

# Breathe

## Assessment methods in respiratory medicine training in Europe: current status and future needs

There is a perceived need for harmonisation of training standards in medicine across Europe. Assessment methodology is a fundamental part of this harmonisation. Assessment may be: summative, measuring competency at the end of training, usually in a pass/fail mode; or formative, assessing strengths and weaknesses during training with a view to optimising performance. A survey of experienced respiratory medicine trainers in Europe was undertaken to discover the formative and summative assessment methods being used, and to explore the use of workplace-based assessments (WBAs). Structured interviews were sought with experienced trainers in adult and paediatric respiratory medicine. 35 trainers from 22 (71%) out of 31 countries were interviewed.

The number of types of summative assessments required at the end of training varied from 0 to 4 including clinical (25%), written (44%) and oral (47%) examinations. Four respondents required a research thesis as a proof of clinical competence.

WBA was not commonly used. Only 14% of respondents reported using a formal case-based discussion, 20% used a form of multisource feedback and 25% described some form of formal assessment of procedural skills. However, 77% of all respondents expressed a wish to have access to case-based discussion and 72% wanted to have a facility for multisource feedback. The majority also wanted training in the use of these tools. Almost half of the respondents had received no formal training in educational supervision and 80% of all respondents expressed a wish to receive such training.

The findings suggest that there is no adequate process of ensuring uniform standards for specialist accreditation in Europe, and demonstrate a need and desire among trainers for more WBA tools and training in their use to be made available.

**Cite as:** Primhak R. Assessment methods in respiratory medicine training in Europe: current status and future needs. *Breathe* 2020; 16: 190314.

For some years the European Respiratory Society (ERS) has made efforts to set standards for training in respiratory medicine across Europe. The HERMES programme (Harmonising Education in Respiratory Medicine for European Specialists) was established for this purpose, with the development of syllabi and curricula in both adult and paediatric respiratory

medicine [1–4]. The HERMES theory examinations in adult and then paediatric respiratory medicine followed shortly afterwards, aiming to assess the knowledge required to practice as a specialist, set against these curricula. However, there are no standardised criteria for specialist accreditation in Europe, as the requirements are set by the individual

 @ERSpublications

**Assessment of respiratory medicine trainees varies considerably across Europe. This will make harmonisation of specialist accreditation problematic. Few countries use formal workplace-based assessment but trainers would like access to such methods.** <http://bit.ly/2uALzMj>



CrossMark



© ERS 2020

regulatory authorities within each country and vary from country to country. While occasional countries have adopted the HERMES examination as one of the assessment criteria, this is by no means widespread.

A knowledge base, as assessed by the HERMES examination, is a necessary part of the competence required of a specialist but does not, in itself, demonstrate competency. Miller's classic pyramidal model for assessing clinical competence suggests four levels: "knows", "knows how", "shows how" and "does" [5]. A theory examination such as HERMES can only get as far as the second level, "knows how", but gives no insight into higher levels of clinical competence; neither does it assess skills, attitudes or behaviour. For these more real-life assessments, there are several tools available, many of which are forms of workplace-based assessment (WBA).

In 2010, the ERS set out a list of possible tools that could be used to assess competency of trainees [4], including assessment of clinic letters, case-based discussion (CbD), direct observation of procedural skills (DOPS) and multisource feedback (MSF). This study set out to assess whether such tools were in use across Europe and, where they were not available, whether trainers perceived a need or desire to have access to such tools.

## Methodology

### Population and methods

Because of the potential for confusion about some of these tools in written questionnaires, it was felt that a verbal interview would be a better way to obtain accurate information about current practice and future wishes. ERS national delegates from each country within the European Union or the Schengen area were asked to provide contact details for one adult specialist and one paediatric specialist in respiratory medicine who were involved in supervision of trainees for the purposes of the survey. A single interviewer then contacted each subject by e-mail to ask if they were willing to participate and arrange a suitable time for interview. Where delegates did not respond to two or three requests, other approaches were made to try and find respondents from these countries. Interviews were conducted by telephone, Skype (Microsoft Corp., Redmond, WA, USA) or Facetime (Apple Inc., Cupertino, CA, USA), or face-to-face when the opportunity arose.

Each respondent was asked the same questions in a semistructured interview (supplementary material) but was allowed to expand on the answers as they wished. Further clarifying questions were used as necessary. The questions first set out to establish the methods of summative assessment of competence at the end of training. An open question, "How do you decide whether a trainee is competent at the end of training?", was followed

by specific questions about different methods of assessment. The questions then focused on WBAs used during training. The specific methods included in the questionnaire were CbD [6], procedural assessment such as DOPS [7], MSF [8], patient feedback [9] and assessment of clinic letters [10].

In each case, the question was asked about a structured method of assessment with a recorded outcome. Whether or not the respondent used such methods, they were asked if they or their colleagues were likely to welcome such a tool being available from the ERS. Furthermore, in the case of CbD and MSF, they were asked if they or their colleagues would welcome training in how to use the tools. The answers to these two questions were coded as "yes", "no" or "maybe" to allow some degree of uncertainty in the response. They were then asked if there were other methods of assessment that they used or whether there were other tools that they would be interested in the ERS developing for such assessments. Finally, the respondent was asked whether they had previously received any formal training in educational supervision and whether they would be interested in further training from the ERS.

## Results

### Population

We included a possible 31 countries in the survey; in 13 cases, we were sent the contact details of possible respondents, and in a further 11 countries, the interviewer identified possible respondents from personal contacts. Most contacts were happy to be interviewed; in one country, no responses could be obtained despite two e-mail approaches to the identified contacts. This resulted in data from 22 (73%) countries and 35 (58%) respondents, of which 19 were adult specialists and 16 were paediatric (table 1). Several of the countries that failed to respond have relatively small populations and, in some cases, do not undertake specialist training. The population of the countries represented in this survey was 461.8 million (88% of the total target population).

### Summative assessments of competency

The type of assessment, if any, at the end of training varied widely between countries, and to some extent within countries, between adult and paediatric training, and is shown in table 2.

16 (44%) of the respondents had a written examination, although only two used the HERMES examination as the assessment. Nine (25%) set a clinical examination, 17 (47%) had an oral examination, 25 (69%) required the trainee to maintain a log book of clinical experience and 28 (78%) required formal supervisor reports. The

**Table 1** *The 31 countries targeted and responses obtained*

Country	Population millions <sup>#</sup>	Adult	Paediatric
Austria	8.60	•	
Belgium	11.20	•	
Bulgaria	7.20		
Croatia	4.20	•	•
Cyprus	0.80		
Czech Republic	10.50	•	•
Denmark	5.70	•	•
Estonia	1.30	•	•
Finland	5.50		
France	66.40	•	•
Germany	81.10		•
Greece	10.90	•	•
Hungary	9.90	•	
Ireland	4.60	•	•
Iceland	0.35		
Italy	60.80	•	•
Latvia	2.00		
Lithuania	2.90	•	•
Luxembourg	0.60		
Malta	0.40	•	
Netherlands	16.90	•	•
Norway	5.30	•	•
Poland	38.00		
Portugal	10.30	•	
Romania	19.80	•	•
Slovakia	5.40		
Slovenia	2.10		
Spain	46.40		•
Sweden	9.70		•
Switzerland	8.50	•	•
UK	66.40	•	•
<b>Total</b>	<b>523.80</b>	<b>19</b>	<b>17</b>

<sup>#</sup>: derived from European Union and national sources [11–13].

details of what methods were used in the “clinical examination” and “oral examination” were not explored. In three cases, no formal assessment was performed; a trainee had merely to complete the required time in training without being dismissed to achieve specialist accreditation. At the other end of the scale, eight respondents used at least four

of the five listed assessment types. To complicate matters, paediatric respiratory medicine was not recognised as an accreditable speciality in half of the 24 countries where data were available.

## WBA tools

The respondents’ reported current use of the specific WBA tools included in the survey, and their estimation of the usefulness and desirability of such tools in their own practice, are shown in figure 1. It can be seen that the most common form of WBA was a clinic letter appraisal and in most cases, this was not a formal structured appraisal but a supervisor review and sign-off of the letter. There was relatively little enthusiasm for access to a formal tool for letter appraisal; many respondents described a template letter that all doctors used, which constrained the content and would make an appraisal tool less useful. In contrast, there was little reported use of a private, one-to-one CbD with a method of recording the outcomes; this was restricted to those countries that have introduced a mandatory system for such assessments, namely Denmark, Ireland, Norway and the UK. Of the 29 (78.4%) respondents who do not currently use a formal CbD, 25 (86.2%) expressed a definite desire to have access to such a tool. Of these 25 who answered positively, 23 (92%) felt that training in the method would also be desirable for them and their colleagues.

MSF was only being carried out in countries where there was a requirement for it and where there were resources to undertake it, namely Denmark, Malta, the Netherlands and the UK. Of the 30 respondents who do not currently use MSF, 26 (86.7%) thought that it would be a useful resource to have available and the other four answered “maybe”, with none answering “no”. When asked whether training in discussing the feedback with the trainee was desirable, 25 (96%) out of 26 answered positively.

Patient feedback on the trainee’s performance was used by only three respondents. Some institutions had consumer surveys in clinics about the whole patient experience, but this did not focus on the medical staff attitudes and the results were rarely linkable to a particular doctor. While just over 54% felt it would be useful, there were more concerns about confidentiality and practicality.

An assessment of procedural skills was reported to be more commonly used, being performed by 24% of respondents. A number of respondents felt that they did not need a tool for formal assessment of skills such as bronchoscopy, as they knew by observation whether the trainee was competent in the procedure and they were always supervised until they were felt to be competent. It was not always clear from the interview whether the procedural assessment, when performed, was standardised against anchor statements or merely a record of a satisfactory procedure.

**Table 2** Summative assessments used by country in adult (n=19) and paediatric (paed) (n=17) respiratory training

	Summative assessment type									
	Written		Clinical		Oral		Log book		Supervisor reports	
	Adult	Paed	Adult	Paed	Adult	Paed	Adult	Paed	Adult	Paed
Austria	●		○		●		●		●	
Belgium	●		●		○		●		●	
Croatia	○	○	○	●	●	●	○	○	○	●
Czech Republic	●	○	●	●	●	●	●	●	○	●
Denmark	○		○		○		●		●	
Estonia	●	●	○	○	○	●	●	●	●	●
France	○	●	○	○	●	●	●	●	●	●
Germany		○		○		●		●		○
Greece	●	○	○	○	●	○	○	○	●	○
Hungary	○		●		●		○		●	
Ireland	○	●	○	○	○	○	●	●	●	●
Italy	○	○	○	○	●	●	●	●	●	●
Lithuania	●	●	○	○	●	○	●	○	○	○
Malta	●		○		○		●		●	
Netherlands	○	○	○	○	○	○	●	●	●	●
Norway	○	○	○	○	○	○	●	●	●	●
Portugal	●		●		●		●		●	
Romania	●	●	●	●	●	○	○	●	●	●
Spain	○	○	○	○	○	○	○	○	○	○
Sweden		○		○		○		●		●
Switzerland		○		○		●		○		●
UK	●	○	○	●	○	○	○	●	●	●
Assessment used	10 (53%)	6 (35%)	5 (26%)	4 (24%)	10 (53%)	7 (41%)	13 (68%)	12 (71%)	15 (79%)	13 (76%)
Overall (n=36)	16 (44%)		9 (25%)		17 (47%)		25 (69%)		28 (79%)	

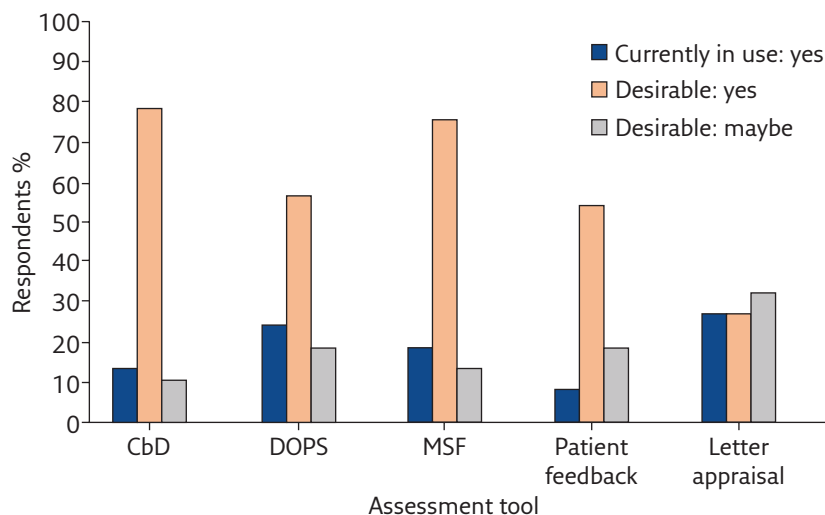
No other assessment tools were identified by any respondent as being in use apart from a research thesis, defended by a *viva voce* examination, which five respondents mentioned as part of the summative assessment. No other suggested assessment tools were identified.

When the respondents were asked what training they had received in educational supervision, 20 (57.1%) out of 35 had received some training; however, 28 (80%) respondents felt that they would like further training in this area if it was available.

## Discussion

Assessments are an important part of training for several reasons. The most important purpose of summative assessment is patient safety, since

it is unjustifiable to allow less than competent trainees to practise independently as specialists. Another benefit of assessment is to drive learning, since trainees will put more effort into acquiring a skill, knowledge or behaviour if they know it is going to be assessed in a high-stakes examination [14]. Assessments should ideally be valid, objective, reliable and feasible [7]. Validity implies that it tests whether the trainee is competent to practice as a specialist, with the appropriate knowledge, skills, attitudes and behaviours required for this role, measuring level 4 of the Miller pyramid (“does”) [5]. Most of the traditional methods that were used by respondents in this study do not measure competence at this level, and few fulfil the other requirements of objectivity and reliability. The idea of harmonising standards across Europe has proved problematic in practice and at present,



**Figure 1** Current usage and desirability of workplace-based assessment tools.

there is little convincing evidence to confirm that trainees in adult or paediatric respiratory medicine are emerging from their training at a similar level of competence regardless of the country of training.

However, it could be argued that if a trainee has successfully completed medical school and basic medical training, and has entered specialist training in adult or paediatric respiratory medicine, they should have the innate intelligence and ability to become competent in the speciality. If a trainee fails a summative assessment at the end of training, this could be because the assessment is invalid or it could be interpreted as a failure of the training programme rather than the trainee. If patient safety is the issue, it might be argued that ensuring that every trainee becomes as good as they can be is at least as important as a process that excludes one or two trainees who have failed an arbitrary standard. For this reason, some countries have introduced formative WBAs into the training programme, to identify problems and attempt to remedy them before completion of training.

WBAs do measure level 4 of the Miller pyramid but are subjective and single assessments are unreliable. Thus, it is necessary to undertake repeated and different WBAs, performed by different supervisors, to achieve a reasonable reliability. While they can be used as evidence to inform the summative assessment of competence at the end of training, they have a far more important role in formative assessment, helping the trainee and supervisor to identify developmental needs and remedy them before the completion of training.

The present study has identified some major inconsistencies in the types of assessments used, both for summative and formative assessment in respiratory medicine training. Similar variability across Europe was recently reported for anaesthetic training [15]. There are a number of deficiencies in our study: it is incomplete, sampling only just over half of the target respondents, although it has information from 73% of the target countries, representing 88% of the population. This lack of completeness does not

invalidate the conclusions, since even within this population, there are considerable inconsistencies in the assessment methods used and few are using formal WBAs. The study was conducted by an interviewer rather than a written questionnaire; this has the advantage of the interviewer being able to clarify questions that were poorly understood and to explore incomplete answers. It may, however, have introduced a bias into the results, since respondents may be more likely to express interest in methods that they do not actually wish to use. Reassuringly, the respondents showed no hesitation in refusing tools, such as a clinic letter assessment tool, which they felt would not be useful in their practice; nevertheless, it is possible that the needs assessment has overstated the likely use of such tools if they were introduced. Some WBA tools were not included in the question set: for example, miniCEX (clinical evaluation exercise) is a widely used method of assessing history-taking and examination skills [16], but it was felt that such skills should be well-established by the time a trainee reaches subspeciality training, and clinical reasoning could be better assessed with Cbd.

A key element in WBA is the feedback and supervision offered by the trainer. Behavioural change in the recipient is dependent on the quality of feedback and mentoring [17–19]. It is therefore of concern that only just over half of the experienced supervisors in this survey had received any formal training in educational supervision and >80% expressed an enthusiasm for further training in this area.

It should be noted that recent ERS training courses [20] have included validated assessment methods for all levels of Miller's pyramid but it is clear that more validated WBAs will be needed both to obtain a global picture of trainee competence and to ensure that trainees achieve their maximum potential at the end of training.

Based on the results of this study, there does seem to be a need and a desire to have access to WBAs. The most desirable assessments are a formalised Cbd, a form of procedural assessment and a tool for MSF.

Alongside the introduction of such tools there must be training available in how to use the tools and how to facilitate the feedback when using them.

## Conclusions

This study has demonstrated a wide variation in methods of summative assessment and

accreditation in respiratory medicine training in Europe, and a paucity of countries using formal WBA either as formative or summative assessments. The findings do not suggest that there is an adequate process of ensuring uniform standards for specialist accreditation in Europe and demonstrate a need and desire for more WBA tools and training to be made available.

## Affiliations

### Robert Primhak

Education Council, European Respiratory Society, Lausanne, Switzerland.

## Acknowledgements

I would like to thank Nathalie Tabin (ERS, Lausanne, Switzerland) for her support in identifying possible respondents and my colleagues from across Europe who cheerfully gave their time for the interviews.

## Supplementary material

This article has supplementary material available from [breath.ersjournals.com](http://breath.ersjournals.com).

## Conflict of interest

R. Primhak has nothing to disclose.

## References

- Loddenkemper R, Severin T, Eisele J-L, *et al.* HERMES: a European core syllabus in respiratory medicine. *Breathe* 2006; 3: 59–69.
- Loddenkemper R, Haslam PL, Severin T, *et al.* European curriculum recommendations for training in adult respiratory medicine. *Breathe* 2008; 5: 80–93.
- Gappa M, Noel J-L, Severin T, *et al.* Paediatric HERMES: a European syllabus in paediatric respiratory medicine. *Breathe* 2009; 5: 237–247.
- Gappa M, Noel J-L, Severin T, *et al.* Paediatric HERMES: European curriculum recommendations for training in paediatric respiratory medicine. *Breathe* 2010; 7: 72–79.
- Miller G. The assessment of clinical skills/competence/performance. *Acad Med* 1990; 65: S63–S57.
- Primhak R, Gibson N. Workplace-based assessment: how to use case-based discussion (CbD) as a formative assessment. *Breathe* 2019; 15: 163–166.
- Davis M, McKimm J, Forrest K. How to assess doctors and health professionals. Oxford, John Wiley, 2013.
- Archer J, Norcini J, Southgate L, *et al.* Mini-PAT (Peer Assessment Tool): a valid component of a national assessment programme in the UK? *Adv Health Sci Educ Theory Pract* 2008; 13: 181–192.
- Al-Jabr H, Twigg M, Scott S, *et al.* Patient feedback questionnaires to enhance consultation skills of healthcare professionals: a systematic review. *Patient Educ Couns* 2018; 101: 1538–1548.
- Crossley JGM, Howe A, Newble D, *et al.* Sheffield Assessment Instrument for Letters (SAIL): performance assessment using outpatient letters. *Med Educ* 2001; 35: 1115–1124.
- Eurostat. Population on 1 January. <https://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tps00001>. Date last updated: June 6, 2019.
- Statistics Norway. Key figures for the population 2019. Date last updated: May 20, 2019. <https://www.ssb.no/en/befolkning/nokkeltall>.
- Swiss federal Statistical Office. Population Growth 2018: provisional figures. [www.bfs.admin.ch/bfs/en/home.assetdetail.7946001.html](http://www.bfs.admin.ch/bfs/en/home.assetdetail.7946001.html). Date last updated: April 9, 2019.
- Wormald BW, Schoeman S, Somasunderam A, *et al.* Assessment drives learning: an unavoidable truth? *Anat Sci Educ* 2009; 2: 199–204.
- Jonker G, Manders LA, Marty AP, *et al.* Variations in assessment and certification in postgraduate anaesthesia training: a European survey. *Br J Anaesth* 2017; 119: 1009–1014.
- Lorwald AC, Lahner F-M, Nouns ZM, *et al.* The educational impact of Mini-Clinical Evaluation Exercise (Mini-CEX) and Direct Observation of Procedural Skills (DOPS) and its association with implementation: a systematic review and meta-analysis. *PLoS One* 2018; 13: e0198009.
- Overeem K, Wollersheim HC, Arah OA, *et al.* Factors predicting doctors' reporting of performance change in response to multisource feedback. *BMC Med Educ* 2012; 12: 52–58.
- Ferguson J, Wakeling J, Bowie P. Factors influencing the effectiveness of multisource feedback in improving the professional practice of medical doctors: a systematic review. *Med Educ* 2014; 14: 76.
- Mehta F, Brown J, Shaw N. Do trainees value feedback in case-based discussion assessments? *Med Teach* 2013; 35: e1166–e1172.
- Farr A, Clementsen P, Herth F, *et al.* Endobronchial ultrasound: launch of an ERS structured training programme. *Breathe* 2016; 12: 217–220.