



A Delphi study to validate competency-based criteria to assess undergraduate midwifery students' competencies in the maternity ward[☆]

M. Embo, RN, RM, PhD (Head Midwifery Department, Researcher)^{a,b,*}, K. Helsloot, RM, MSc (Lecturer, Researcher)^a, N. Michels, MD, PhD (General practitioner, Assistant Professor)^c, M. Valcke, PhD (Professor of Instructional Sciences)^d

^a Midwifery Department, University College Arteveldehogeschool Ghent, Voetweg 66, 9000 Ghent, Belgium

^b Department of Educational Studies, Faculty of Psychology and Educational Sciences, Ghent University, H. Dunantlaan 2, 9000 Ghent, Belgium

^c Skills lab and Centre for General Practice, Faculty of Medicine and Health Sciences Antwerp, Universiteitslaan 1, 2610 Antwerp, Belgium

^d Department of Educational Studies, Faculty of Psychology and Educational Sciences, Ghent University, H. Dunantlaan 2, 9000 Ghent, Belgium

ARTICLE INFO

Keywords:

Competency-based education
Competency-based assessment
Assessment criteria
Delphi-study
Midwifery education
Postnatal care

ABSTRACT

Background: workplace learning plays a crucial role in midwifery education. Twelve midwifery schools in Flanders (Belgium) aimed to implement a standardised and evidence-based method to learn and assess competencies in practice. This study focuses on the validation of competency-based criteria to guide and assess undergraduate midwifery students' postnatal care competencies in the maternity ward.

Method: an online Delphi study was carried out. During three consecutive sessions, experts from workplaces and schools were invited to score the assessment criteria as to their relevance and feasibility, and to comment on the content and their formulation. A descriptive quantitative analysis, and a qualitative thematic content analysis of the comments were carried out. A Mann-Whitney U-test was used to investigate differences between expert groups.

Findings: eleven competencies and fifty-six assessment criteria were found appropriate to assess midwifery students' competencies in the maternity ward. Overall median scores were high and consensus was obtained for all criteria, except for one during the first round. Although all initial assessment criteria ($N=89$) were scored as relevant, some of them appeared not feasible in practice. Little difference was found between the expert groups. Comments mainly included remarks about concreteness and measurability.

Conclusion: this study resulted in validated criteria to assess postnatal care competencies in the maternity ward.

Introduction

The move to competency-based education fits the aim to enhance the readiness of graduates, to ensure the provision of safe care, and to push their learning orientation throughout their professional career (Frank et al., 2015). Health care competency frameworks are, at the core, educational initiatives to improve patient care. They help to articulate entry-to-practice competencies (Frank et al., 2015). Clearly stated educational outcomes help students' understanding of what is expected of them and guide staff to focus precisely on what students have to achieve during their clinical practice (Shumway and Harden, 2003; Ossenbreg and Henderson, 2015). Assessment plays a major role in competency-based education. Competency-based education and

performance assessment are closely related paradigms (Shumway and Harden, 2003). The decision whether students have attained requisite learning outcomes is based on their actual performance. Assessing competencies in practice settings therefore requires competency-based assessment instruments (Frank et al., 2010) that offer criteria to assess student performance (Gulikers et al., 2010; Richardson and Flynn, 2011; Fastré et al., 2014).

Current socio-cultural theories of workplace learning claim that learning and learning outcomes result from active participation in activities and in interaction with complex and dynamic systems of the clinical work environment (Mann, 2011; Govaerts and van der Vleuten, 2013). Assessing learners in a competency-based education context requires a radically different (a) integrated and (b) holistic assessment

[☆] Names of the institutions at which the research was conducted: Midwifery Department, University College Arteveldehogeschool Ghent, Belgium.

* Corresponding author at: Midwifery Department, University College Arteveldehogeschool Ghent, Voetweg 66, 9000 Ghent, Belgium.

E-mail addresses: mieke.embo@arteveldehs.be (M. Embo), kaat.helsloot@arteveldehs.be (K. Helsloot), nele.michels@ua.ac.be (N. Michels), martin.valcke@UGent.be (M. Valcke).

<http://dx.doi.org/10.1016/j.midw.2017.07.005>

Received 15 February 2017; Received in revised form 26 June 2017; Accepted 4 July 2017

0266-6138/ © 2017 Elsevier Ltd. All rights reserved.

method. The integrated approach acknowledges competency as a complex combination of knowledge, attitudes, skills and personal values that are demonstrated at a defined level of proficiency in the particular context of practice (Leung, 2002; Stoof et al., 2002; Lane, 2010; Yanhua and Watson, 2011). The holistic approach on the other hand, takes into account the cultural and social context in which competencies are assessed. The concept ‘professional competence’ is a holistic term that considers that the competent person ‘not only possesses the requisite competencies but is also able to use them and make appropriate decisions and judgements according to the context’ (Epstein and Hundert, 2002).

Assessment criteria for competency-based education are either competency-based (what the student is able to do) or performance-based (what the student has to do) (Fastré et al., 2014). Competency- and performance-based criteria should be seen as a continuum linked to different levels in professional education or training. Competency-based criteria are less meaningful for novice students compared to advanced students, because the former have not yet attained the required level of integration of knowledge, skills and attitudes (Fastré et al., 2010). There is however, a direct relationship between competency-based and performance-based criteria, because the latter specify context-specific performance in relation to the competencies (Crossley and Jolly, 2012). The nature and quality of effective assessment criteria in competency-based education is an under-researched topic (Fullerton et al., 2016). Some studies revealed performance-based criteria resulted in more accurate assessment because of the directly observable and task-specific nature of these criteria (Idrissi et al., 2016). This confirms the results of Lurie (2012), who indicates that it is important to define assessment criteria in terms of the situations to which they are relevant, rather than as global personal characteristics (Fastré et al., 2014). Similar concepts for assessment criteria were found in the literature, such as observable or measurable behaviour standards (Frank, 2011) and behavioural cues (Ossenberg and Henderson, 2015); often used interchangeably. They all share the fact that they provide more clarity and transparency to the assessment process in that they describe routine behaviours pertinent to care delivery.

Providing high-quality measurable outcomes is an important challenge to improve competency-based assessment in practice. Nowadays, educators are increasingly mandated to measure and report detailed competence measures but they are hindered by the lack of valid and reliable criteria across competencies (Holmboe and Snell, 2011; Kalet et al., 2016). Lack of valid assessment criteria is one of the barriers to (1) clearly understand and communicate standards of practice from regulatory bodies (Ossenberg and Henderson, 2015); (2) align assessment cultures with the profession's values (Watling 2016); (3) make assessment authentic or relevant to students' future work (Harrison et al., 2016); (4) promote learner centeredness (Frank 2011); (5) address the needs of learners who do not meet standards or expectations (Kalet et al., 2016); (6) prevent reductionism (breaking

competencies down into the smallest observable units of behavior, creating endless nested lists of abilities that frustrate learners and teachers alike) (Frank et al., 2010); and (7) involve stakeholders in the development of logbooks which ensures local acceptability and feasibility with the additional benefit of combining training and agreed standards (Schüttpelz-Brauns et al., 2016).

The present study aimed to validate assessment criteria for a newly defined competency framework for undergraduate midwifery students' postnatal care in the maternity ward. It builds on a collaboration between twelve midwifery schools who designed a new competency-based educational profile (Flemish Education Council, 2014; Embo and Valcke, 2016). This education profile challenged midwifery educators to reconsider workplace learning and assessment during midwifery practice. It was the catalyst to standardise education, to introduce a workplace learning model based on current socio-cultural learning theories and to put an end to adopting too many different assessment instruments and criteria (Embo and Valcke, 2016). The midwifery schools agreed to gradually implement a competency-based continuous workplace learning model that integrates learning and assessment in a six-step method (Embo et al., 2015). The Achilles heel is the competency framework with concrete and easy-to-use criteria that proved being suitable to monitor learning progress and to support assessment of competencies and professional competence (Embo et al., 2010, 2014). To validate the assessment criteria educational and professional experts could be involved (Epstein, 2007; Ossenberg and Henderson, 2015). The study is embedded in a research project funded by the Mustela Foundation Award 2015 and the education department of the University College Arteveldehogeschool Ghent.

Method

Design

We adopted an online Delphi-survey study to collect opinions from a wide range of experts in order to develop a consensus about the criteria. Consensus methods are used by healthcare professionals and educators because of their presumed capacity to extract the profession's ‘collective knowledge’ which is often considered tacit knowledge that is difficult to verbalise and formalise (Foth et al., 2016). Consensus methods are supposed to offer quantitative estimates through qualitative approaches, characterised by anonymity, iteration, controlled feedback, statistical group response and structured interaction (Foth et al., 2016). The competencies and assessment criteria were presented to a multidisciplinary expert panel over three iterations as depicted in Fig. 1. The study was conducted in Flanders (Belgium) between April 4th and May 29th 2016. It was approved by the Ethics Committee of the Ghent University Hospital (B-Nr 2016/0250) and the Ethical Review Board of the Dutch Association for Medical Education (NVMO, the Netherlands-Nr 657). Informed consent was obtained from all participants.

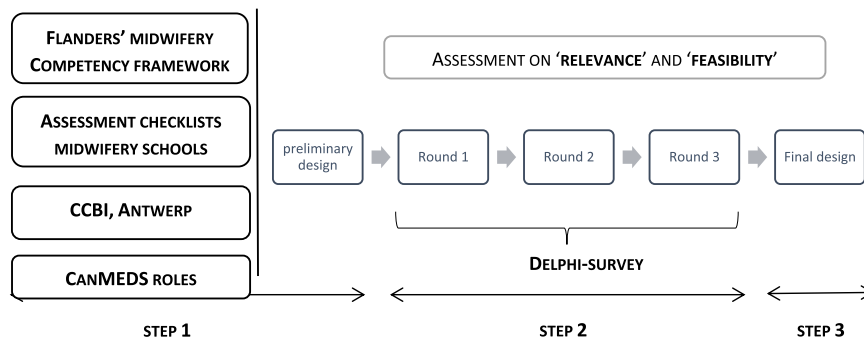


Fig. 1. Overview of the method.

Delphi expert panel

A group of 90 educational and workplace experts was invited to participate in the Delphi study, taking into account a dropout of 70%. Expert selection built on two criteria: at least three-year experience as a professional and as an assessor of midwifery students in maternity wards. Educational experts - i.e. midwifery teachers and experts in competency-based education - were selected from all current midwifery schools in Flanders ($N=12$). Workplace experts (midwives, obstetricians and paediatricians) were purposefully selected from 15 out of the 70 maternity wards in Flanders. We wanted to include both regional and speciality settings; the 15 wards included university and non-university hospitals as well as wards from regional hospitals. In each hospital setting, the heads of the midwifery departments, maternity wards and medical staff assisted with the recruitment of the experts. Forty-eight experts were willing to participate in the study. They all received information about the study, a copy of the starting document, access to the online surveys (Lime Survey), and practical tips on how to complete the survey. Informed consent was obtained at the start of the Delphi round.

Delphi procedure

The first two authors designed the starting document and used the validated midwifery competency framework (Flemish Education Council, 2014) that also forms the basis for the Belgian professional profile (Federal Government of Health, 2015). To facilitate reading of competencies, the keywords from the official documents of the Ministers for Education and Health were copied (see also Table 3). They collected all the criteria from the midwifery assessment instruments actually being applied in the different midwifery schools in Flanders, the 'CanMEDS Competency Based Inventory' or CCBI used to assess medical doctors at the University of Antwerp (Michels et al., 2012), and the latest CanMEDS framework (Frank et al., 2015; Fig. 1). One author linked these criteria to specific competencies ($N=11$) and both authors discussed this list. Overlapping criteria were removed from the list and special attention was given to an unambiguous formulation. As such, the resulting instrument comprised of criteria-linked statements related to 11 competencies, referring in total to 89 assessment criteria.

In three consecutive Delphi rounds, experts were asked to consider two main research criteria (relevance and feasibility) and to comment on the content and formulation of assessment criteria, bearing in mind its use in undergraduates' final year assessment. Specifically, experts were asked whether competencies and assessment criteria were: 1) appropriate for assessing observable behaviour in the maternity ward (validity), 2) sufficiently concrete for reliable assessment (reliability), 3) formulated without an overlap (sensitivity), 4) feasible or measurable in practice (feasibility) and 5) acceptable for mentors and students (acceptability). In the literature, these questions have been found adequate to ensure criterion quality (Campbell et al., 2003). The first three questions addressed 'relevance' and the last two questions addressed 'feasibility' (Table 1). The experts gave ratings, building on a 6-point Likert scale (1 = strongly disagree; 6 = strongly agree) when answering the five above questions.

The first Delphi round finished after one week. Data were analysed and discussed by the researchers. In the second round the outcomes of the first round, were returned - anonymously - to all the first-round respondents ($N=48$). They received an analysis of frequencies and remarks, as well as a description of reasons why and how criteria were deleted and/or adapted. Provided with this input, the experts were invited to score once again all the criteria as to their relevance and feasibility, taking into account median scores, level of consensus and the remarks of the experts. Additionally, experts were asked to provide new suggestions and remarks (Michels et al., 2012), as well as to accept the proposal to delete specific criteria. The procedure during the third

Table 1
Delphi study: research criteria and questions.

Research Criteria	Questions
research criterion 1: Relevance (validity, reliability, sensitivity)	1) Is the formulation of the competencies and assessment criteria appropriate to evaluate observed behaviour in the workplace? 2) Are the competencies and assessment criteria formulated sufficiently concrete and assessable? 3) Is there an overlap between competencies or assessment criteria?
Research Criterion 2: Feasibility (feasibility and acceptability)	4) Are the competencies or assessment criteria measurable in the workplace? 5) Is the formulation acceptable for mentors and students?

round was similar to the second round. Analysis of the third-round data resulted in a set of validated competency-based assessment criteria.

Delphi data analysis

A descriptive analysis of the scores on relevance and feasibility and a qualitative thematic content analysis of the comments were carried out. Descriptive statistics (medians, level of consensus) were calculated after each round.

Consensus indicated the level of agreement in responses between experts. Consensus was defined when more than 30% of the scores were in the first or third tertile and more than 30% of the scores occurred were not simultaneously observed in the third and first tertile. When consensus was reached in relation to a specific criterion AND when a median score of 4 or larger was attained, the criterion was considered as valid. When no consensus was attained or the median score < 4, this specific criterion was deleted. When specific recommendations were offered by experts, an adapted version was developed and presented in the consecutive round. The predefined stop criterion was achievement of consensus and saturation of the round's comments (Foth et al., 2016). The suggestions and comments of the experts were registered anonymously and literally. At the end of the Delphi study, thematic analysis was carried out of all comments (Braun and Clark, 2006). Building on the proportions of themes emerging from the analysis, Mann-Whitney U-test was performed with SPSS (statistics 22) to investigate potential differences in scoring behaviour of the two types of experts involved in the validation procedure (school experts and workplace experts).

Findings

Thirty-two of the 48 experts actively responded in the first Delphi-round (response rate of 67%). The other 16 experts expressed their interest to participate, and some started responding to the survey but did not finish due to a lack of time, the duration of the survey, or other reasons that were not specified. In the second and third Delphi-round, 34 experts (71%) and 28 experts (58%) completed the survey. The composition of the expert panel varied in the subsequent rounds (1/2/3) as documented in Table 2.

Delphi round 1

Median scores reflected that all competencies ($N=11$) were perceived as relevant and feasible in the maternity ward (median scores ≥ 4). A different picture emerged when looking at the results in relation to the assessment criteria ($N=89$). Although all criteria were rated as relevant, four criteria seemed not feasible. A Mann-Whitney U-test revealed little differences in opinion between the two types of experts (workplace and school). Two assessment criteria, 'Knowledge and

Table 2
Composition of the expert panel.

	Delphi round 1	Delphi round 2	Delphi round 3
N=	32	34	28
Obstetrician	0	1	1
Paediatrician	1	2	1
Midwifery practitioner	20	18	7
Midwifery teacher	5	9	13
Educational expert	6	4	6
Job combination Midwifery practitioner & teacher	11	10	4
University hospital	3	3	2
Non-University hospital	9	9	8
Not in hospital	20	22	18
Midwifery practice	21	21	16
Education (school)	22	23	19
Supervision experience > 10 years	20	16	15
Assessing experience > 10 years	17	16	15
Professional experience at the maternity ward > 10 years	17	20	15
Professional experience teaching midwifery > 10 years	8	10	10

insight in medication' and 'Assesses a potential risk situation adequately', were significantly rated as more relevant by the workplace than the school experts. Interestingly, although the overall median scores of the first round were high, many remarks and suggestions were presented ($N=770$).

Analysis of the themes reflected in the comments in relation to the five quality questions (validity, reliability, sensitivity, feasibility and acceptability), showed the vast majority were about the assessment criteria and not about the competencies. Experts raised concerns about the extent to which the criteria were sufficiently 'concrete' ($N=233$ or 30.3%), reflected too much 'overlap' ($N=56$ or 7.3%), or were less 'appropriate' for the maternity ward ($N=32$ or 4.5%). Other comments referred to the feasibility, perceiving them as 'unmeasurable' because of limited or infrequent learning opportunities during practice ($N=279$ or 36%). Most comments were conveyed about the assessment criterion 'Prescribes medication'. Twenty-two experts stated this criterion was not observable in the context of the maternity ward. Another feasibility problem stated it was not 'acceptable' that students carried out some learning opportunities autonomously ($N=20$ or 2.6%), e.g. 'hand over the care of a patient/client with complex pathology to another health care professional' was only acceptable when under supervision of a registered midwife. Next to comments about the relevance and feasibility, experts made suggestions to improve the quality of the assessment criteria: adding criteria to make the instrument more concrete ($N=4$ or 0.5%), combining criteria in order to reduce the extent of the assessment instrument ($N=31$ or 4%), changing terminology (e.g. learning outcome instead of competency) ($N=16$ or 2%), and formulating attitudes as independent criteria at the end of the list ($N=3$ or 0.38%). Finally, some minor comments were made about the academic career level and assessment (not assessable in the first year). Also, a proportion of comments was outside the scope of the present study ($N=99$ or 12.9%).

Building on the comments, the researchers decided to adapt the formulation of 17 assessment criteria (including 13 criteria with a median score of four) and to delete nine overlapping assessment criteria. Criteria with a median score of five or six were accepted as valid, without changes.

Delphi round 2

During the second round, all assessment criteria resulting from the first round were presented again to the panel. The experts were invited

to react to the same 5 quality questions. They were additionally informed about the changes to the 17 criteria and the deletion of nine criteria; they could comment on this.

In this round, overall median scores for relevance and feasibility were high (median scores ≥ 4), with the exception of the mean feasibility score for the assessment criterion 'Prescribes medication' (median score 2). Consensus was obtained for all of the assessment criteria and the majority of experts (85%) accepted to drop the nine overlapping criteria. Mann-Whitney U-test results reflected no significant differences between types of experts.

Content analysis of the comments ($N=106$) mirrored similar results in relation to relevance and feasibility of the criteria: weak appropriateness in view of assessment ($N=37$ or 34.9%), not sufficiently concrete ($N=15$ or 14%), formulated with an overlap ($N=2$ or 1.9%), not appropriate ($N=1$ or 0.94%), not measurable because of too limited learning opportunities during practice ($N=33$ or 31.1%), and not possible to perform autonomously ($N=3$ or 2.83%). Also during this round, the experts made suggestions for improvement: add a list to explain the criteria ($N=1$ or 0.9%), split criteria ($N=15$ or 14.1%) and consider attitudes as discrete criteria ($N=2$ or 1.9%). A small number of comments questioned the level at which the assessment could be carried out or were outside the focus of the present study ($N=7$ or 6.6%).

Analysis of these findings resulted in a minor revision of the instrument, by adding five new assessment criteria and adaptation of five criteria. This new list was represented to the expert panel during the third round.

Delphi round 3

In this third and final Delphi-round, all assessment criteria turned out to be validated (relevance and feasibility) and consensus was obtained. All but one of the assessment criteria achieved median scores of 5 or 6. The single criterion reflecting a feasibility score of 4 (both types of experts) was 'Acts in accordance with the legal and ethical rules concerning the profession of midwifery'. No significant differences between the expert types were revealed by the Mann-Whitney U-test.

In the third round, a rather small number of comments were put forward by the experts ($N=31$): focusing on weak appropriateness ($N=10$ or 32.2%), not sufficiently concrete ($N=12$ or 38.7%), and formulated with an overlap ($N=2$ or 6.4%). Comments focusing on poor feasibility addressed the problem of limited learning opportunities during the internship. Some minor suggestions were made to split criteria ($N=3$ or 9.6%), and the fact some criteria were impossible to apply in a first year ($N=2$ or 6.4%). Some experts explicitly commented ($N=5$ or 6.1%) that they could observe how their input was considered by the researchers and that the subsequent adjustments enhanced the quality of the assessment instrument.

Building on the three rounds, the final validated instrument consists of 11 competencies and 56 assessment criteria to assess postnatal care competencies in the maternity ward (Table 3).

Discussion

Providing clear learning objectives with observable and measurable outcomes is an important challenge to improve competency-based education in practice settings (Holmboe and Snell, 2011). The present study aimed to validate a set of competency-based criteria to assess undergraduate midwifery students' postnatal care competencies in the maternity ward. We adopted an online Delphi-survey design to collect multidisciplinary expert opinions. A descriptive analysis of relevance and feasibility and a qualitative thematic content analysis of the comments were carried out.

In total, three Delphi rounds were needed to validate the competencies and assessment criteria. The high overall median scores demonstrated that all competencies ($N=11$) and assessment criteria

Table 3

Validated competency-based assessment instrument: 11 competencies, 56 assessment criteria.

	Delphi-round	1	2	3
Competency 1: Physiology, care and guidance		✓		
Diagnoses, supports and promotes a physiological postnatal recovery and supports the care for the newborn. Promotes, supports and guides breastfeeding.				
1	Systematically collects all necessary data and (vital, physical, obstetric) parameters of the client.		✓	
2	Interprets the collected data correctly (and drafts a (differential) diagnosis).	✓		
3	Drafts a care plan to prepare the care.	✓		
4	Prepares the client correctly and collects the necessary materials.	✓		
5	Admits mother and newborn correctly after birth and has attention for continuity of care.			✓
6	Administers hygienic care to the mother correctly during a normal low-risk postpartum.			✓
7	Administers hygienic care to the newborn correctly during a normal low-risk postnatal period.			✓
8	Executes blood and/or specimen sampling correctly from the mother and gives the necessary health information.			✓
9	Executes blood and/or specimen sampling correctly from the newborn and gives the necessary health information.			✓
10	Gives adequate and proper assistance and advice on breastfeeding.	✓		
11	Gives adequate and proper assistance and advice on the expression of breastmilk (manually, electric), and the storage, handling and administration of expressed breast milk.	✓		
12	Gives adequate and proper assistance and advice on the administration, preparation, storage and handling of formula.	✓		
13	Gives information and training on pelvic floor education correctly.			✓
14	Cleans up material and room correctly and installs the client comfortably.	✓		
15	Prepares the client correctly for discharge and attends to a proper follow-up in function of continuity of care.	✓		
16	Orally reports correctly and adequately.	✓		
17	Reports correctly and adequately in writing (medical record).	✓		
18	Calculates, prepares and administers medication to the client and gives the necessary health information.	✓		
19	Has knowledge and understanding of the indication, side effects, dose and the way of administering medication.	✓		
Competency 2: Risk detection/selection				✓
Detects risks and complications within the fields of the profession (autonomously), acts adequately, consults with physicians and other caregivers and refers in a timely manner.				
20	Detects (potential) risks and complications that can present a threat to the physiological course of the postnatal period.			✓
21	Consults a midwife or physician based on risk detection (and refers on adequately and in a timely manner if indicated).	✓		
22	Responds correctly and appropriately in emergency situations (within his/her own competences) until the midwife/physician takes over.	✓		
Competency 3: Pathology, care and guidance		✓		
Provides adequate care and guidance in elevated risk situations or when complications occur, on the authority of and in cooperation with physicians and other caregivers, within the domain of obstetrics. Administers the necessary care, autonomously if necessary.				
23	Prepares the care and client correctly and at a good pace in elevated risk situations and gives the necessary health information.			✓
24	Administers hygienic care to the mother correctly after cesarean section or after complications.			✓
25	Implements the nursing care correctly and gives the necessary health information (i.e. IV, bladder catheter, wound care).			✓
26	Administers blood and blood derivatives correctly and gives the necessary health information.	✓		
27	Administers phototherapy correctly and gives the necessary health information.	✓		
Competency 4: Psychosocial context		✓		
Situates the mother and her environment in the family, social, societal and cultural contexts and interacts targetedly, taking into account the diversity and intercultural perception. Recognises psychosocial crisis situations and refers on to physicians and other caregivers when appropriate.				
28	Shows respect and empathy for the cultural background, values and standards of the client, the family and the environment.			✓
29	Offers support for baby blues, bonding, the transition to parenthood and relational challenges.			✓
29	Recognises psychosocial crisis situations, offers support and responds correctly.			✓
Competency 5: Ethics and legislation		✓		
Acts ethically within the boundaries of deontology and legislation.				
30	Opens up ethical situations for discussion.			✓
31	Acts in accordance with the legal and deontological rules concerning the profession of midwifery.			✓
32	Manages personal values and beliefs adequately within midwifery care.			✓
33	Respects the professional secrecy.	✓		
Competency 6: Health promotion		✓		
Targets on prevention, autonomously and in cooperation, and stimulates the (physical, psychosocial, relational and sexual) health of mother and child in accordance with the principles of health promotion.				
34	Provides adequate and timely information and advice and helps families to make informed choices.	✓		
35	Empowers the mother and her family and stimulates self-care.			✓
36	Provides adequate information on the physiological changes of mother and newborn and teaches (self-) care skills.			✓
Competency 7: Collaboration and communication with health care providers		✓		
Ensures optimal interdisciplinary cooperation and adequate and up-to-date practical organisation. Communicates both orally and in writing in a professional and adequate way with colleagues, physicians, and other caregivers.				
37	Communicates correctly (orally and/or in writing) with colleagues, physicians and other caregivers.			✓
38	Co-operates with colleagues, physicians and other caregivers to ensure the continuity and efficiency of the care.			✓
39	Executes administrative, organisational and coordinating tasks correctly.			✓
Competency 8: Coaching		✓		
Coaches and supports colleagues (fellow students) of his/her own profession and other professions, in order to ensure quality care.				
40	Gives feedback appropriately.			✓
41	Is a role model for others.			✓
Competency 9: Professional development		✓		
Works on permanent profiling and the professionalisation of the profession by constantly critically questioning his/her own professional functioning and by permanent training. Reflects on his/ her own practice and processes these reflections in his/her professional actions.				
42	Self-reflects and adjusts his/her actions if necessary.			✓
43	Seeks feedback adequately and deals with it constructively.			✓
45	Maintains a personal development plan and actively works on his/her own professional development.			✓
Competency 10: Evidence-based care				
Provides evidence-based care by integrating new scientific understandings and by participating in practice-oriented scientific research.				

(continued on next page)

Table 3 (continued)

		Delphi-round		
		1	2	3
46	Analyses the care critically.			✓
47	Acts correctly based on evidence-based midwifery practice.			✓
48	Integrates the theory learned into practice.		✓	
49	Co-operates with the registration of data meant for optimising perinatal care.		✓	
Competency 11: Quality of care				
Autonomously provides professional, high-quality and safe care. Develops a personal vision with regard to safeguarding and promoting the quality of the care and contributes to the implementation of innovations.				
50	Applies the following basic principles correctly: hygiene, sterility, comfort, safety, ergonomics, economy, ecology.		✓	
51	Uses a systematic approach to plan and execute care.		✓	
52	Organises the overall care for the client correctly.		✓	
53	Is reliable and has a sense of responsibility.			✓
54	Is independent and takes proper initiative.			✓
55	Is decisive.			✓
56	Is flexible.			✓

✓ Accepted in round 1, 2 or 3.

($N=89$) were already considered relevant and feasible in the first round (median score of $\geq 4/6$), with the exception of feasibility for four assessment criteria. Importantly, there was a high level of consensus between experts and expert-groups. Michels et al. (2012) found comparable results when validating the competency-based inventory applicable for workplace assessment of medical doctors in training, based on the CanMEDS profile. In the present study, small differences in opinion were found in the first round for the criteria 'knowledge and insight in medication' and 'assesses a potential risk situation adequately'. These criteria are essential to guarantee patient safety and are rated as more relevant by workplace than the school experts. These results might be linked to the tension between education and patient safety in workplace learning (Dornan et al. 2012). The results present a very positive picture, but the researchers had to deal with a large amount of comments in relation to the assessment criteria. This invoked a revision process to delete, adapt and add criteria during the three rounds. The large amount of comments might reflect the concerns of users when dealing with assessment of competencies, as referred to in the introduction.

First, most comments reflected concerns about the feasibility of the criteria. Especially concerns about limited or infrequent learning opportunities during internships could be identified (36.23%/ 31.1%/ 32.5%). A challenge for workplace learning is indeed finding sufficient opportunities for students to incorporate competencies. Limited learning opportunities are inherent to working in unstructured learning environments (Dornan et al., 2012). This reality emphasises the importance of logbooks (Schüttelz-Brauns et al., 2016) and the design of a consistent programmatic assessment approach (Schuwirth and van der Vleuten, 2011; Van der Vleuten et al., 2012), incorporating simulated standardised assessment for those criteria that cannot be assessed in the workplace (Cooper et al. 2012; Van der Vleuten and Heeneman, 2016).

Building on the results, we also feel it critical to focus on the formulation and the number of criteria. Similar to the CanMEDS change in 2015, we observed an emphasis on accessible language that supports practical application of the criteria. This has also been reported by other researchers (Ossenberg and Henderson, 2015; Kalet et al., 2016). Not surprisingly, we found contradicting perceptions in the comments (adding versus deleting, and combining versus splitting). We can explain this by considering the continuum in performance- versus competency-based assessment; some experts might have had novices (single performances) or advanced students (performance of a variety of tasks) in mind (Fastré et al., 2010, 2014). These assumptions should be addressed in future research.

Finally, four sets of comments can be linked to the theoretical foundations of competency-based education. To start, experts sometimes worried about the feasibility of criteria linked to autonomous performance. As competency-based education has an entry-to-practice

approach (Frank et al., 2015), criteria that cannot be assessed independently during undergraduate education reiterate an on-going discussion about the length (three to five years) and level (bachelor or master) of midwifery programmes in higher education. Furthermore, it emphasises the importance of monitoring progress and giving guidance, also during unsupervised performances. The different steps in Embo's model are designed to facilitate continuous and self-regulated competency development in practice because it integrates learning, assessment and supervision (Embo et al., 2015). The assessment criteria, validated in this study, are observable during postnatal care because of the task-specific nature of these criteria. As revealed by Idrissi et al. (2016), future research in Flanders will be important to investigate the impact of these criteria on the capacity to enhance transparency of the assessment process and to stimulate active involvement in learning and assessment. They are also valuable to support a collaborative workplace learning culture where all stakeholders (students, supervisors and teachers) share responsibility for the quality of competency-based education in clinical practice. Second, recurring comments raised during the three rounds, reflected a preference for separating criteria to assess attitudes. This is in conflict with the basis of competency-based education, which aims to integrate knowledge, skills and attitudes (Fastré et al. 2014). This result fits earlier evidence that assessing competencies and professional competence in practice is difficult and that experts observe and assess novice students who have not yet attained the required level of integration (Fastré et al., 2010). This brings us to the third issue, the perceived difficulty to assess ethical competencies in practice. More work is needed to establish a holistic competency view in practice training settings. Also other authors stress how high-quality care implies health care professionals take into consideration legal, and ethical issues in complex health care environments (Ossenberg and Henderson, 2015). The fourth set of comments focused on the feasibility of the criterion 'Prescribes medication'. Twenty-four comments stated that this criterion was not observable in the maternity ward. This is not surprising because this outcome was recently legally regulated in 2015. This competency was immediately included in the curriculum but midwives are not yet adopting this competency in the maternity ward of hospitals.

The present Delphi-study built on a unique collaboration between professionals involved in midwifery education. As recommended in the literature (Foth et al., 2016), we included a large number of experts, with strong practical experience (> 10 years), from different geographical areas, representing diverse expertise areas (education and health care), working in different settings (university and non-university) and representing different disciplines (midwives, teachers, educational specialists, obstetricians and a pediatrician). The present authors state that a consensus from this group reflects sufficient alignment with the values of the profession, and contributes to the authenticity of the

assessment approach (Harrison et al., 2016). Nevertheless, we note that recruitment from workplace professionals was more difficult than expected. Especially medical doctors were difficult to motivate in participating (obstetricians ($N=2$), pediatrician ($N=1$)). This is a point of concern since midwives and medical doctors work together on a daily base in clinical setting. This brings us to the inter-professional education literature showing that much remains to be done to improve the collaboration between professions in competency-based health care education (Reeves et al., 2016).

Notwithstanding its strengths, the current study has also some limitations related to well-known Delphi-criticisms (see e.g., Boulkedid et al., 2011). Some of these criticisms include the question of 'imposed consent' (Haug, 2015), and that a Delphi-study possibly does not provide a space for discussion and debate (Foth et al., 2016). This can explain why we had to delete 4.5% of the comments considered 'irrelevant', and why we lack information to explain differences in concreteness as described above. Another criticism states that consensus methods reflect 'pool ignorance'. Due to the anonymity of the process, it has been argued that experts might not be accountable for the views they express and the judgements they make (Foth et al., 2016). Indeed, we could not trace and track the individual experts from one round to the other. Therefore, information about consistency in their answers is missing. The dropout of experts also needs addressing (Boulkedid et al., 2011). We invited 90 experts, reached the following response rates ($N=32/34/28$), a dropout percentage of 70%. This may be due to the length of the survey and the short completion periods for each round (3/2/2 weeks) (Trevelyan, 2015).

The lessons learned from this study are currently influencing the re-design of assessment criteria in other midwifery settings, where a similar validation Delphi study will be set-up. Building on the current and the next studies, future research will investigate if the adoption of the validated competency-based assessment instruments in Flanders will improve the competencies of newly graduated midwives and consequently the quality and safety of perinatal care.

Conclusion

High-quality competency-based assessment requires clear outcomes and valid assessment criteria. The present Delphi-study validated criteria to assess undergraduate midwifery students in the maternity ward with a multidisciplinary expert panel. After three Delphi rounds, 56 assessment criteria within 11 competencies were found appropriate to guide and assess midwifery students' competencies. Overall median scores and level of consensus were high and almost no differences were found between the expert groups. Many comments addressed current assessment problems: lack of feasibility because of limited learning opportunities, contradicting views on the number and formulation of criteria and limited knowledge of competency-based theoretical underpinnings. Future research is necessary to support a programmatic learning and assessment approach, to better understand the nature of effective competency-based assessment criteria and to enhance inter-professional competency-based education.

Ethical statement

Conflict of interest

The authors report no declarations of interest.

Ethical approval

The study was approved by the Ethics Committee of the Ghent University Hospital (B-Nr 2016/0250) and the Ethical Review Board on the Dutch Association for Medical Education (NVMO, the Netherlands-Nr 657). Informed consent was obtained from all participants.

Funding sources

The study was financed by the Mustela Foundation Award 2015 and by the University College Arteveldehogeschool Ghent (B), Educational Department.

Clinical Trial Registry and Registration number (if applicable)

Not applicable.

Acknowledgments

We thank the experts for their participation in this study. We also thank Linda Sweet, Associate Professor of the Flinders University Adelaide for editing this manuscript.

References

- Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., Alberti, C., 2011. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 6, 1–7.
- Braun, V., Clark, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 77–101.
- Campbell, S.M., Braspenning, J., Hutchinson, A., Marshall, M.N., 2003. Research methods used in developing and applying quality indicators of care. *British Medical Journal* 362, 816–819.
- Cooper, S., 2012. Simulation based learning in midwifery education: a systematic review. *Women and Birth* 25, 64–78.
- Crossley, J., Jolly, B., 2012. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Medical Education* 46, 28–37.
- Dornan, T., Muijtjens, A., Graham, J., Scherpbier, A., Boshuizen, A., 2012. Manchester Clinical Placement Index (MCPPI). Conditions for medical students' learning in hospital and community placements. *Advances in Health Sciences Education* 17, 703–716.
- Embo, M.P., Driessen, E.W., Valcke, M., van der Vleuten, C.P., 2010. Assessment and feedback to facilitate self-directed learning in clinical practice of Midwifery students. *Medical Teacher* 32, e263–e269.
- Embo, M.P., Driessen, E.W., Valcke, M., van der Vleuten, C.P., 2014. A framework to facilitate self-directed learning and supervision in clinical practice: a qualitative study of supervisors' perceptions. *Nurse Education in Practice* 14, 441–446.
- Embo, M.P., Driessen, E.W., Valcke, M., van der Vleuten, C.P., 2015. Integrating learning, assessment and supervision in a competence framework for clinical workplace education. *Nurse Education Today* 35, 341–346.
- Embo, M., Valcke, M., 2016. Workplace learning in midwifery education in Flanders (Belgium). *Midwifery* 33, 24–27.
- Epstein, R.M., 2007. Assessment in medical education. *New England Journal of Medicine* 356, 387–396.
- Epstein, R.M., Hundert, E.M., 2002. Defining and assessing professional competence. *JAMA* 287, 226–235.
- Fastré, G.M.J., van der Klink, M.R., van Merriënboer, J.J.G., 2010. The effects of performance-based assessment criteria on student performance and self-assessment skills. *Advances in Health Sciences Education* 15, 517–532.
- Fastré, G.M.J., van der Klink, M.R., Amsing-Smit, P., van Merriënboer, J.J.G., 2014. Assessment criteria for competency-based education: a study in nursing education. *Instr Sci* 42, 971–994.
- Federal Government of Health, 2015. Professional Profile of the Belgian Midwife [Beroepsprofiel van de vroedvrouw]. Available from: (<http://www.vroedvrouwen.be/sites/default/files/Beroepsprofiel%2012%2005-1.pdf>).
- Flemish Education Council, 2014. Educational Profile of the Flemish Midwife [Opleidingsprofiel van de vroedvrouw]. Not published.
- Foth, T., Efstatiou, N., Vanderspank-Wright, et al., 2016. The use of Delphi and Nominal Group Technique in nursing education: a review. *International Journal of Nursing Studies* 60, 112–120.
- Frank, J.R., Snell, L.S., ten Cate, O., et al., 2010. Competency-based medical education: theory to practice. *Medical Teacher* 32, 638–645.
- Frank J.R., 2011. Beginning with the end in mind: defining graduate abilities. In: Sherbin J, Frank J.R., (Eds.) *Educational Design. A CanMEDS Guide for the Health Professions*. Ottawa Ontario K1S 5N8 Canada. pp 3–4.
- Frank J.R., Snell L., Sherbino J., 2015. *CanMEDS 2015. Physician Competency Framework*. Royal College of Physicians and Surgeons of Canada, Ottawa, Canada. pp 5–7.
- Fullerton, J.T., Johnson, P., Lobe, E., et al., 2016. A Rapid assessment tool for affirming good practice in midwifery education programming. *Midwifery* 34, 36–41.
- Govaerts, M., van der Vleuten, C.P.M., 2013. Validity in work-based assessment: expanding our horizons. *Medical Education* 47, 1164–1174.
- Gulikers, J.T.M., Baartman, L.K.J., Biemans, H.J.A., 2010. Facilitating evaluations of innovative competency-based assessments: creating understanding and involving multiple stakeholders. *Evaluation and Program Planning* 33, 120–127.
- Harrison, C.J., Könings, K.D., Dannefer, E.F., Schuwirth, L.W.T., Wass, V., van der

- Vleuten, C.P.M., 2016. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. *Perspectives on Medical Education* 5, 276–284.
- Haug C., 2015. What is consensus and how is it achieved in meetings? Four practices of consensus decision-making. In: Allen J.A., Lehmann-Willenbrock N., Rogelberg S. G. (Eds.) *The Cambridge Handbook of Meeting Science?* Cambridge University Press, UK.
- Holmboe E.S., Snell L., 2011. Principles of competency-based education: better preparation of learners for practice. In: Sherbin J., Frank J.R., (Eds.) *Educational Design. A CanMEDS Guide for the Health Professions.* Ottawa Ontario K1S 5N8 Canada. pp 9–15.
- Idrissi M., Hnida M., Bennani S., 2016. Competency-Based Assessment: From Conceptual Model. In: Cano E., Ion G. (Eds.) *Innovative Practices for Higher Education Assessment and Measurement.* IGI Global, Hershey, PA., pp 57.
- Kalet, A., Guerrasio, J., Chou, C.L., 2016. Twelve tips for developing and maintaining a remediation program in medical education. *Medical Teacher* 38, 787–792.
- Lane, I.F., 2010. Professional competencies in health sciences education: from multiple intelligences to the clinic floor. *Advances in Health Sciences Education* 15, 129–146.
- Leung, W., 2002. Competency based medical training: review. *BMJ* 325, 693–696.
- Lurie, S.J., 2012. History and practice of competency-based assessment. *Medical Education* 46, 49–57.
- Mann, K.V., 2011. Theoretical perspectives in medical education: past experience and future possibilities. *Medical Education* 45, 60–68.
- Michels, N.R.M., Driessen, E.W., Denekens, J., Van Gaal, L.F., Bossaert, L.L., De Winter, B.Y., 2012. A Delphi study to construct a CanMEDS competency-based inventory applicable for workplace assessment. *BMC Medical Education* 12, 86.
- Ossenberg, C., Henderson, A., 2015. Determining attainment of nursing standards: the use of behavioural cues to enhance clarity and transparency in student clinical assessment. *Nurse Education Today* 35, 12–15.
- Richardson D., Flynn L., 2011. The road map: learning objectives. In: Sherbin J., Frank J.R., (Eds.) *Educational Design. A CanMEDS Guide for the Health Professions.* Ottawa Ontario K1S 5N8 Canada. pp. 33–35.
- Reeves, S., Fletcher, S., Barr, H., et al., 2016. A BEME systematic review of the effects of interprofessional education: beme Guide No. 39. *Medical Teacher* 38, 656–668.
- Schüttpelz-Brauns, K., Narciss, E., Schneyinck, C., et al., 2016. Twelve tips for successfully implementing logbooks in clinical training. *Med Teach* 38, 564–569.
- Schuwirth, L., van der Vleuten, C.P.M., 2011. Programmatic assessment: from assessment of learning to assessment for learning. *Medical Teacher* 33, 478–485.
- Shumway, J.M., Harden, R.M., 2003. AMEE Guide No. 25: the assessment of learning outcomes for the competent and reflective physician. *Medical Teacher* 25, 569–584.
- Stoof, A., Martens, R., Van Merriënboer, R., Bastiaens, T., 2002. The boundary approach of competence: a constructivist aid for understanding and using the concept of Competence. *Human Resource Development Review* 1, 345–365.
- Trevelyan, E.G.R., Robinson, N., 2015. Delphi methodology in health research: how to do it? *European Journal of Integrative Medicine* 754, 428.
- Van der Vleuten, C.P.M., Schuwirth, L.W.T., Driessen, E.W., et al., 2012. A model for programmatic assessment fit for purpose. *Medical Teacher* 34, 205–214.
- Van der Vleuten, C.P.M., Heeneman, S., 2016. On the issue of costs in programmatic assessment. *Medical Teacher* 5, 303–307.
- Watling, C., 2016. The uneasy alliance of assessment and feedback. *Perspectives on Medical Education* 5, 262–264.
- Yanhua, C., Watson, R., 2011. A review of clinical competence assessment in nursing. *Nurse Education Today* 31, 832–836.