university of groningen

University Medical Center Groningen

University of Groningen

Progress testing in intensive care medicine training

Van Mook, Walther N. K. A.; Arbous, Sesmu M.; Delwig, Hans; Van Hemel-Rintjap, Tina J. D.; Tepaske, Robert; Tulleken, Jaap. E.; Van der Vleuten, Cees P. M.

Published in:
Minerva anestesiologica

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Van Mook, W. N. K. A., Arbous, S. M., Delwig, H., Van Hemel-Rintjap, T. J. D., Tepaske, R., Tulleken, J. E., & Van der Vleuten, C. P. M. (2016). Progress testing in intensive care medicine training: useful and feasible?! *Minerva anestesiologica*, *82*(6), 711-719. http://www.minervamedica.it/en/journals/minerva-anestesiologica/article.php?cod=R02Y2016N06A0711

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 29-10-2022

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

© 2015 EDIZIONI MINERVA MEDICA
Online version at http://www.minervamedica.it
Minerva Anestesiologica 2016 June;82(6):711-9

EXPERTS' OPINION

Progress testing in intensive care medicine training: useful and feasible?!

Walther N.K.A. VAN MOOK [1-3] *, Sesmu M. ARBOUS [3, 4], Hans DELWIG [3, 5], Tina J.D. VAN HEMEL-RINTJAP [3, 4], Robert TEPASKE [3, 6], Jaap. E. TULLEKEN [3, 5], Cees P.M. VAN DER VLEUTEN [2]

[1]Department of Intensive Care Medicine, Maastricht University Medical Centre, Maastricht, The Netherlands; [2]Department of Medical Education Development and Research, Faculty of Health, Medicine, and Life Sciences, Maastricht University, Maastricht, The Netherlands; [3]Education and Training Working Group of the Joint Intensivists Committee, The Netherlands; [4]Department of Intensive Care Medicine, Leiden University Medical Centre, Leiden, The Netherlands; [5]Department of Intensive Care Medicine, University Medical Centre, Groningen, The Netherlands; [6]Department of Intensive Care Medicine, Academic Medical Centre, Amsterdam, The Netherlands

*Corresponding author: Walther N.K.A. van Mook, Department of Educational Development and Research, Faculty of Health, Medicine, and Life Sciences, Maastricht University, P. Debyelaan 25, 6202 AZ Maastricht, The Netherlands.
E-mail: w.van.mook@mumc.nl

ABSTRACT

So far the in-training assessment of knowledge is perhaps underrepresented in postgraduate assessment frameworks in intensive care medicine (ICM). In most contemporary training programs a predominant emphasis is placed on workplace based learning and workplace based assessment. This article provides a concise general background on the nature and use of progress testing, and touches upon potential strengths, and constraints regarding its potential implementation and use in the postgraduate ICM training programs.

(*Cite this article as*: van Mook WNKA, Arbous SM, Delwig H, van Hemel-Rintjap TJD, Tepaske R, Tulleken JE, *et al.* Progress testing in intensive care medicine training: useful and feasible?! Minerva Anestesiol 2016;82:711-9)

Key words: Intensive care units - Educational measurement - Knowledge.

Semi-recently there has been a progression towards defining curricula in terms of educational outcomes through competency-based training.[1] Since significant differences exist between specialties, the specialty-specific tailoring of the general competency framework is relegated to specialty-specific committees. In the UK, a comprehensive competency-based training program for intensive care medicine (ICM) was developed in 2001.[2] In 2006, the CoBaTrICE program published competencies for an international training program in ICM.[3] The latter were endorsed by the European Union of Medical Specialists (UEMS) in 2008. A national training and assessment program for ICM training can subsequently be constructed based on this competency framework regarding knowledge, skills and behavior. Although UEMS acknowledges that some of the knowledge and skills of ICM are universal to many medical and surgical specialties (and are herein trained and assessed), in-training- assessments in ICM should nevertheless include formative and summative assessments, respectively to encourage learning and provide evidence of the achievement of competence.

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

End-of-training, summative assessment in the Netherlands is currently performed using the two-part European Diploma in Intensive Care format (EDIC) consisting of a theoretical, knowledge part I, and an oral/clinical part II. Assessment-in-training is mainly performed during workplace-based, formative assessments.[4] Details regarding each separately defined knowledge, skills and attitude item is found on the CoBaTrICE website. Since, in an international perspective, different ICUs offer varied opportunities for assessment, CoBaTrICE provides general assessment guidance rather than a proscriptive pathway, allowing details of the assessment process to be established at a national level. Knowledge testing during ICM training in the Netherlands is so far neither a mandatory nor integrated part of all local training programs. Nevertheless, the UEMS advocates in-training evaluations of knowledge with multiple choice questions (MCQs) of *viva voce*.[5] The formal assessment of knowledge in ICM in the Netherlands seems underrepresented, with a predominant emphasis on workplace based learning and assessment.[4] Recently it was re-emphasized that "*Expert problem solving cannot take place without a well organized knowledge database and requires expert knowledge*".[6, 7] Participation by engaging in deliberate practice, learning by doing, should thus be complemented by some kind of knowledge testing.[8] A frequently used method for testing knowledge aspects during medical training is the so called progress testing. Progress testing involves the principle of longitudinally and repeatedly testing of trainees' knowledge.[9] Progress testing is based on the application of subsequent equivalent, yet different tests, after which the aggregation of the derived data is used as an indicator of the growth of the (functional) medical knowledge of the candidate. Whereas each individual progress test provides feedback to the learner regarding areas of strengths and potential improvements and thus contributes to assessment for learning, the aggregated results over a predefined timeframe are also used for summative purposes, on which pass fail decisions can be based. Extensive discussion of the more generic aspects of assessment in general and programmatic assessment specifically is beyond the scope of this article, and is summarized elsewhere.[8, 10, 11] Although primarily developed and used for undergraduate medical education,[9] one might wonder whether progress testing can and should also be applied to postgraduate training in general, and intensive medicine training programs in particular as part of a framework of programmatic formative and summative assessment.[10, 11] The next sections will consecutively touch upon the potential strengths, constraints and remaining questions regarding the scientific background, practical use and implementation of systemic progress testing frameworks in general,[9, 12] and in postgraduate medicine specifically.

What is progress testing?

When progress testing is applied to undergraduate medical education, the questions aim to represent relevant knowledge at the graduate level (the level students' should master at graduation). Each test provides a comprehensive sample of questions covering broad domains of relevant medical knowledge. The framework for the assessment of competence and performance as proposed by Miller [13] is often used to illustrate the relative position and use of commonly assessment instruments in medical education, and is visualized in Figure 1. Questions used in progress testing preferably aim at the functional level of students' (so called applied) knowledge, thus level 2 (not the base) of Miller's pyramid.[13] For comparison in a longitudinal fashion, multiple equivalent, but different assessments are used, often resulting in a number of sets per year throughout the entire training program. The production of each separate test is centrally coordinated, reviewed, administered and analyzed. Since each assessment aims at testing the knowledge that should have been incorporated at the end of medical school, it is relatively independent on the local training program and curricula, and can therefore be used in "*a multi-centre collaborative production and administration framework thus reducing, costs, increasing efficiency,*

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

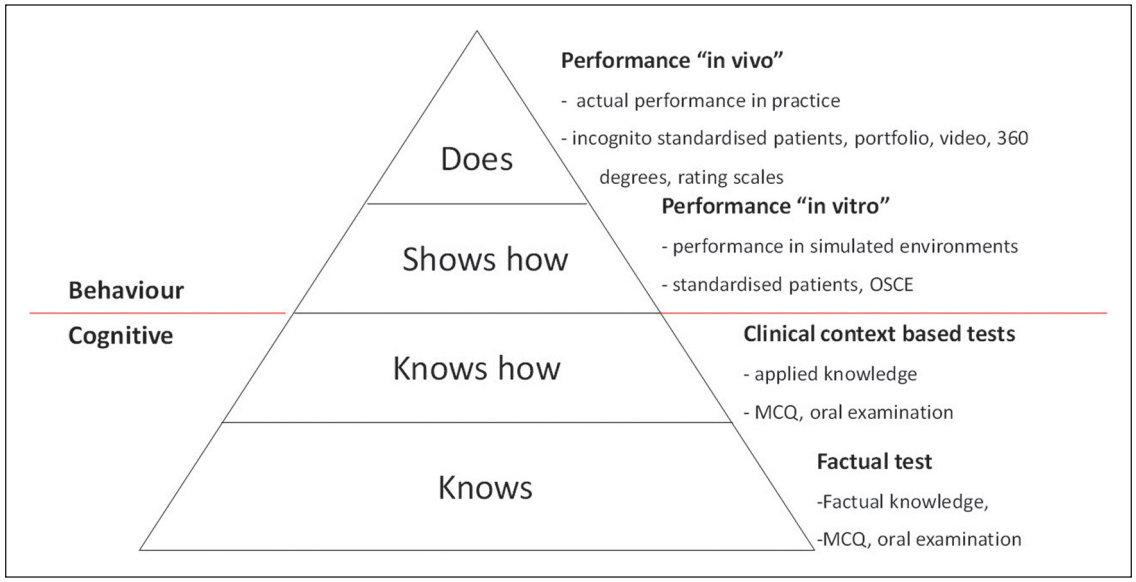

Figure 1.—The framework for assessing competence as proposed by Miller: Miller's pyramid.[13]

and allowing for constant benchmarking".[9] In medical school each progress test is therefore completed by all students in all year classes. Although knowledge gradually increases when applying progress testing to all years of training, this is obviously also accompanied with occasional peaks and troughs,[9] these being damped when progress testing is summative in nature.[14] Since all years of students are assessed, the most junior students will have difficulty answering most questions, leading to guessing. To discourage this phenomenon a correct-minus-false formula scoring system as well as a question mark option are offered. A drawback of offering a question mark option is that it decreases psychometric performance.[9] After completion of each test, the individual assessment results are aggregated numerically and graphically to provide an overview about the overall growth of the individual student's knowledge as compared to this student's peers, and forms the basis for a decision on progressing to the next year in medical school.[15] Additionally, both the trainee and the training program director thus receive feedback on performance, respectively on an individual and institutional curriculum level, after each separate test.

What can one expect from progress testing?

It is clear that a progress test use is a unique and demonstrable measurement of the growth of knowledge as well as the effectiveness of knowledge gathering,[12] measuring both process and outcome. In addition, progress testing has a positive influence on learning, causing students to study more continuously while building a better foundation of knowledge,[16] that retains over the years.[9] It can easily be envisaged that in the view of the comprehensiveness of the test, strategic preparation for progress testing is difficult, and continuous learning remains as the most optimal preparation.[9] Test-driven strategies are thereby discouraged, or rendered useless.

Furthermore, it is generally assumed that less examination stress is experienced when using progress testing. In progress testing a single bad result does not directly have consequences in the perspective of a previous series of good results; this has indeed been reported to provide little or only moderate stress in the majority of participants.[17] Resits for a one-off unsatisfactory result are therefore also unnecessary,[9] another practical advantage. However, a series of poor test results can compara-

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

bly not be compensated by a single positive result, and subsequent resits, remediation attempts, and alternative testing approaches are all stressful events for students with such experiences.[9]

Looking at progress testing from a medical education perspective, the longitudinal aggregation of results adds to the reliability of the decisions based on progress testing, and the sampling properties of progress testing have proven more important than the test's structure. There is a trade-off between frequency and number of questions (items) in the test. Examples of this phenomenon are *e.g.* published by Rickets for a progress test framework example in the United Kingdom (*e.g.* 2 tests of 200 items per year yields more reliable results than 5 times year 100 items).[18] More frequent sampling is not simply always better, increases in quantity does not automatically contribute to quality. A comparable analysis was performed by Wrigley *et al.* for the Maastricht progress testing. In the more advanced years of medical school training a reliability of 0.9 was reached using 4 tests per year of 200 questions each.[12]

However, determining what constitutes the content and nature of the items in the progress test is even more difficult and important than setting the frequency of tests and the number of questions per test. What knowledge is considered relevant, and necessary at graduation level, and what knowledge should be regarded as *e.g.* too focused, in-depth knowledge for example to be gathered during additional postgraduate training remains challenging however.[9] The same is true regarding insurance of the equivalence of the individual assessments within a time frame.[9, 14] The recent guide on progress testing by the Association for Medical Education in Europe (AMEE) summarizes such issues stating *"Multiple choice question tests in medicine… show considerable sensitivity for "construct-irrelevant variance"*,[12, 19] and examples include items judged to reflect non-core medical knowledge, use of flawed test items and imprecise terminology (the former resulting score lowering of up to 15%!),[20, 21] and variation of test difficulty.[12] Thus attention should be paid to ensure consistency and uniformity regarding progress test construction, content, administration, testing conditions and scoring procedures.[12] For this purpose the AMEE guide describes an empirically based systematic framework for progress test practices.[12] So far considerable variation in both content and application of progress testing is noticed.[12] However, a common denominator is formed by the necessity of establishing consensus on a 'blue print' representing the end-of-training objectives, thereby providing clarity on the topics and subjects assessed. However, in depth discussion of these recommendations is beyond the scope of this paper.[12]

Finally it should be noted that the context in which progress testing is used may at least partly determine its use and outcome.[9] The progress test is always part of a systematic, programmatic approach to assessment in which it is only one out of different methods used. Progress testing is complimentary to, but not a substitute for the use of other assessment instruments.[9]

Once the abovementioned constraints have been addressed, collaboration between institutions regarding production and administration of the (same) progress test is feasible due to the curriculum independence.[9, 22, 23] This strategy obviously contributes to cost-effectiveness.[9] Notwithstanding its current popularity in undergraduate medical education.[9, 12], progress testing is so far neither an integral part of ICM training, nor of many other postgraduate (residency) training programs.

Progress testing in ICM training?

Comparable to changes in undergraduate medical training, developmental shifts towards competency-based residency training programs are observed,[24-28] in attempts to "*assist future specialists in responding to the innumerable challenges as health-care providers-…- while providing the best specialty care*".[24-28] Under the aegis of the European Society of Intensive Care Medicine (ESICM), the Competency-Based Training program in Intensive Care in Europe (CoBaTrICE) identified

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

12 competency domains for ICM in 2006.[29] Likewise, the contemporary training program for ICM in the Netherlands is currently undergoing major reform. Although initial developments in competency based training emphasized practice-oriented teaching and learning, and knowing how to use resources adequately, the pivotal role of a well-organized knowledge database as a contributor to medical expertise is now again focus of attention.[30] Medical expertise indeed appears to be based on a doctor's well developed, highly structured and re-shapeable knowledge networks.[31,32] Since focus beyond postgraduate training is likewise on lifelong and active learning, assessment of long term retrieval and structure of knowledge such as progress testing could potentially play a contributory role in the postgraduate training programs or even beyond.[32] Furthermore, it has indeed been hypothesized that the sort of knowledge transferred during every day training differs from that tested using written tests: experience based knowledge and evidence based knowledge.[33]

Prior research has revealed that undergraduates and junior physicians in the UK, Europe and the US alike, in general lack knowledge in many aspects of acute care,[34] although formal training in this field has notably improved since then.[35] The CoBaTrICE competency framework, using an "*empirical competency based approach rooted in everyday practice at the bedside*",[36] does not provide a detailed description of the nature and level of specific knowledge required during each stage and at the end of ICM training. In 2011 the CoBaTrICE collaboration defined international standards for programs of training in ICM. Regarding assessment each training program should incorporate clear assessment processes, focusing on the workplace and providing structured feedback; assessment of knowledge is not specifically addressed.[37] In 2009, the CoBaTrICE collaboration reported that only 50% of European national coordinators for ICM training performed in-training assessment using formal structured assessment of knowledge, skills and attributes during routine clinical activities.[35] In our experience formal knowledge testing in the Netherlands during routine clinical activities using workplace based assessment tools is extremely limited, if not absent. In contrast, most European countries, including the Netherlands, require trainees to complete an examination at the end of training, of which the format and content varies; oral examination, MCQ and clinical examination have been reported to be used most commonly (96%, 69%, 46% respectively).[35] The current format of the European Diploma of Intensive Care Medicine (EDIC) combines two assessment methods: MCQ and an objective structured assessment, is the only international examination available, but is so far used only in a minority of European countries, often without a requirement to pass for certification.[36]

Apart from necessity of knowledge testing, the duration of ICM postgraduate training is another aspect to consider from a feasibility perspective. Although the format and the duration of intensive care training vary between European countries, the most common duration is 24 months (for half the countries participating in the CoBaTRICE collaboration, range 10-60 months),[35] a time frame in which a progress testing framework could be adequately fitted.

Thus, considering the potential for improvement of in-training assessment of ICM knowledge in the European training programs, introduction of a progress test in ICM thus seems worthwhile considering. In this perspective, results and conclusions from previously published studies on progress testing in other postgraduate training programmes should be acknowledged.

Examples of progress testing in postgraduate training

Unfortunately, an absence of studies on progress testing in postgraduate ICM training can be noted. Nevertheless, some outstanding examples of analysis of progress testing in other postgraduate training programs, such as general practice, obstetrics and gynecology, radiology and osteopathic surgery are noticeable, and consecutively discussed in the sections below.

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

General practice

Progress testing in postgraduate medicine has most extensively been studied in the general practice training programs, *e.g.* in the Netherlands and Belgium.[33,38-40] For example, the Dutch Knowledge Test for General Practice is a written test administered to all trainees in the Netherlands three times a year, and consists of 80 patient cases with a total of 160 items, with a true-false-don't know response format. The test format and questions content were found to be highly liked and the majority of items (80%) was considered transferable between the Netherlands and Belgium.[40] The construct and content validity of the test was found to be satisfactory, assessing knowledge closely related to the GP's daily work.[38] Reliability (Cronbach's alpha) was 0.61 to 0.76 and remained stable between 1992 and 1999.[39] When comparing the previous two-year curriculum to the newly introduced three-year curriculum, the latter resulted in a higher level of knowledge compared to the former.[39]

Obstetrics and gynecology

Dijksterhuis *et al.* evaluated the validity and reliability on 10-year data of a national progress test in Obstetrics and Gynecology training in the Netherlands, a test consisting of 150 questions (true false, later single best option MCQs) taken at a yearly interval.[7] Validity and reliability using this framework proved unsatisfactory. Several suggestions for improvement were provided including *e.g.* the inclusion of only relevant items, with a correct item format preferably constructed to reflect medical expertise and/or case based questions in an item bank and increasing sample size or test frequency.[7]

Radiology

Ravesloot *et al.* studied the quality of progress testing in radiology (the Dutch Radiology Progress test) over a five year period in the Netherlands, a test which is taken every half year, and consists of 200 questions (true-false-don't know).[41] The test monitors both knowledge items and visual skills items. Reliability and validity were high. After the fourth year of training no significant increase in test scores on knowledge could be measured on most tests. A comparable pattern was found for the visual skills.[41]

Osteopathic surgery

Shen *et al.* reported on the results of four-year experiment of the American College of Osteopathic Surgeons Resident Examinations. The once-a-year test consisted of 300 one-best answer items. Reliability was high and varied between 0.87 and 0.90 for each of the exams.[42]

In summary, progress testing for postgraduate training has significant potential, and if taken test construction properties into account, appears feasible and acceptable. So far pilot studies in ICM are thus lacking. One may wonder however, whether ICM trainees actually perceive a need for formal assessment including knowledge testing.

Perceived need for more formal knowledge testing by ICM trainees?

Knowledge aspects are considered prerequisite for practice under supervision, as well as independent practice after graduation.[29] Surprisingly, medical knowledge and technical skills were *not* mentioned by ICM trainees as *the* most important aspects of intensivists' professionalism in a nationwide Dutch study.[43] One could therefore speculate whether gathering and retaining knowledge and mastering technical skills become more tacit, are more implicitly (*versus* explicitly) considered important when a resident advances and becomes progressively independent.[44, 45] One could also wonder whether many of the technical dimensions of tacit (applied medical) knowledge are largely incorporated during prior specialist training and early ICM training, whereas the development of other aspects, like *e.g.* the cognitive dimension of tacit knowledge (beliefs, perceptions, values, emotions) is catalyzed in the intensive care unit (ICU) due to its specific context by reflective and deliberate practice

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

and assessed by means of workplace based assessment methods.[43] In another Dutch study exploring the value of different professionalism elements to trainees in ICM, striving for excellence was found to be the most important factor, which included a good knowledge base and technical skills.[46] These do provide some indirect evidence for the importance of addressing knowledge gathering. Combined with the trainees' non-specific request to increase the use of formal teaching courses, progress testing could perhaps fill an important void here. Preliminary thoughts and a proposal for a (hypothetical) framework for progress testing in the ICM training program in the Netherlands are provided in the Appendix.

Conclusions

In the contemporary era of competency based training, knowledge gathering and testing should neither be forgotten nor ignored. Although reports on the application and usefulness of progress testing in postgraduate training are not numerous, progress testing nevertheless seems a possible and promising useful and feasible adjunct to the current and future competency based ICM curricula. Scaffolding of implementation of progress testing in ICM can be performed using existing frameworks as a starting point. European collaborative efforts regarding the transferable test format and content are worthwhile to consider.

Key messages

— Expert problem solving cannot take place without a well organised knowledge database and requires expert knowledge.

— In-training assessment of knowledge is underrepresented in postgraduate assessment frameworks in intensive care medicine, with a predominant emphasis on workplace based learning and assessment.

— Progress testing appears to be a promising useful and feasible adjunct to the current and future competency based intensive care medicine curricula.

References

1. Van Mook W, Bion J, van Der Vleuten C, Schuwirth LW. Integrating education, training and assessment: competency-based intensive care medicine training. Neth J Crit Care 2011;15.
2. The Intercollegiate Board for Training in Intensive Care Medicine Medicine. The curriculum for the CCT in intensive care medicine competency-based training and assessment; Part I: Reference Manual for Trainees and Trainers; [Internet]. Available from: http://wwwrcoaacuk/ibticm/docs/CCT%20in%20ICM%20(Part%201)(Jan%2009%20Logo).pdf [cited 2010, Jun 8].
3. Bion JF, Barrett H. Development of core competencies for an international training programme in intensive care medicine. Intensive Care Med 2006;32:1371-83.
4. Van Mook W, Bion J, WIlde J, Schuwirth L, on behalf of the CoBaTrICE collaboration. Workplace based assessment: how to use it in the ICU. Neth J Crit Care 2010;16:162-73.
5. European Union of Medical Specialists. Common training framework. Training requirements for the core curriculum of multidisciplinary intensive care medicine; [Internet]. Available from: www.uems.net [cited 2016, Jun 16].
6. Van de Wiel MW. Knowledge encapsulation: studies on the development of medical expertise. Thesis. Ponsen & Looijen. 1997.
7. Dijksterhuis MG, Scheele F, Schuwirth LW, Essed GG, Nijhuis JG, Braat DD. Progress testing in postgraduate medical education. Med Teach 2009;31:e464-8.
8. van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. Best Pract Res Clin Obstet Gynaecol 2010;24:703-19.
9. Schuwirth LW, van der Vleuten CP. The use of progress testing. Perspectives on medical education 2012;1:24-30.
10. Dijkstra J, Van der Vleuten CP, Schuwirth LW. A new framework for designing programmes of assessment. Adv Health Sci Educ Theory Pract 2010;15:379-93.
11. van der Vleuten CP, Schuwirth LW, Driessen EW, *et al.* A model for programmatic assessment fit for purpose. Med Teach 2012;34:205-14.
12. Wrigley W, van der Vleuten CP, Freeman A, Muijtjens A. A systemic framework for the progress test: strengths, constraints and issues: AMEE Guide No. 71. Med Teach 2012;34:683-97.
13. Miller G. The assessment of clinical skills/competence/performance. Acad Med65(suppl) 1990;65 (suppl):S63-7.
14. Albano MG, Cavallo F, Hoogenboom R, Magni F, Majoor G, Manenti F, *et al.* An international comparison of knowledge levels of medical students: the Maastricht Progress Test. Med Educ 1996;30:239-45.
15. Muijtjens AM, Timmermans I, Donkers J, Peperkamp R, Medema H, Cohen-Schotanus J, *et al.* Flexible electronic feedback using the virtues of progress testing. Med Teach 2010;32:491-5.
16. Norman G, Neville A, Blake JM, Mueller B. Assessment steers learning down the right road: impact of progress testing on licensing examination performance. Med Teach 2010;32:496-9.
17. Blake JM, Norman GR, Keane DR, Mueller CB, Cunnington J, Didyk N. Introducing progress testing in McMaster University's problem-based medical curriculum: psychometric properties and effect on learning. Acad Med 1996;71:1002-7.
18. Ricketts C, Freeman A, Pagliuca G, Coombes L, Archer J. Difficult decisions for progress testing: how much and how often? Med Teach 2010;32:513-5.
19. Downing SM. Threats to the validity of locally developed multiple-choice tests in Med Educ: construct-irrelevant

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

variance and construct underrepresentation. Adv Health Sci Educ Theory Pract 2002;7:235-41.
20. Downing SM. The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in Med Educ. Adv Health Sci Educ Theory Pract 2005;10:133-43.
21. Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. Med Educ 2008;42:198-206.
22. van der Vleuten CP, Schuwirth LW, Muijtjens AM, Thoben AJ, Cohen-Schotanus J, van Boven CP. Cross institutional collaboration in assessment: a case on progress testing. Med Teach 2004;26:719-25.
23. Schuwirth L, Bosman G, Henning RH, Rinkel R, Wenink AC. Collaboration on progress testing in medical schools in the Netherlands. Med Teach 2010;32:476-9.
24. Lee AG, Beaver HA, Boldt HC, Olson R, Oetting TA, Abramoff M, *et al.* Teaching and assessing professionalism in ophthalmology residency training programs. Surv Ophthalmol 2007;52:300-14.
25. Edelstein SB, Stevenson JM, Broad K. Teaching professionalism during anesthesiology training. J Clin Anesth 2005;17:392-8.
26. Rowley BD, Baldwin DC, Jr., Bay RC, Cannula M. Can professional values be taught? A look at residency training. Clin Orthop Relat Res 2000:110-4.
27. The Royal College of Physicians and Surgeons in Canada; [Internet]. Available from: http://rcpscmedicalorg/canmeds/indexphp [cited 2006, Oct 10].
28. Modernisering Medische Vervolgopleidingen; [Internet]. Available from: http://wwwmedischevervolgopleidingennl/pages/contentaspx?content=10010000000029_2_10000000001651&contentcode=competentiesspecialist [cited 2016, Oct 10].
29. Barrett H, Bion JF. Development of core competencies for an international training programme in intensive care medicine. Intensive Care Med 2006;32:1371-83.
30. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. Acad Med 1990;65:611-21.
31. Cox K. Knowledge which cannot be used is useless. Med Teach 1987;9:145-54.
32. Verhoeven BH. Progress testing: The utility of an assessment concept. Thesis Drukkerij C Regenboog, Groningen ISBN 90-36701805-8 2003.
33. Van Leeuwen YD, Düsman H, Mol SS, Pollemans MC, Drop MJ, Grol RP, *et al.* Factors influencing the growth in knowledge of trainees in general practice. Adv Health Sci Educ Theory Pract 1997;2:61-70.
34. Smith CM, Perkins GD, Bullock I, Bion JF. Undergraduate training in the care of the acutely ill patient: a literature review. Intensive Care Med 2007;33:901-7.
35. Wilde J, Bion J, collaboration obotC. The educational environment for training in Intensive Care Med: structures, processes, outcomes and challenges in the European region. Intensive Care Med 2009;35:1575-83.
36. Bion J, Rothen HU. Models for intensive care training. A European perspective. Am J Respir Crit Care Med 2014;189:256-62.
37. Collaboration C. International standards for programmes of training in Intensive Care Med in Europe. Intensive Care Med 2011;37:385.
38. Van Leeuwen YD, Pollemans MC, Mol SS, Eekhof JAH. The Dutch knowledge test for general practice: issues of validity. Eur J Gen Pract 1995;1:113-7.
39. Kramer AW, Dusman H, Tan LH, Jansen KJ, Grol RP, van der Vleuten CP. Effect of extension of postgraduate training in general practice on the acquisition of knowledge of trainees. Fam Pract 2003;20:207-12.
40. Remmen R, Wens J, Damen A, Duesman H, Verhoeven V. Knowledge assessment of trainees and trainers in general practice in a neighboring country. Making a case for international collaboration. BMC Fam Pract 2012;13:103.
41. Ravesloot C, van der Schaaf M, Haaring C, Kruitwagen C, Beek E, Ten Cate O, *et al.* Construct validation of progress testing to measure knowledge and visual skills in radiology. Med Teach 2012;34:1047-55.
42. Shen L. Progress Testing for Postgraduate Med Educ: a Four-Year Experiment of American College of Osteopathic Surgeons Resident Examinations. Adv Health Sci Educ Theory Pract 2000;5:117-29.
43. van Mook WN, De Grave WS, Gorter SL, Zwaveling JH, Schuwirth LW, van der Vleuten PM. Intensive Care Med trainees' perception of professionalism: a qualitative study. Anaesth Intensive Care 2011;39:107-15.
44. Sturmberg JP, Martin CM. Knowing--in medicine. J Eval Clin Pract 2008;14:767-70.
45. Wyatt JC. Management of explicit and tacit knowledge. J R Soc Med 2001;94:6-9.
46. van Mook WN, de Grave WS, Gorter SL, Muijtjens AM, Zwaveling JH, Schuwirth LW, *et al.* Fellows' in Intensive Care Med views on professionalism and how they learn it. Intensive Care Med 2010;36:296-303.
47. Bjorner JB, Chang CH, Thissen D, Reeve BB. Developing tailored instruments: item banking and computerized adaptive assessment. Qual Life Res 2007;16 Suppl 1:95-108.
48. Roex A, Degryse J. A computerized adaptive knowledge test as an assessment tool in general practice: a pilot study. Med Teach 2004;26:178-83.
49. Gershon R. Computer adaptive testing. Journal of applied measurement 2005;6:109-27.
50. Green B, Bock R, Humphreys L, Linn R, RJ N. Technical guidelines for assessing computerized adaptive tests. J Educ Meas 1984;21:347-60.

Authors' contributions.—Walther N.K.A. van Mook wrote the initial and subsequent versions of the the manuscript together with Cees P.M. van Der Vleuten. Sesmu M. Arbous, Hans Delwig, J.D. van Hemel-Rintjap, Robert Tepaske and Jaap. E. Tulleken commented on the preliminary drafts, and the paper was than finalised in an iterative process.

Conflicts of interest.—The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Article first published online: November 17, 2015. - Manuscript accepted: November 13, 2015. - Manuscript revised: November 10, 2015. - Manuscript received: June 17, 2015.

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

COPYRIGHT© 2016 EDIZIONI MINERVA MEDICA

This document is protected by international copyright laws. No additional reproduction is authorized. It is permitted for personal use to download and save only one file and print only one copy of this Article. It is not permitted to make additional copies (either sporadically or systematically, either printed or electronic) of the Article for any purpose. It is not permitted to distribute the electronic copy of the article through online internet and/or intranet file sharing systems, electronic mailing or any other means which may allow access to the Article. The use of all or any part of the Article for any Commercial Use is not permitted. The creation of derivative works from the Article is not permitted. The production of reprints for personal or commercial use is not permitted. It is not permitted to remove, cover, overlay, obscure, block, or change any copyright notices or terms of use which the Publisher may post on the Article. It is not permitted to frame or use framing techniques to enclose any trademark, logo, or other proprietary information of the Publisher.

Appendix I.—Some preliminary thoughts on a blue print for progress testing: a hypothetical framework for the Intensive Care Medicine training program in the Netherlands

Intensive care medicine (ICM) training in the Netherlands is scaffolded into a so-called multi-subspecialty model. Apart from trainees from internal medicine and anesthesiology, an increasing number of colleagues have a background in cardiology, pulmonary medicine, and neurology. The training program lasts two years. Knowledge on ICM specific aspects of care is expected to be acquired during these two years, whereas knowledge aspects of medicine in general and the subspecialty specifically are expected to be acquired during training *prior to* ICM training. Progress testing in ICM should thus specifically focus on aspects of ICM. Progress testing aims at providing formative feedback to the trainees, although it can be imagined that at the end of the two year training program, the individual's results could aid in making summative decision regarding certification. Progress testing thus fits into a framework of programmatic assessment including workplace-based assessment, progress testing and end-of-training assessment, for example using the European Diploma in ICM format.

Items pertaining to aspects of ICM be constructed by the staff members (and their consultants if necessary) of all nine ICM training centers in the Netherlands, and deposited in a national item bank. The paper-based testing will be performed using a 3 to 4-times a year testing frequency, each time with 200 not recently used multiple choice questions with (scenario-based) single best option response format. The test is administered in a central location, preferably during regular national educational activities for the intensive care trainees. The maximal duration of the test is four hours. After the test, the answers to the questions in the test are provided, so the trainee can study the literature provided for questions wrongly answered. Individual participants results are expressed as an absolute and percentage score relating the individuals' performance to that of the reference group. Standard setting will include development of minimum knowledge standards. Comparing the scores over time reveals a trend of the trainees' progression over time, areas of strengths, as well as areas in which potential improvements are possible and necessary. The scores of the reference group may provide insight and feedback to what extent and where improvements in the formal and informal ICM curriculum are possible.

Future developments could be directed towards online progress testing on the one hand (paper based testing will become obsolete sooner or later), and Computer Adaptive Testing (CAT) on the other.

Online testing has several advantages including the possibility to include newer item formats (such as images, sound, video), more flexibility on when and where to take the test for individual trainees, and the possibility to immediately provide rich feedback on individual items but also on passing or failing the test. Using online testing multiple tests of equal difficulty is needed. This can be achieved by applying psychometric analysis using Item Response Theory (IRT) to estimate the difficulty level of the questions, resulting in a calibrated item bank.[47-49]

Computer Adaptive Testing refers to tailoring the test to the ability of the individual trainee.[49] Every next question is purposefully selected based on performance on previous questions, based on the assumption that question that are too simple or too difficult for a certain trainee contribute little information about that trainees' ability.[50] This will significantly shorten the duration of the test without compromising reliability. Using this approach guessing becomes less of an issue, and the don't know option thus superfluous.[12] The use of CAT has for example been piloted in general practice.[48]