and ‘competency’ were used across the different studies interchangeably. In our analysis, there was not an attempt to interrogate this use of terminology and the definition of competency provided by Carraccio et al, which describes the overlap in meaning, has been assumed, ‘Competencies are a complex set of behaviors built on the components of knowledge, skills, attitudes...’. Seventeen studies focused on a change in behaviour following the intervention and were assessed in a range of ways. Twelve tools with validity evidence that measured behaviour are presented in Table 1 and other studies measuring behaviour used a standardised patient checklist (n = 1), a team performance evaluation by faculty (n = 1), a standardised grading rubric based on OSCE evaluation (n = 1) or a 100 mm visual analogue (VAS) clinical decision-making scale (n = 1).

3.6 | QuAILS results

After contacting the authors/developers of each tool (n = 11), 45.4% responded (n = 5) and 33.3% (n = 4) of the tools’ critical appraisal were revised based on the supporting evidence provided.

3.6.1 | Qualitative features

In the 12 tools with validity evidence, the scales and subscales pertaining to a range of behaviours and competencies that map to the WHO Interprofessional Learning Domains and the IPEC core competencies for interprofessional collaborative practice. The length of tools varied from four items to 42 items. These tools either assessed the behaviour of the individual, the team or both.

For more about qualitative features, see Appendix S1.

3.6.2 | Psychometric properties

Psychometric properties for each tool were appraised using the three standards: validity, reliability and scale and score, which are presented in Table 2.

3.6.2.1 | Individual performance tools

All tools partially met standards for validity. IPA and iTOFT met a greater proportion of the standards. All tools reported evidence of reliability except IPEC, with the IPA tool again demonstrating the greatest quality in this standard.

All standards in scales and scores were met in iTOFT only, with other tools showing evidence in a few or just one of these standards.

3.6.2.2 | Team performance tools

PACT-Expert and PACT-Novice met most of the validity standards with the former performing more superiorly than the latter. All team performance tools partially met the reliability standards. Scale and score standards were met only in PACT tools.

3.6.2.3 | Individual and team performance tools

All tools met similar standards for validity. Reliability standards for these tools were partially met. Tools reported reliability but did not report the data measurement error. Scale and score standards were met for McMaster-Ottawa scale while only being partially met for IUICR & IUTCR and ORTAS.

For more about the psychometric properties, see Appendix S1.

3.7 | Decision aid

A flow chart has been developed to facilitate the decision making of an educator looking to select an appropriate IPE assessment tool for their IPE intervention (see Figure 2). It is recommended that the educator consider the following questions to aid the selection of the appropriate tool and note recommendations based on the results of this review:

- What is the aim of the IPE intervention?
- Who is being assessed?
- What are the domains of interest?
- What tools are available that will measure the achievement of that aim in that context?
- What tools have the most evidence to support their reliability, validity and scales and scoring?

3.8 | Quality assessment

The average quality score, assessed by the Joanna Briggs Institute Critical Appraisal of Evidence Effectiveness Tools, of the included studies was 70.1% (Appendix S1), which indicated good methodological quality. The quasi-experimental studies demonstrated mixed quality, with the most common deficiencies being lack of a control group, incomplete follow-up and the absence of pre-intervention measurements. In the cross-sectional studies, confounding factors were rarely identified or controlled for when measuring outcomes, a common weakness...
| Tool measurement characteristics | Level of performance | Competency domains | Textual feedback | Available online | Limitation of use |
|----------------------------------|---------------------|-------------------|-----------------|-----------------|------------------|
|                                  | Individual | Team | Individual | Team | Interprofessional | Collaboration | Leadership | |
| **Appropriate behavioural tool to be chosen** | | | | | | | | |
| Interprofessional Professionalism Assessment tool (IPA) | ✓ | x | ✓ | ✓ | x | x | ✓ | ✓ | – |
| Individual Teamwork Observation and Feedback Tool (iTOFT) | ✓ | x | x | ✓ | x | ✓ | ✓ | ✓ | – |
| Interprofessional Collaborator Assessment Rubric (ICAR) | ✓ | x | x | ✓ | ✓ | x | ✓ | ✓ | – |
| The Collaborative Behaviors Observational Assessment Tools (CBOATs) | ✓ | x | ✓ | ✓ | ✓ | x | Unknown | x | Nursing and medical professions only in end of life workshops |
| Interprofessional Competence Scale (IPEC) | ✓ | x | x | ✓ | ✓ | x | x | ✓ | – |
| The Performance Assessment for Interprofessional Communication and Teamwork (PACT-expert) | ✓ | ✓ | x | ✓ | x | ✓ | x | – | – |
| The Performance Assessment for Interprofessional Communication and Teamwork (PACT-novice) | ✓ | ✓ | x | ✓ | x | ✓ | x | – | – |
| Creighton-Interprofessional Collaborative Evaluation (C-ICE) | x | ✓ | x | ✓ | ✓ | x | ✓ | – | Rating is yes or no |
| The Team Performance Observation Tool (TPOT) | x | ✓ | x | ✓ | x | x | ✓ | – | – |
| McMaster-Ottawa scale | ✓ | ✓ | x | ✓ | ✓ | x | x | ✓ | – |
| Indiana University Individual Communication Rubric (IUICR) and Indiana University Team Communication Rubric (IUTCR) | ✓ | ✓ | x | ✓ | x | x | Unknown | x | – |
| Operating Room Teamwork Assessment Scales (ORTAS) | ✓ | ✓ | x | ✓ | x | x | ✓ | x | Limited use to the operating theatre |
### TABLE 2  Evidence of validity, reliability, and scales and scores in tools which measure behavioural change outcomes after interprofessional education

| Evidence                  | Individual Level performance tools | Team level performance tools | Both |
|---------------------------|-----------------------------------|-------------------------------|------|
|                           | IPA | ITOFT | ICAR | CBOAT | IPEC | PACT-Expert | PACT-Novice | C-ICE | TPOT | OTTAWA | IUICR & IUTC | ORTAS |
| (1) Test content          |     |       |      |       |      |             |             |       |      |         |              |       |
| Was a pool of items generated? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | NC | Yes | Yes | Yes |
| Were criteria for item pool reduction specified? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | NA | NA | Yes | Yes | Yes |
| Were criteria for item design quality specified? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Was instrument trialled in an environment aligned with the intended purpose of instrument? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Were characteristics of study participants reported? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Were characteristics of participants involved in the development of the instrument reported? | Yes | Yes | Yes | No | No | No | No | No | No | Yes | No | Yes | Yes |
| Were pre-qualification health students involved in the development of the instrument? | No | No | Yes | No | No | No | No | No | No | No | No | No | No |
| (2) Internal structure    |     |       |      |       |      |             |             |       |      |         |              |       |
| Have statistical approaches been used to investigate the internal structure of the instrument? | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Was Cronbach's alpha reported for each subscale or factor and within acceptable range .70-.90*? | Yes | Yes | No | No | No | Yes | No | NA | Yes | NA | Yes | No | No |
| Was any bias for items among different subgroups investigated? | No | No | No | No | No | No | No | No | No | No | Yes | No | No |
| (3) Response processes    |     |       |      |       |      |             |             |       |      |         |              |       |
| Have analyses of individual responses been completed? | Yes | Yes | No | Yes | No | Yes | Yes | No | No | Yes | Yes | Yes | Yes |
| (4) Relationship to other variables |     |       |      |       |      |             |             |       |      |         |              |       |
| Have any correlational studies been completed? | Yes | Yes | No | Yes | No | No | No | No | No | Yes | No | No | No |
| (5) Consequences of testing |     |       |      |       |      |             |             |       |      |         |              |       |
| Have the benefits of using the instrument been described? | No | No | No | No | No | No | No | No | No | No | No | No | No |
| Overall assessment against Standards relating to validity | / | / | / | / | / | / | / | / | / | / | / | / | / |
| Reliability               |     |       |      |       |      |             |             |       |      |         |              |       |
| Has reliability been measured and reported? | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Have measurement error data been reported? | Yes | No | No | No | No | No | No | No | No | No | No | No | No |
| Overall assessment against Standards relating to reliability: | + | / | / | / | - | / | / | / | / | / | / | / | / |

(Continues)
TABLE 2 (Continued)

| Evidence                           | Individual Level performance tools | Team level performance tools | Both |
|-----------------------------------|-----------------------------------|------------------------------|------|
|                                   | IPA                               | PACT-Expert                  |      |
|                                   | ITOFT                             | PACT-Novice                  |      |
|                                   | ICAR                              | C-ICE                        |      |
|                                   | CBOAT                             | TPOT                         |      |
|                                   | IPEC                              | OTAWA                        |      |
|                                   |                                   | IUICR                        |      |
|                                   |                                   | IUTCR                        |      |
|                                   |                                   | ORTAS                        |      |
| Scales and scores                 | Yes                               | Yes                          |      |
| Has information relating to the   | Yes                               | Yes                          |      |
| construction of score scales      | No                                | No                           |      |
| scores been provided?             | No                                | No                           |      |
| Has the score calculation protocol| No                                | Yes                          |      |
| been specified?                   | Yes                               | No                           |      |
| Is the scoring protocol           | NA                                | Yes                          |      |
| appropriate given the wording of  | NA                                | NA                           |      |
| items (ie positive and negative   | NA                                | NA                           |      |
| wording?                          |                                   | NA                           |      |
| Has the interpretation and        | Yes                               | Yes                          |      |
| meaning of scores been described? | No                                | No                           |      |
| Overall assessment against        | /                                 | +                            |      |
| Standards relating to scales and  | /                                 | /                            |      |
| scores                         | /                                 | /                            |      |

Abbreviations: C-ICE, Creighton-Interprofessional Collaborative Evaluation; CBOATS, Collaborative Behaviors Observational Assessment Tools; ICAR, Interprofessional Collaborator Assessment Rubric; IPA, Interprofessional Professionalism Assessment; IPEC, Interprofessional Competence Scale; ITOFT, Individual Teamwork Observation and Feedback Tool; IUICR, Indiana University Individual Communication Rubric; IUTCR, Indiana University Team Communication Rubric; NA, not applicable; NC, not clear; OTAWA, McMaster-Ottawa scale; ORTAS, Operating Room Teamwork Assessment Scales; PACT-Expert, Performance Assessment Communication and Teamwork Tool-Expert; PACT-Novice, Performance Assessment Communication and Teamwork Tool-Novice; TPOT, Team Performance Observation Tool; /, standard partially met; +, standard met; -, standard not met.

* Acceptable range for Cronbach’s alpha based on Streiner.63

4 | DISCUSSION

This review has identified assessment tools that measured the impact of IPE interventions across the levels of the Kirkpatrick/Berlanga model (except for change in organisational practice). Those tools measuring change in attitudes and competencies. No single tool was identified as meeting all of the quality standards as articulated by Oates et al. However, tools that have the most evidence supporting their reliability and validity in assessing performance at the individual level are IPA and ITOFT, whereas the PACT tools have the most credibility in assessing performance at the team level. The 12 tools measuring behaviour change present educators with the opportunity to measure team and/or individual behaviours and competencies. A single tool was developed to improve patient outcomes in smoking cessation post-appointment and at follow-up in the interprofessional care group. However, sadly this tool did not translate to improvements in smoking cessation outcomes.13

Similar tools which measured patient benefit were tailored to the patient outcome rather than the IPE intervention, limiting a direct comparison and therefore contributing more meaningfully to the evidence base of smoking cessation interventions. However, few studies have used self-assessment tools to measure change in attitudes and competencies. No single tool was identified as meeting all of the quality standards as articulated by Oates et al. However, tools that have the most evidence supporting their reliability and validity in assessing performance at the individual level are IPA and ITOFT, whereas the PACT tools have the most credibility in assessing performance at the team level. The 12 tools measuring behaviour change present educators with the opportunity to measure team and/or individual behaviours and competencies. A single tool was developed to improve patient outcomes in smoking cessation post-appointment and at follow-up in the interprofessional care group. However, sadly this tool did not translate to improvements in smoking cessation outcomes.13

SIMILAR TOOLS WHICH MEASURED PATIENT BENEFIT WERE TAILORED TO THE PATIENT OUTCOME RATHER THAN THE IPE INTERVENTION, LIMITING A DIRECT COMPARISON AND THEREFORE CONTRIBUTING MORE MEANINGFULLY TO THE EVIDENCE BASE OF SMOKING CESSION INTERVENTIONS. HOWEVER, FEW STUDIES HAVE USED SELF-ASSESSMENT TOOLS TO MEASURE CHANGE IN ATTITUDES AND COMPETENCIES. NO SINGLE TOOL WAS IDENTIFIED AS MEETING ALL OF THE QUALITY STANDARDS AS ARTICULATED BY OATES ET AL. HOWEVER, TOOLS THAT HAVE THE MOST EVIDENCE SUPPORTING THEIR RELIABILITY AND VALIDITY IN ASSESSING PERFORMANCE AT THE INDIVIDUAL LEVEL ARE IPA AND ITOFT, WHEREAS THE PACT TOOLS HAVE THE MOST CREDIBILITY IN ASSESSING PERFORMANCE AT THE TEAM LEVEL. THE 12 TOOLS MEASURING BEHAVIOUR CHANGE PRESENT EDUCATORS WITH THE OPPORTUNITY TO MEASURE TEAM AND/OR INDIVIDUAL BEHAVIOURS AND COMPETENCIES. A SINGLE TOOL WAS DEVELOPED TO IMPROVE PATIENT OUTCOMES IN SMOKING CESSATION POST-APPOINTMENT AND AT FOLLOW-UP IN THE INTERPROFESSIONAL CARE GROUP. HOWEVER, SADLY THIS TOOL DID NOT TRANSLATE TO IMPROVEMENTS IN SMOKING CESSATION OUTCOMES.13

Similarly, tools which measured patient benefit were tailored to the patient outcome rather than the IPE intervention, limiting a direct comparison and therefore contributing more meaningfully to the evidence base of smoking cessation interventions. However, few studies have used self-assessment tools to measure change in attitudes and competencies. No single tool was identified as meeting all of the quality standards as articulated by Oates et al. However, tools that have the most evidence supporting their reliability and validity in assessing performance at the individual level are IPA and ITOFT, whereas the PACT tools have the most credibility in assessing performance at the team level.
level. If both individual and team performance are of interest, the McMaster-Ottawa scale most rigorously meets the quality criteria.

Behaviour change within the pre-licensure curriculum is often the most optimal outcome of assessment, as interprofessional behaviours will increase the potential of students to contribute to effective team working in practice and lead to better patient outcomes. Also, given the mixed exposure to, and experience of real-life clinical practice within pre-licensure professional programmes, measuring patient outcomes as a consequence of educational interventions is challenging due to the episodic nature of many interventions as well as the resources required to measure the patient benefit. Despite the endeavours of these tools to assess performance, the challenge of context specificity remains, that is the observation that an individual's performance on a particular problem or in a particular situation is only weakly predictive of the same individual's performance on a different problem or in a different situation64. Educators must consider that their own subjective use of these tools within the context of the specific IPE intervention offers only one ‘truth’ of a student’s or team's performance. It has been suggested that to arrive at a more ‘true’ summative decision about a performance (whether team or individual), a collection of assessments over some time period (including versions of the ‘truth’ from different assessors in different contexts) are reviewed by a committee or similar to make a collective, coherent (rather than ‘objective’) determination65.

The decision aid developed in this review illustrates which domains are measured by each tool and is offered here to simplify the constructive alignment of aim and outcome. Used iteratively over time and across IPE interventions, this could lead to a set of assessments that yield evidence about the development and demonstration of interprofessional knowledge, skills and behaviours of individuals and teams.

This review has synthesised the evidence to help educators navigate the expanse of literature on IPE assessment and a range of tools have been critiqued for use with a range of IPE interventions. However, no single tool has emerged to suit all IPE interventions. There are tools that have more rigorous evidence of quality and improve educational assessments. However, they need to be used where there is coherence in aim, subjects of assessment, intervention (including domain of interest) and outcome. Thistlewaite58 describes this constructive alignment as crucial to ensure that added value from IPE over uniprofessional learning is effectively investigated.

4.1 | Recommendation/future work

Future work should focus on consistent and coherent reporting and alignment of IPE interventions and their assessments. Guidelines produced by Phillips et al.6463 which provide educators
and researchers with a comprehensive checklist, can facilitate the reporting of educational interventions for evidence-based practice. This would mean the individually reported IPE interventions are transparent, reproducible and make valuable contributions to the literature and inform practice. Where possible, experimental designs, with a defined control group (with no IPE intervention), will advance knowledge in the field about the impact of IPE experiences.

This review has limitations. Screening of the studies at the title and abstract stage was only conducted by two researchers independently for 30 studies. Consequently, some studies may have been excluded due to potential screening fatigue. Some of the information reported by the published studies was limited, which limited data extraction. Individual authors of studies could have been contacted to provide more detail, but time and resource limited this as an option. However, some of the synthesis was carried out by two researchers independently, and all interpretations were discussed within the research team. Only the lead authors of the tools included in the critical appraisal were contacted to inform this aspect of the study.

5 | CONCLUSION

There is no single IPE tool suited to all IPE interventions. There are a wide range of possible approaches to assess IPE, with more tools being developed in recent years that measure change in behaviour and competency. The selection of an appropriate assessment tool is a decision to be taken alongside the conception and design of the IPE intervention.

This review provides educators with a decision aid about which tools might be most appropriate for the purpose of their planned IPE intervention.

Future work is required on using these tools with stronger research designs to discern if any change detected in students is as a result of the education being interprofessional in nature. Such an approach will provide evidence about the value of IPE within pre-licensure health care curricula.

ACKNOWLEDGEMENTS

Hailah Almoghirah acknowledges the funding that she received as a scholarship from King Saud University, Saudi Arabia.

CONFLICT OF INTEREST

None.

AUTHOR CONTRIBUTIONS

All authors contributed to the design of the work. One author [HA] undertook the initial searching and screening of abstracts. A second author [HN] also screened a sample (n = 30) of the abstracts as a quality assurance measure and discussed the application of inclusion and exclusion criteria with the first author to ensure a standard approach. Another author was consulted where there was any doubt of eligibility [JI, HN]. The full texts were read and screened by at least two authors [HA and HN/JI]. Data abstraction was undertaken independently by two authors [HA, HN], and a third was consulted where discrepancies arose [JI]. [HA] contributed to the writing original draft of the paper. [JI] and [HN] contributed to the reviewing and editing the paper. All authors approved the final version for publication. All authors have agreed to be accountable for all aspects of the work.

ETHICAL APPROVAL

There were no ethical issues identified in conducting this review, therefore no applications for approval were sought.

ORCID

Hailah Almoghirah https://orcid.org/0000-0001-7893-5015
Hamde Nazar https://orcid.org/0000-0002-3168-5386
Jan Illing https://orcid.org/0000-0001-6218-9775

REFERENCES

1. West M, Coia D. Caring for Doctors Caring for Patients. United Kingdom: General Medical Council; 2019.
2. de Vries-Erlich J, Reuchlin K, de Maajer P, van de Ridder JM. Identifying facilitators and barriers for implementation of interprofessional education: perspectives from medical educators in the Netherlands. J Interprof Care. 2017;31(2):170-174.
3. Hammick M, Freeth D, Koppel I, Reeves S, Barr H. A best evidence systematic review of interprofessional education: BEME guide no. 9. Med Teach. 2007;29(8):735-751.
4. Rodger S, Hoffman S, World Health Organization Study Group on Interprofessional Education and Collaborative Practice. Where in the world is interprofessional education? A global environmental scan. J Interprof Care. 2010;24(5):479-491.
5. Barr H. Interprofessional Education Today, Yesterday and Tomorrow. United Kingdom: Centre for the Advancement of Interprofessional Education; 2002.
6. World Health Organization. Framework for Action on Interprofessional Education & Collaborative Practice. Geneva, Switzerland: World Health Organization; 2010.
7. General Pharmaceutical Council. Consultation on Initial Education and Training Standards for Pharmacists. Great Britain: General Pharmaceutical Council; 2019.
8. General Medical Council. Domain 3: communication partnership and teamwork. 2019. https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice/domain-3---communication-partnership-and-teamwork.. Accessed November 2, 2019.
9. Nursing and Midwifery Council. Standards to Support Learning and Assessment in Practice: NMC Standards for Mentors, Practice Teachers and Teacher. United Kingdom: Nursing and Midwifery Council; 2008.
10. Biggs J. Enhancing teaching through constructive alignment. High Educ. 1996;32(3):347-364.
11. Reeves S, Boet S, Zierler B, Kitto S. Interprofessional education and practice guide no. 3: evaluating interprofessional education. J Interprof Care. 2015;29(4):305-312.
12. Freeth D, Hammick M, Koppel I, Reeves S, Barr H. A Critical Review of Evaluations of Interprofessional Education. United Kingdom: Higher Education Academy, Health Sciences and Practice Network; 2002.
13. Oates M, Davidson M. A critical appraisal of instruments to measure outcomes of interprofessional education. Med Educ. 2015;49(4):386-398.
14. Blue A, Chesluk B, Conforti L, Holmboe E. Assessment and evaluation in interprofessional education: exploring the field. J Allied Health. 2015;44(2):73-82.
15. Reeves S, Fletcher S, Barr H, et al. A BEME systematic review of the effects of interprofessional education: BEME guide no. 39. Med Teach. 2016;38(7):656-668.

16. Lapkin S, Levett-Jones T, Gilligan C. A systematic review of the effectiveness of interprofessional education in health professional programs. Nurse Educ Today. 2013;33(2):90-102.

17. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-269.

18. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club. 1995;123(3):A12-A13.

19. National Health Service. Five Year Forward View. England: National Health Service; 2014.

20. Joanna Briggs Institute. About us. https://joannabriggs.org/about. Accessed April 20, 2020.

21. Popay J, Roberts H, Sowden A, et al. Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC methods programme. 2006.

22. Frost J, Hamer D, Nunez L, et al. The intersection of professionalism and interprofessional care: development and initial testing of the interprofessional professionalism assessment (IPA). J Interprof Care. 2019;33(1):102-115.

23. Abbott-Anderson K, Hsinhuei Sheen C, Burk BN. Spring EngAGEment: an interprofessional collaboration to enhance community interaction and university education. Perspect ASHA Special Int Grps. 2019;4(1):91-96.

24. Acquavita SP, Anne Van Loon R, Smith R, et al. The SBIRT interprofessional curriculum and field model. J Soc Work Pract Addict. 2019;19(1-2):10-25.

25. Iverson L, Todd M, Ryan Haddad A, et al. The development of an instrument to evaluate interprofessional student team competency. J Interprof Care. 2018;32(5):531-538.

26. Wong AKC, Wong FKY, Chan LK, Chan N, Ganotice FA, Ho J. The effect of interprofessional team-based learning among nursing students: a quasi-experimental study. Nurse Edu Today. 2017;53:13-18.

27. Racic M, Joksimovic BN, Cicmil S, et al. The effects of interprofessional diabetes education on the knowledge of medical, dentistry and nursing students. Acta Med Acad. 2017;46(2):145-154.

28. Popkess A, Poirier TI, Wilhelm M, et al. Interprofessional error disclosure simulation for health professional students. Clinim Simul Nur. 2017;13(11):573-582.

29. Monteiro K, Dumenco L, Collins S, et al. An interprofessional education workshop to develop health professional student opioid misuse knowledge, attitudes, and skills. J Am Pharm Assoc. 2017;57:511S-517.

30. Feather R, Carr D, Garletts D, Reising D. Nursing and medical students teaming up: results of an interprofessional project. J Interprof Care. 2017;31(5):661-663.

31. Lie D, Richter-Lagra Forest C, Walsh A, Lohenny K. When less is more: validating a brief scale to rate interprofessional team competencies. Med Educ Online. 2017;22(1):1314751.

32. Reed T, Horsley TL, Muccino K, et al. Simulation using TeamSTEPPS to promote interprofessional education and collaborative practice. Nurse Educ. 2017;42(3):E1-E5.

33. Randall K, Steinheider B, Isaacson M, et al. Measuring knowledge, acceptance, and perceptions of telehealth in an interprofessional curriculum for student nurse practitioners, occupational therapists, and physical therapists. J Interact Learn Res. 2016;27(4):339-353.

34. Kent F, Martin N, Keating JL. Interprofessional student-led clinics: an innovative approach to the support of older people in the community. J Interprof Care. 2016;30(1):123-128.

35. Hwang Ji, Yoon TY, Jin HJ, Park Y, Park JY, Lee BJ. Patient safety competence for final-year health professional students: perceptions of effectiveness of an interprofessional education course. J Interprof Care. 2016;30(6):732-738.

36. Brushers V, Erickson JM, Blackhall L, Owen JA, Thomas SM, Conaway MR. Measuring the impact of clinically relevant interprofessional education on undergraduate medical and nursing student competencies: a longitudinal mixed methods approach. J Interprof Care. 2016;30(4):448-457.

37. Zhang C, Miller C, Volkman K, Meza J, Jones K. Evaluation of the team performance observation tool with targeted behavioral markers in simulation-based interprofessional education. J Interprof Care. 2019;29(3):202-208.

38. Djukic M, Adams J, Fulmer T, et al. E-Learning with virtual teammates: a novel approach to interprofessional education. J Interprof Care. 2015;29(5):476-482.

39. Dauenhauer JA, Glose S, Watt C, Design, delivery, and outcomes from an interprofessional fall prevention course. Gerontol Geriatr Educ. 2015;36(3):278-301.

40. Paige JT, Garbee DD, Kozenko V, et al. Getting a head start: high-fidelity, simulation-based operating room team training of interprofessional students. J Am Coll Surg. 2014;218(1):140-149.

41. Chiu C-J. Development and Validation of Performance Assessment Tools for Interprofessional Communication and Teamwork (PACT). University of Washington: ProQuest LLC; 2014.

42. Anderson ES, Thorpe LN. Students improve patient care and prepare for professional practice: an interprofessional community-based study. Med Teach. 2014;36(6):495-504.

43. Saini B, Shah S, Keary P, Bosnic-Anticevich S, Grootjans J, Armour C. Instructional design and assessment: an interprofessional learning module on asthma health promotion. Am J Pharm Educ. 2011;75(2):30.

44. Nango E, Tanaka Y. Problem-based learning in a multidisciplinary group enhances clinical decision making by medical students: a randomized controlled trial. J Med Dent Sci. 2010;57(1):109-118.

45. LeFlore JL, Anderson M. Alternative educational models for interdisciplinary student teams. Simul Healthc. 2009;4(3):135-142.

46. Lindblom P, Scheja M, Torell E, Astrand P, Felländer-Tsai L. Learning orthopaedics: assessing medical students’ experiences of interprofessional training in an orthopaedic clinical education ward. J Interprof Care. 2007;21(4):413-423.

47. Reeves S, Freeth D, McCorrie P, Perry D. ’It teaches you what to expect in future …’: interprofessional learning on a training ward for medical, nursing, occupational therapy and physiotherapy students. Med Educ. 2002;36(4):337-344.

48. Liang En W, Koh GC, Lim VK. Caring for underserved patients through neighborhood health screening: outcomes of a longitudinal, interprofessional, student-run home visit program in Singapore. Acad Med. 2011;86(7):829-839.

49. Luetsmer MT, Cloud BA, Youdas JW, Pawlina W, Lachman N. Simulating the multi-disciplinary care team approach: enhancing student understanding of anatomy through an ultrasound-anchored interprofessional session. Anat Sci Educ. 2018;11(1):94-99.

50. Rotz ME, Duenas GG, Zanoni A, Grover AB. Designing and evaluating an interprofessional experiential course series involving medical and pharmacy students. Am J Pharm Educ. 2016;80(5):85.

51. Myers Virtue S, Rotz ME, Boyd M, Lykon J, Waldron EM, Theodorou J. Impact of a novel interprofessional dental and pharmacy student tobacco cessation education programme on dental patient outcomes. J Interprof Care. 2018;32(1):52-62.

52. Branch-Mays GL, Pittenger AL, Williamson K, Milone A, Hein E, Thierer T. An interprofessional education and collaborative practice model for dentistry and pharmacy. J Dent Educ. 2017;81(12):1413-1420.

53. Sadowski CA, Li JC, Pasay D, Jones CA. Interprofessional peer teaching of pharmacy and physical therapy students. Am J Pharm Educ. 2015;79(10):155.
54. Jones VE, Karydis A, Hottel TL. Dental and dental hygiene intraprofessional education: a pilot program and assessment of students' and patients' satisfaction. J Dent Educ. 2017;81(10):1203-1212.
55. Berkowitz O, Brisotti MF, Gascon L, Henshaw M, Kaufman LB. The impact of an interprofessional oral health curriculum on trainees. J Physician Assist Educ. 2017;28(1):2-9.
56. Anders PL, Scherer YK, Hatton M, Antonson D, Austin-Ketch T, Campbell-Heider N. Using standardized patients to teach interprofessional competencies to dental students. J Dent Educ. 2016;80(1):65-72.
57. McLelland G, Perera C, Morphet J, et al. Interprofessional simulation of birth in a non-maternity setting for pre-professional students. Nurse Educ Today. 2017;58:25-31.
58. Thistlethwaite J, Dallest K, Moran M, et al. Introducing the individual Teamwork Observation and Feedback Tool (iTOFT): Development and description of a new interprofessional teamwork measure. J Interprof Care. 2016;30(4):526-528.
59. Bandali KS, Craig R, Ziv A. Innovations in applied health: evaluating a simulation-enhanced, interprofessional curriculum. Med Teach. 2012;34(3):e176-e184.
60. Curran V, Hollett A, Casimiro LM, et al. Development and validation of the interprofessional collaborator assessment rubric (ICAR). J Interprof Care. 2011;25(5):339-344.
61. Carraccio C, Wolfshal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. Acad Med. 2002;77(5):361-367.
62. Interprofessional Education Collaborative. Core Competencies for Interprofessional Collaborative Practice: Report of an Expert Pane. Washington, DC: Interprofessional Education Collaborative; 2011.
63. Streiner DL. Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. Journal of Personality Assessment. 2003;80(1):99-103.
64. Phillips AC, Lewis LK, McEvoy MP, et al. Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). BMC Med Educ. 2016;16(1):237.
65. Ten Cate O, Regehr G. The power of subjectivity in the assessment of medical trainees. Acad Med. 2019;94(3):333-337.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Almoghirah H, Nazar H, Illing J. Assessment tools in pre-licensure interprofessional education: A systematic review, quality appraisal and narrative synthesis. Med Educ. 2021;55:795–807. https://doi.org/10.1111/medu.14453