A conversation about the role of medical regulators

Lesley Southgate & Cees P M van der Vleuten

This article is part of a series in Medical Education entitled 'Dialogue'. Each publication in the series will be a transcription of an e-mail discussion about a current issue in the field held by two scholars who have approached the issue from different perspectives. For further details, see the editorial published in Med Educ *2012;**46** (9):826–7. In this volume, Dame Lesley Southgate, St George's Hospital Medical School, London, UK, and Professor Cees P M van der Vleuten, Faculty of Health, Medicine and Life Sciences, Maastricht University, the Netherlands, discuss the role of regulators in the quality assurance and continuing professional development of the health professional workforce.*

Lesley Southgate:

Writing this piece feels risky because it means exposing my uncertainties and, worse, certainties that may seem like arrogance, but here goes anyway! While I am uncertain about the scope of what we might write, what I would like to do to get us going is to say something about current regulatory bodies in the UK and North America to identify what I see as their principal achievement.

In thinking about assessment and regulation of the medical profession, I suppose the first thing to say is that assessment activities generate evidence that is of prime importance for regulators. Sometimes, as I look at the current UK medical scene, it seems as if the UK regulator, the General Medical Council (GMC), is everywhere, and the rebellious words '*mission creep*' come into my head. But looking further afield, and taking a reality check from a Canadian colleague, I was reminded that the Medical Council of Canada, which maintains the national registry of physicians and their qualifications throughout their professional careers, has similarly developed its initial mandate to create a national licensing examination into: 'Initiate and promote, with partners, a national integrated strategy of assessment of physicians throughout their careers'.[1]

The present purpose of the UK GMC[2] is enshrined within UK law. It is: '…to protect, promote and maintain the health and safety of the public by ensuring proper standards in the practice of medicine.'[2] Subsumed within that purpose are four roles: (i) to keep up-to-date registers of qualified doctors; (ii) to foster good medical practice; (iii) to promote high standards of medical education and training, and (iv) to deal firmly and fairly with doctors whose fitness to practise is in doubt.

The principal document that holds all of these roles together is *Good Medical Practice* (2013), which sets out the values for all registered UK doctors and medical students.[3] This document sets out 'the principles and values on which good practice is founded; these principles together describe medical professionalism in action'.[3] In this conversation I want to reflect on how *Good Medical Practice* (GMP) has influenced assessment programmes I have worked on by making explicit the linkage between the regulator and the assessment of students, trainees and doctors in practice. The principles contained have now been elaborated by the GMC and form a basis for all UK postgraduate and undergraduate medical curriculum content in that they mandate the qualities and abilities about which evidence must be gathered within the assessment programmes that are part of those curricula. A similar approach is now in place for revalidation of the UK medical profession. In Canada, the CanMEDS programme serves a similar function by setting out roles for physicians, and the Accreditation Council for Graduate Medical Education (ACGME) competencies fulfil a similar function in the USA.[4,5] What I want to argue is that the emergence of these high-level statements represents the greatest achievement of the regulatory

Correspondence: Cees P M van der Vleuten, Department of Educational Development and Research University of Maastricht, Postal Address: P.O. Box 616, 6200 MD Maastricht, the Netherlands. Tel: +31 43 3885725; Fax: +31 43 3885779; E-mail: c.vandervleuten@maastrichtuniversity.nl

doi: 10.1111/medu.12309

bodies, but that many of the problems in contemporary assessment approaches lie in the way in which they are presently used and frequently misunderstood. Balanced partnerships with other organisations, both professional and public, are the key here. The dynamic between them in developing policies around doctor regulation and the evidence for it provided by assessment programmes is, in my view, not presently achieved. Further, it is inadequately explored in the literature, although it is sensationally covered in the press from time to time.

Cees P M van der Vleuten:
This all makes perfect sense to me. I can also resonate with what you are saying by reflecting on it using my personal theory of assessment, which has been moulded by many years of experience, both in the practice of assessment and in the research we have engaged in. Lambert Schuwirth and I have voiced the basis of this 'theory' in a set of recent papers.[6–8] A very basic question in assessment is the principle of fitness for purpose. So what is the purpose of assessment for regulators? Naturally that purpose is good health care delivered by a competent workforce. But how do we achieve a competent workforce? It can happen only if the individuals in the workforce keep learning. In my view, the assurance of lifelong learning is the prime aim for which a regulator should strive. So the issue here is to develop assessment strategies that help learning. The next purpose for the regulator is to guarantee patient safety by safeguarding the public from incompetent individuals in the workforce. These two purposes should be separated, even firewalled, and treated differently in developing an assessment strategy.
To foster learning, I think a regulator's role is to provide the conditions for learning to occur. You are right; developing competency frameworks and reports on assessment strategy visions provide those conditions. But in fostering learning through assessment it is imperative to connect the assessment as much as possible to the learning itself. Allowing any disconnection is asking for trouble and often leads to the unwanted effect of having learners jump through hoops. So as a regulator, I would not engage in assessment itself (e.g. by proposing certification examinations). Most lifelong learning occurs in the work itself, in the informal learning. A health professional becomes an expert by infinite practice and learning from clinical outcomes. Only a very small part of this learning is formal. Assessment should reinforce both types of learning; assessment should be personalised and relevant. As we know from the continuing medical education and continuing professional development literature and through implementation, interventions are effective only if they are multimodal and relevant to the individual person and address a felt need.[9] As a regulator, I would enforce accredited lifelong learning programmes that *include* embedded assessment. The challenge is to make the assessment relevant to the learning. It therefore needs to be feedback-oriented, both in terms of a screening function (a concept we once termed directed 'self-assessment'[10]) and in terms of an evaluation of what has been learnt.

LS:
That concept leads me to reflect on the design and implementation of the Performance Procedures for the UK GMC.[11] In designing an assessment programme for doctors referred to the GMC for seriously deficient performance, who could face losing their license to practise, we were in uncharted territory, with the only certainty being that we would face legal challenge every step of the way. Several factors gave us the structure we needed, the most important being the concurrent development by the GMC of *Good Medical Practice*[3] and the decision of the development working group to assess performance in actual practice followed by tests of competence. Through the GMP headings and the content of actual practice, we developed the overall framework for determining the content of what to look at and gather evidence about. We applied it to all of the doctors who were assessed through approaches designed for every specialty under the same general blueprint. The other essential condition for starting was that the purpose of the assessment programme was enshrined in regulation in that the evidence we would gather would be used in making a decision about the doctor's fitness to practise. This, of course, had profound implications for the standard against which assessors would judge the evidence from workplace assessment, and set standards for tests of competence.
But the important point for our conversation is that all of the separate judgements made over the 2 days of the peer review visit to practice were collated to the headings within GMP, irrespective of the assessment method which produced them. The methods were tools, not an end in themselves, and some of the data were 'stronger' than others. This process was aided by the use of a computerised template into which assessors (two medical and one non-medical) entered their findings during the visit. There were no rules for considering the large amount of information except that a single piece of evidence could only be used to support one GMP heading and that every piece of evidence must be included. Assessors weighed the evidence from different sources

about the same GMP heading and then reached judgements using the principles of triangulation. Overall, in practice, 500–700 independent judgements were recorded during each peer review. The number of judgements supporting each major heading of the assessors' report was important only insofar as it demonstrated that the assessors had covered all of the areas of GMP during the visit.[12,13]
To my mind, this example illustrates the first consistent use of an overarching framework proposed by the regulator, applied to every specialty, which did not dictate assessment methods, but, rather, forced attention on the content of actual practice. As a consequence, evidence from many sources was judged by assessors who had spent many hours together discussing both the meaning of *Good Medical Practice* and the standards it must be judged against. This seems to me to come right up against the direction you have been proposing in recent years and, although the stakes for these doctors were the highest that could be imagined, in many cases in which remediation was the outcome, it certainly was assessment for learning!

CPMvdV:
I am very impressed by the work you have done in this arena. And I am very impressed by the National Health Service (NHS) National Clinical Assessment Service and its work.[14] Doctors in need of help, usually referred by their employers, are evaluated thoroughly, receive help and support, and, in a small number of cases, are withheld from further practice. As a regulator I would make sure to have an assessment programme in place for underperforming doctors. We don't have such a system in the Netherlands and I am sure we would have fewer incidents if we did. They have been reported regularly in our national press in recent years.
For both types of assessment (of learning and for learning) we need a multitude of assessment methods and a lot of judgements. I welcome the move towards non-standardised assessment in recent years. Standardised assessment has great limitations. If we want to assess complex skills that are also mostly work-embedded or work-related, we have to rely on professional judgement. Imagine a clinical practice in which the health care professional was not allowed to make clinical judgements. That would not be acceptable. Why do we have a different attitude when it comes to assessment? Objectivity and standardisation are typically driven by the psychometric paradigm. There is one 'truth' in this paradigm and everything that deviates from this truth is error or bias. In such a conception of the world, we need to train and calibrate our assessors and standardise assessment protocol in order to reduce any bias. It is not that I am against being objective. Rather, I become worried when striving for objectivity causes us to throw out the baby with the bathwater and I think we have been doing that a lot in our assessment practices. We have been measuring the measurable and ignoring the (so-called) '*too difficult* to measure'. Judgement by knowledgeable people (including the self) is imperative for assessing complex performances. As with a clinical decision, I would assemble a lot of information before I take a high-stakes decision. The high-stakes decision should be a credible one and one that can be defended in court if necessary. Your practice performance procedures are a great example of that, actually proving that (many) subjective judgements may lead to defensible high-stakes decisions.

LS:
As usual in this conversation, I have left it a while before coming back. I think that is because I keep having experiences that are relevant to our exchange, which reinforce my conviction that we are on the right lines, but how to articulate those experiences is less clear. When I was watching the chief executive officer of the British NHS give televised evidence to a parliamentary committee investigating a scandal of excess deaths in one of our hospitals recently,[15] I was so struck by one of his replies that I wrote it down. It was something like: '… everyone was in favour of quality – but everyone thought it meant something different.'[15] I think that in breaking down the regulators' high-level statements into the fragments that are now used in so many assessment checklists or, worse, using a word like 'professionalism' without previous discussion and training, the attempt to achieve meaning results in meaningless or trivial evidence. And, even worse, cynicism all round.[16] I guess that we need to trust our assessors and judges more, provided there are many judgements, and put time and energy into dealing with bias where possible. I think you agree?

CPMvdV:
I absolutely agree. Building rigour into judgements is about more than having many judgements though. We have published a paper in which we proposed procedural measures to further reduce bias and to build credibility or 'trustworthiness' into high-stakes decision making, if that decision calls for a judgement.[17] Having a panel of judges may help, the size of the panel may matter, the independence of the panel from the evidence gathering may matter, the amount of deliberation may matter, the justification of the decision may matter, the

creation of a paper trail may matter. All these strategies of due diligence bring rigour to the decision and are essentially drawn from methodological strategies in qualitative research. The whole point is that judgement is imperative when assessing complex things and we do have ways to deal with subjectivity other than by constraining the judgement so much that one kills what is supposed to be evaluated.
So overall, I completely agree on the vital role of the regulators. By calling for competencies that are very relevant for patient and health care, by stimulating and assuring learning trajectories, by connecting meaningful assessment to learning and by sometimes taking harsh decisions, we are much better prepared for the future. I also think we have come a long way and made huge progress in learning and assessment as a result thereof.

LS:
Reading this back it looks as if the priority is to try to develop much better support for those busy clinicians who, every day, make judgements, not just about patient care, but about the progress and difficulties of trainees and medical students and each other! Greater respect for their input would go a long way to combat the widespread cynicism expressed about the use of some current assessment tools. And of course, by clinicians I don't just mean doctors! And while I am thinking along these lines, we might talk sometime about role models and feedback. It's all connected…

Contributors: this paper is a transcript of an original e-mail correspondence between LS and CPMvdV.
Acknowledgements: none.
Funding: none.
Conflicts of interest: none.
Ethical approval: not applicable.

REFERENCES

1 Medical Council of Canada. Vision, mission and strategic goals. http://mcc.ca/about/vision-mission-goals/. [Accessed 2 August 2013.]
2 General Medical Council. Duties of a doctor. http://www.gmc-uk.org/about/index.asp. [Accessed 2 August 2013.]
3 General Medical Council. Good Medical Practice 2013. gmc-uk.org/gmp2013. [Accessed 2 August 2013.]
4 Royal College of Physicians and Surgeons of Canada. CanMEDS Physician Competency Framework. http://www.royalcollege.ca/portal/page/portal/rc/canmeds. [Accessed 7 October 2012.]
5 Accreditation Council for Graduate Medical Education. ACGME Outcome Project. http://www.acgme.org/outcome/comp/compMin.asp. [Accessed 7 October 2012.]
6 van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Med Educ* 2005;**39**:309–17.
7 Schuwirth LW, van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach* 2011;**33** (6):478–85.
8 van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, van Tartwijk J. A model for programmatic assessment fit for purpose. *Med Teach* 2012;**34** (3):205–14.
9 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;**362** (9391):1225–30.
10 Sargeant J, Mann K, van der Vleuten C, Metsemakers J. 'Directed' self-assessment: practice and feedback within a social context. *J Contin Educ Health Prof* 2008;**28** (1):47–54.
11 Southgate L, Cox J, David T *et al.* The assessment of poorly performing doctors: the development of the assessment programmes for the General Medical Council's Performance Procedures. *Med Educ* 2001;**35** (Suppl 1):2–8.
12 Southgate L, Cox J, David T *et al.* The General Medical Council's performance procedures: peer review of performance in the workplace. *Med Educ* 2001;**35** (Suppl 1):9–19.
13 Schuwirth LWT, Southgate L, Page GG, Paget NS, Lescop JMJ, Lew SR, Wade WB, Barón-Maldonado M. When enough is enough: a conceptual basis for fair and defensible practice performance assessment. *Med Educ* 2002;**36**:925–30.
14 NHS National Clinical Assessment Service. Who we are. http://www.ncas.nhs.uk/about-ncas/ncas-is-changing/. [Accessed 3 August 2013.]
15 Report of the Mid-Staffordshire NHS Foundation Trust Public Inquiry. http://www.parliamentlive.tv/Main/Player.aspx?meetingId=12729. [Accessed 4 August 2013.]
16 Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Med Educ* 2012;**46**:28–37.
17 van der Vleuten CPM, Schuwirth LWT, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol* 2010;**24** (6):703–19.