



Developing self-assessment skills amongst doctors in Nepal

Katrina Butterworth

To cite this article: Katrina Butterworth (2010) Developing self-assessment skills amongst doctors in Nepal, Medical Teacher, 32:2, e85-e95, DOI: [10.3109/01421590903206372](https://doi.org/10.3109/01421590903206372)

To link to this article: <https://doi.org/10.3109/01421590903206372>



Published online: 17 Feb 2010.



Submit your article to this journal [↗](#)



Article views: 925



View related articles [↗](#)



Citing articles: 1 View citing articles [↗](#)

WEB PAPER

Developing self-assessment skills amongst doctors in Nepal

KATRINA BUTTERWORTH

Patan Academy of Health Sciences, Nepal

Abstract

Background: Accurate self-assessment is essential to direct life-long learning. Most research on self-assessment is from the West. This study takes place in Kathmandu, Nepal.

Aim: To develop tools to aid the development of self-assessment skills in Nepali doctors.

Methods: Fifteen doctors were asked to complete three self-assessment tasks per month over a 6-month period; one mini-clinical evaluation exercise, one clinical case review and one significant event analysis. Self-assessment was compared with mentor assessment for each task. Changes over time for each individual were noted. Results were analyzed using SPSS 10.0. Self and tutor scores were compared using Pearson's correlation coefficient. Reliability of the tools was assessed using Cronbach's Alpha. Participants completed a qualitative questionnaire regarding each tool.

Results: All three tools had high content and face validity, as well as reliability. The use of the "intra individual" approach, with multiple assessments over time demonstrated that most doctors were able to accurately self-assess in some areas. Feedback from a senior tutor was vital. Doctors appreciated feedback that was immediate, specific and delivered in a safe environment. Even where self-assessment was less accurate, the process itself helped to develop awareness of key learning issues.

Conclusions: These self-assessment tools are feasible, reliable and valid for the hospital setting in Nepal.

Introduction

Internationally, it is recognized that doctors must be life-long learners in order to maintain good quality services. Medical knowledge is expanding so rapidly that it is not possible to keep up with unless doctors undertake continuous professional development. In Calman's report (1999) on continuing professional development (CPD) in general practice (GP) in the UK, he states:

To survive in a constantly changing environment life-long learning must become a way of life.

One of the key components of life-long learning is the ability to accurately assess ones' strengths and weaknesses in order to efficiently direct learning to address "gaps" in knowledge or skills. The General Medical Council in the UK describes the skills of a life-long learner as: "Reflect on practice, be self-critical, carry out an audit of their own work and others" (General Medical Council 2003).

The overwhelming evidence in the literature is that doctors do not self-assess well in all domains of knowledge, skills and attitudes. Self-assessment is a learnt skill and requires practice (Gordon 1991; Pinsky & Fryer-Edwards 2004).

Two areas that seem to be very important in aiding accurate self-assessment are the use of precise standards and the availability of feedback.

Analytic global rating scales (where performance is broken into component parts that are scored individually first and then

Practice points

- The mini-CEX can be usefully adapted as a self-assessment tool.
- Clinical case analysis helped doctors to develop a habit of self-study and encouraged evidence-based practice.
- The use of the "intra individual approach" enabled important differences in self-assessment ability to be identified.
- The process of self-assessment helped to enhance critical thinking.
- Constructive feedback from a tutor was important.

summed up to give an overall performance score) have been shown to be valid and reliable for assessing both students and physicians in practice (Cohen et al. 2002; Hodges & McIlroy 2003). Such scales need to be calibrated to ensure both tutors and students are using them accurately in the same way (Ward et al. 2002; Holmboe et al. 2003; Norcini 2005). Familiarity with a tool through regular use and training is also important if it is to be used effectively (Lindemann & Jedrychowski 2002; Norman et al. 2004).

Medical education and social cognition literature both clearly demonstrate that individual reflection is insufficient for accurate self-assessment. Feedback needs to be systematically and routinely sought out from reliable and valid external

Correspondence: K. Butterworth, C/o United Mission to Nepal, PO Box 126, Kathmandu, Nepal. Tel: 977 1 5535586; fax: 977 1 5548008; email: mkbutter@wlink.com.np

resources (Pinsky & Fryer-Edwards 2004). Many studies expound the importance of peers or mentors for providing such support and feedback in an acceptable way to practicing physicians (Hoftvedt & Mjell 1993; Challis et al. 1997; Evans et al. 2002). Feedback needs to be timely (Mattheos et al. 2004) and given in a constructive way.

There has been very little research done on the self-assessment ability of doctors from non-Western cultures. A systematic review of the literature on the accuracy of self-assessment by Davis et al. (2006) deliberately excluded non-Western country studies from their data. No explanation was given for this. Mattheos et al. (2004) comment that cultural diversity and its influence on self-assessment patterns needs further research.

Self-assessment is not widely used and accepted in Nepal's medical system, although this is starting to change. In this, Nepal is not so different from the rest of the world, where most doctors are largely untrained in techniques of self-assessment (Spivey 2005).

The specific objectives of this study are:

- (1) To design self-assessment tools that can be used in the context of daily work in a hospital setting in Nepal.
- (2) To evaluate the effectiveness of these tools in terms of improving the ability of junior doctors to self-assess.
- (3) To evaluate the validity and feasibility of these new tools for use in a postgraduate setting.
- (4) To undertake a qualitative assessment of the usefulness and acceptability of these self-assessment tools amongst junior doctors in Nepal.

The working hypothesis of this study was that with practice and feedback each individual's self-score would come closer to agreeing with the tutor score, demonstrating an improvement in ability to self-assess.

Methodology

Setting

Patan Hospital is a 318-bed tertiary-level hospital in Kathmandu, Nepal. Junior doctors who have just completed their 1-year internship work in the General Outpatient and Emergency Department (OPD/ER; which is run by the Department of General Practice) to gain clinical experience prior to applying for postgraduate training posts.

Participants

Fifteen junior doctors, 12 working in the OPD/ER and three from the postgraduate GP training programme participated. Amongst participants, four were women and the others men. This roughly reflects the gender distribution amongst junior doctors in our hospital of just under one-third women. Of the doctors working in the OPD/ER, experience varied from 5 months to 2 years after completing internship and age ranged from 24 to 28 years. The doctors in the postgraduate GP training programme were older (27 to 30 years) and had more experience, varying from 2 to 4 years.

Doctors had to be planning to work in Patan Hospital for at least the 6-month study period. The high turnover amongst our junior doctors in OPD/ER meant that only 12 of our 26 doctors were eligible. All 12 agreed to participate. There was no major difference between the group of doctors who were eligible and those who were not.

Study design

Each participant was asked to complete three self-assessment tasks per month; one mini-clinical evaluation exercise (mini-CEX), one clinical case review and one significant event analysis. Self-assessment was compared with tutor assessment for each task and changes over time for each individual were noted. The study period was 6 months.

The results were analyzed using SPSS 10.0. Doctors' ability to self-assess was examined on an individual basis (an intra individual approach: Ward et al. 2002), using Pearson's correlation (PC) coefficient to look for changes in the difference between self and tutor scores for both individual criteria and overall scores occurring over the 6-month time period.

Internal consistency reliability was assessed for each of the three tools, using Cronbach's Alpha coefficient.

At the end of the study, participants were asked to complete a qualitative questionnaire regarding the acceptability and usefulness of each tool. The author and another member of the GP Department analyzed the responses to the qualitative questionnaire independently, and key themes were identified.

Design of self-assessment tools

During review of the literature, it became clear that the two key attributes of an effective self-assessment tool are the use of precise standards and criterion to assess performance and the availability of feedback from a peer or tutor.

Three different self-assessment tools were designed bearing in mind these attributes:

- Self-assessment using the framework of the mini-CEX (Appendix 1).
- Self-assessment of a clinical case reflective journal (Appendix 2a and 2b)
- Self-assessment of a significant event analysis (SEA; Appendix 3a and 3b)

Mini-CEX tool

The mini-CEX, used widely in the UK and the USA for formative assessment, was modified to include a self-assessment component. The validity and reliability of the mini-CEX as an assessment tool in the clinical setting has been well established internationally. Explicit standards were set, and the scale was calibrated using video clips of a simulated patient, so that both students and researcher assessed performance in the same way (Ward 2002; Holmboe 2003).

The researcher directly observed each consultation assessed, using the mini-CEX rating scale to guide feedback.

Consultations took place in the OPD. Self-assessment and tutor-assessment were undertaken independently, immediately after the consultation. The respective scores were then compared. Discussion took place regarding differences, and an educational plan was made in consultation with the student.

The researcher was familiar with the use of both the mini-CEX and the SEA tools, having used them extensively for the previous 2 years in the GP training programme running in Patan Hospital.

Self-assessment of clinical case and SEA

Two further self-assessment tools were developed that would be less dependent on the immediate availability of a mentor: one for reflection on clinical case and the other for SEA. A structured format was given to guide residents and a rating scale was developed that referred directly to each of the main components. For each of the criteria a five-point Likert scale was used.

The SEA followed a fairly standard format used and validated in many institutions in the UK and the USA. The clinical case required the resident to identify a learning need based on some aspect of the case. The resident then had to go to the literature to find the answer to their question. They were asked to summarize their findings and then reflect on how they might alter their practice in the future based on this new knowledge.

On completion, residents were asked to use the relevant rating scale to add in a self-assessment component. The aim was to further enhance the learning that had taken place and encourage a more critical and reflective approach to their work.

The mentor used the same rating scale to make an assessment of the residents' work, prior to reading the residents' self-assessment. The mentor also added more detailed written feedback.

Results

Mini-CEX self-assessment tool

Each month the tutor observed participants during a consultation in outpatients, using the mini-CEX tool – a total of six observations over time. Following each observed consultation participants completed a self-assessment score, while the tutor completed an identical scoring sheet. Six pairs of scores (self and tutor) were thus obtained for each of the 12 criteria and also 6 summative, "overall" scores. These were analyzed using PC coefficient. Table 1 shows the subanalysis of individual criteria in the mini-CEX for each of the 15 study participants, showing the correlation between self and tutor assessment scores (accuracy of self-assessment) and how these change over the 6-month study period.

The majority of doctors (13/15) showed consistent accuracy in self-assessment for one or more of the individual criteria used for the mini-CEX ("a" in Table 1). Some were accurate in their assessment of their diagnosis and management (Students 13 and 15), others in appropriateness of investigations (Students 11, 13 and 15), others in the respect shown to patients (Student 14) or sensitivity to patient comfort (Student 9). A larger number were able to consistently, accurately assess the appropriateness of their referrals (Students 1, 11, 13 and 14) and the organization of the consultation (Students 2, 7, 9 and 14). Two doctors were accurate in their assessment of

Table 1. PC for mini-CEX: changes in accuracy of self-assessment of individual criteria over 6 successive months.

Student code	History a	History b	Exam a	Exam b	Communication skill a	Communication skill b	Diagnosis	Investigation	Respect	Referral	Organization	Overall
2	0.414	-0.414	-0.169	0.186	-0.828*	-0.414	-0.393	-0.845	-0.655	-0.277	a	0.169
3	-0.707	0.000	0.354	0.447	0.354	0.258	-0.707	0.000	0.354	0.000	-0.577	-0.707
4	0.393	-0.414	0.717	0.447	-0.655	0.039	0.707	0.523	0.414	a	0.923**	0.414
5	-0.548	0.775	a	-0.632	0.674	0.447	0.632	0.000	0.674	0.945	-0.632	0.447
1	0.393	0.488	-0.393	-0.289	-0.098	-0.621	-0.104	-0.169	0.414	a	0.781	0.098
6	-0.474	0.707	-0.884*	-0.577	0.289	-0.289	-0.354	-0.971	0.354	a	0.289	0.289
9	-0.354	-0.884*	-0.354	a	-0.289	0.289	0.000	-0.098	-0.354	-0.189	a	a
10	0.703	0.828*	-0.393	-0.338	-0.655	0.098	0.828*	0.832	-0.355	a	-0.169	0.878*
11	-0.355	-0.078	0.153	-0.414	-0.262	-0.598	-0.064	a	0.878*	a	0.262	0.131
7	-0.131	0.131	-0.663	-0.338	-0.828*	-0.655	-0.393	0.000	-0.655	-0.971	a	-0.338
8	-0.207	0.381	0.131	0.390	-0.213	0.071	-0.828*	-0.835	0.338	-0.866	-0.355	-0.655
12	-0.131	-0.828*	-0.923**	0.594	-0.338	-0.442	-0.393	0.169	-0.393	a	0.488	a
13	0.497	-0.169	-0.828*	-0.431	-0.235	-0.355	a	a	-0.703	a	0.414	-0.621
14	-0.131	-0.828*	-0.393	-0.655	-0.131	-0.507	0.497	0.000	a	a	a	-0.655
15	-0.655	0.447	-0.255	0.814	-0.272	-0.169	a	a	-0.655	0.420	0.639	-0.845*

Notes: A negative correlation shows the doctor initially over-assessed his/her work and then moved either to accurate or under-assessment over the 6-month study period. A positive correlation shows the doctor initially under-assessed his/her work and then moved to either accurate or over-assessment over the 6-month study period.

History a: Facilitates patient telling of story by listening and appropriate use of open and closed questions. *History b:* Responds appropriately to verbal and non-verbal cues. *Exam a:* Efficient logical sequence appropriate to clinical problem. *Exam b:* Sensitive to patients' comfort and modesty. *Communication skills a:* Exploring patient perspectives demonstrating empathy. *Communication skills b:* Clearly explains problem and shared plan of action with patient. *Diagnosis:* Appropriate diagnosis and management plan. *Investigation:* Appropriate selective use of investigations. *Respect:* Shows respect and compassion. *Referral:* Aware of limitations and refers appropriately. *Organization:* Timely and prioritizes. *Overall:* A global "impression" of how the consultation went; a: cannot be computed because at least one of the variables is constant; * $p < 0.05$; ** $p < 0.01$.

overall performance (Students 9 and 12), though not their total score.

Over time, 12 doctors moved from under- or over-assessment of their performance in individual criteria of the mini-CEX to accurate assessment. There was a wide range observed over which actual criterion improved including: the physical examination (Students 10 and 13), clinical diagnosis and management (Students 8 and 10), response to non-verbal cues (Students 10 and 14), respect for patients (Student 11), considering the patient's perspective (Students 1 and 7) and overall care (Student 10).

In addition, four doctors moved from over-assessment of their performance to accurate and then on to under-assessment. This again occurred for several different criteria in the mini-CEX: response to non-verbal cues (Students 9 and 12), overall care (Student 15) and physical examination (Student 6). Looking at the actual scores assigned by the tutor, around 50% of the time the under-assessment occurred

because the doctor had moved toward an excellent performance.

Only two individuals (Students 10 and 14) demonstrated a significant improvement in the overall accuracy of their self-assessment (total score) in the mini-CEX over the 6 months (Table 4). Total score refers to the sum of each of the individual criteria. The results for Student 10 are shown in Figure 1.

Clinical case analysis self-assessment tool

Over six successive clinical case analyses, pairs of self and tutor scores were obtained for the five individual criteria used as well as a "total" score for each case. The difference between self and tutor scores for each of these criteria was calculated and analyzed using PC coefficient. Table 2 shows the correlation between time (pairs of scores taken every month for 6 months) and the changes in accuracy of self-assessment of individual criteria.

There were less significant correlations noted for the clinical case analysis than for the mini-CEX tools. Student 2 moved from under to accurate and then to over-assessment of the quality of his answer. Student 12 moved from over-assessment of the quality of her differential diagnosis to accurate and then on to under-assessment. Tutor assessment showed gradual improvement. While feedback enabled Student 12 to address some deficiencies in her differential diagnosis, she then became overly critical of her performance.

Only one individual (Student 1) demonstrated a clear correlation between self and tutor total score (and thus accurate self-assessment) for the clinical case analysis (Table 4). In this instance, the resident recognized where he had performed better or less well each time but generally under assessed his performance. This is clearly shown in Figure 2.

SEA self-assessment tool

In the SEA, pairs of scores were obtained for each of the five criteria as well as an overall score. The difference between self and tutor scores for each criterion, as a reflection of accuracy

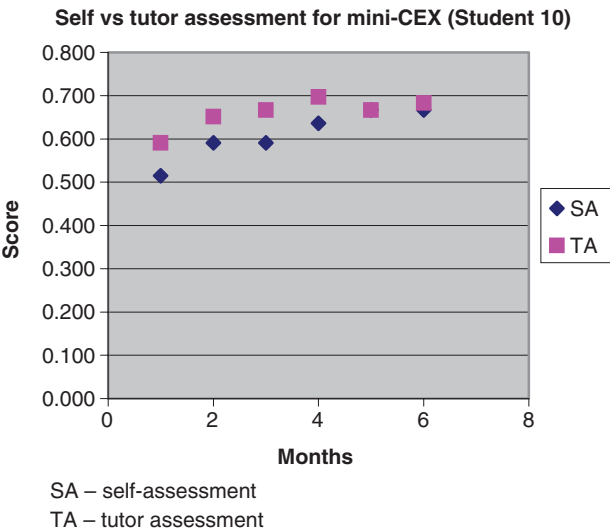


Figure 1. Relation between self and tutor assessment of total score using mini-CEX (Student 10).

Table 2. PC for clinical case analysis: changes in accuracy of self-assessment of individual criteria over 6 successive months.

Student code	History	Differential diagnosis	Question	Answer	Application
2	0.567	0.213	0.224	0.956*	0.886
3	0.131	0.000	−0.869*	−0.717	0.488
5	0.718	0.755	0.258	0.775	a
1	0.478	−0.071	−0.497	−0.272	0.000
6	0.354	−0.433	−0.237	−0.139	−0.189
9	−0.338	0.064	0.169	−0.293	−0.488
10	0.866	0.289	0.289	−0.567	−0.728
11	−0.500	a	a	−0.500	0.866
7	−0.112	−0.289	−0.474	−0.577	−0.354
8	0.781	0.131	0.338	−0.621	−0.704
12	−0.354	−0.949*	0.224	−0.567	−0.671
14	0.338	0.489	0.131	−0.311	0.098

Notes: A negative correlation shows the doctor initially over-assessed his/her work and then moved either to accurate or under-assessment over the 6-month study period. A positive correlation shows the doctor initially under-assessed his/her work and then moved to either accurate or over-assessment over the 6-month study period; a: cannot be computed because at least one of the variables is constant;

* $p < 0.05$.

of self-assessment, and correlation with change over the 6-month time period of the study was analyzed. The results for the nine participants who completed sufficient SEA for analysis are shown in Table 3.

Only Student 9 showed improvement in overall accuracy of self-assessment for the SEA (Table 4), although there was no correlation between individual criteria (Figure 3).

Subanalysis of individual criteria for the other participants, however, revealed that compared to tutor assessment, Student 2 moved from over-assessment of his recognition of strengths and weaknesses to an accurate assessment and then, toward the end of the study period, under-assessment. Tutor assessment showed a steady improvement, though not to excellence.

Student 8 moved from accurate to under-assessment of the awareness of his strengths and weaknesses. During this period, tutor assessment showed a gradual improvement in his actual score.

Student 10 showed perfect agreement between self and tutor assessment in the criteria “Awareness of strengths and weaknesses”, as she demonstrated a gradual improvement in performance over time.

Feasibility of the self-assessment tools

Residents were assessed during routine work in the OPD. The time taken for the consultation to be completed varied from 10 to 20 min with 5 min for self and tutor assessment followed by discussion and feedback. The tutor chose the timing of the assessment to fit in with their schedule and the presence of the resident in the OPD. This worked well.

The clinical case study and the SEA proved more difficult to implement as the timing was in the hands of the residents. Only a few residents were able to complete the tasks according to the schedule. Many seemed to find the SEA more difficult than the clinical case.

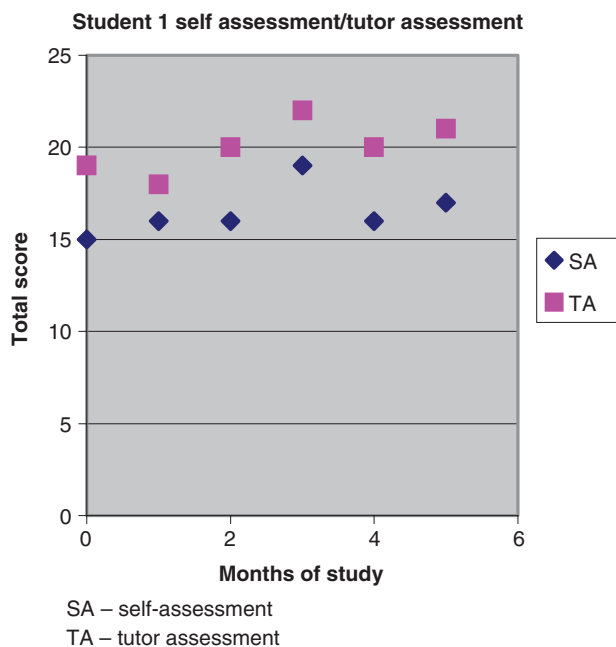


Figure 2. Relation between self and tutor assessment of total score in clinical case analysis (Student 1).

Table 4. . PC for total score student assessment compared to tutor assessment for each of the three assessment tools.

Student code	Assessment tools		
	Mini-CEX	Clinical case analysis	Significant event analysis
1	NS	0.821*	NS
2	NS	NS	NS
3	NS	NS	□
4	NS	NS	□
5	NS	NS	□
6	NS	NS	NS
7	NS	NS	NS
8	NS	NS	NS
9	NS	NS	0.818*
10	0.861*	NS	NS
11	NS	NS	NS
12	NS	NS	□
13	NS	□	□
14	0.913*	NS	NS
15	NS	□	□

Note: NS denotes not significant; □ insufficient data returned data returned to make a calculation; * $p < 0.05$

Table 3. PC for significant event analysis: changes in accuracy of self-assessment of individual criteria over 6 successive months.

Student code	Event	Underlying issues	Reflection on actions	Plan for future	Awareness strength and weakness
2	−0.866	−0.707	−0.756	−0.354	−0.894*
1	0.707	−0.834	−0.289	0.567	0.000
6	0.000	−0.189	0.112	−0.722	−0.224
9	−0.311	−0.447	0.676	0.163	0.213
10	0.000	0.139	−0.289	0.832	a
11	a	0.866	0.866	−0.500	1.000**
7	−0.671	−0.447	−0.493	−0.224	−0.348
8	−0.683	−0.207	−0.414	−0.414	−0.831*
14	−0.207	0.071	0.098	0.488	−0.104

Notes: A negative correlation shows the doctor initially over-assessed his/her work and then moved either to accurate or under-assessment over the 6-month study period. A positive correlation shows the doctor initially under-assessed his/her work and then moved to either accurate or over-assessment over the 6-month study period; a: cannot be computed because at least one of the variables is constant; * $p < 0.05$; ** $p < 0.01$.

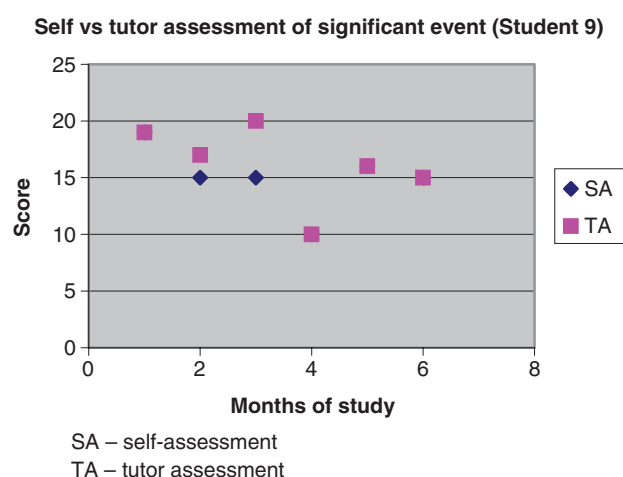


Figure 3. Relation between self and tutor assessment of total score in significant event analysis (Student 9).

Reliability of the tools

As has been found in previous studies, the mini-CEX had high internal consistency with a Cronbach's Alpha of 0.89. The clinical case and SEA also had a high level of internal consistency, with Cronbach's Alpha of 0.84 and 0.847, respectively.

Feedback from residents on usefulness of the self-assessment tools

The majority of doctors were very positive about the mini-CEX finding it useful and practical. The immediacy of feedback from a senior doctor was one of the key factors that made this popular, together with the opportunity for discussion on management of that particular case. A few doctors were less happy with this tool, describing it as being very case dependent, so not giving a representative view of their skills. Some found it intimidating to see a case in front of a senior in case they made a mistake.

Many of these junior doctors commented on how the clinical case analysis in particular had, by compelling self-study, generated a new enthusiasm for learning. Over the 6-month period, they had developed a habit of searching for up-to-date information using the Internet and considering the evidence base for their practice.

The SEA was found most difficult to use. The key issues were difficulty in understanding the process and answering the questions, together with finding an appropriate event. However, some residents found the SEA helpful in learning how to deal with mistakes and handling difficult patients.

Many residents commented on how difficult it was to do self-assessment properly. This was seen as the hardest part of the study. Some expressed a fear of over or underestimating themselves. In this regard, feedback and assessment from a tutor was much appreciated. With time and familiarity, the tools were found to be more useful.

The criteria were felt to add more specificity to the assessment process. They helped individuals to identify

strengths and weaknesses, enabling work to address deficiencies.

Some found the criteria to be repetitive, while others found that the logical sequence of questions helped develop an approach to learning. One individual requested more detailed descriptors. Another wanted more open questions.

One of the key themes to arise from this qualitative questionnaire was the importance of feedback from a senior, experienced tutor. The manner in which feedback was given was found to be important. Interaction and two-way discussion in a non-judgmental atmosphere was appreciated by the majority.

One resident found the feedback "too nice" and would have preferred more criticism. Another resident suggested providing group feedback as well as individual to increase learning.

The residents were overall fairly positive about the value of self-assessment, particularly, as it helped them to develop awareness of the learning process and of their strengths and weaknesses. Writing out the self-assessment helped to reinforce learning.

A key theme was how using these self-assessment tools had changed individuals' thoughts about the learning process. The clinical case analysis, in particular, had helped residents to see the clinical encounter as a rich learning opportunity.

Discussion

A review of these results shows that there was a wide variation between individuals in their ability to self-assess. While many individuals showed improvement over time, it was often in different areas of the consultation. A comparison of self and tutor assessment for just total score rather than breaking down into individual criteria would have missed many significant correlations (Biernat et al. 2003).

Mini-CEX

This study confirms that analytical global rating scales, such as that used in the mini-CEX are useful in the context of analyzing higher levels of clinical competence, particularly, in areas that are conceptually difficult to describe or measure using objective tests, for example, "empathy", "respect for patients" and "awareness of strengths and weaknesses" (Cohen 2002; Hodges 2003; Allen & Velden 2005).

While the mini-CEX method used in this study was not as academically strong an assessment as the use of standardized patients (Cohen 2002; Biernat 2003), it had a number of other significant advantages. First, it had high face and content validity, being rooted firmly in every day practice, so it was directly relevant to these doctors' clinical work and therefore highly transferable knowledge was gained (Little & Hayes 2003). Second, the cost was minimal both in terms of financial expenditure and also in terms of time.

Residents appreciated the formative nature of the assessments, recognizing that they were designed to improve patient care and further their own professional development. The opportunity for dialogue, together with the development of an

action plan, made it a powerful learning tool (Amery & Lapwood 2004).

Calibration of the rating scale for the mini-CEX using a videotape (Holmboe et al. 2003) of three different levels of performance, using a standardized patient did appear to increase the accuracy of self-assessment.

This study, together with the author's experience of implementing the mini-CEX tool within Patan Hospital over the previous 2 years, suggests that the mini-CEX can be readily used in a non-Western environment.

Clinical case and SEA

There was much less accuracy in self-assessment for the individual criteria of the clinical case analysis and SEA. This suggests that doctors found the criteria and rating scales much more difficult to use for these two tools than for the mini-CEX.

What may have improved the use of these tools would be to involve residents in the choosing of criteria and using more detailed descriptors in the rating scale (Burgess et al. 1999; Orsmond et al. 2004).

Despite the difficulties, participants did feel criteria added more specificity to the assessment process, directing learning and helping to identify particular areas of weaknesses or strengths.

Rating scales

Reviewing the actual scores assigned for each criterion in all three tools by both "self" and "tutor" no one gave a "perfect six". Apart from this, the tutor used the full range of the scoring system. Students were less willing to use the full range. There was more agreement between self and tutor assessment where middle scores were assigned by the tutor and less agreement at the extremes of the scale.

The qualitative analysis suggests that some students were tempted to give a "socially desirable" response (as described by Allen & Velden 2005) not wanting to appear boastful or overconfident.

Acceptability of tools

Feedback showed that different people appreciated different tools for learning and self-assessment. This has also been found in the literature (Laidlaw et al. 1995; Davis 1998; Amery & Lapwood 2004) and confirms that offering a range of tools is important in the development of life-long learning skills.

Overall the most popular learning tool was the clinical case analysis. The reason given was that it helped doctors to develop a habit of self-study and encouraged evidence-based practice. This is an important finding as one of the overall aims of this study was to encourage doctors to become life-long learners.

The least popular tool was the SEA. SEA is not routinely done in Nepal, so it was a totally new concept for these doctors. This was reflected in the lack of confidence many expressed in using this tool, although some found it easier to use with practice. Six doctors did not complete enough significant events for data analysis. A few residents found this

the most useful tool of all as it helped them work out how to analyze an event and therefore develop an ability to handle difficult situations, that could not be learnt from a textbook.

External feedback

All participants stressed the importance of having an experienced tutor to aid in accuracy of assessment and to provide further input. This is also noted in the literature (Taras 2003; Pinsky & Fryer-Edwards 2004; Eva & Regehr 2005).

A single researcher was acting as tutor and providing feedback. The purpose of this study was to look at *changes* in accuracy of self-assessment rather than actual marks achieved. Hence, the results of the study are less reliant upon the tutor's actual marks being a "gold standard", as long as they marked consistently.

Certain characteristics of feedback were particularly important. Doctors appreciated its immediacy, the safe non-judgmental environment and its specific, practical nature. The provision of a safe environment may be a particular issue in Nepal where the medical culture is often one of intimidation of juniors by senior staff.

Another important issue was who actually gives the feedback. These junior residents clearly valued a senior person (rather than junior faculty) to be the individual providing feedback. This could be a cultural issue for Nepal, related to the extremely hierarchical nature of medical practice (and social norms) in Nepal society. This is an area that could be explored in more depth.

The intra individual approach

This study suggests that the intra individual approach, using multiple assessments over time, is a useful method of measuring accuracy of self-assessment (Ward et al. 2002). Individuals varied so much as to which tool they found easiest to use and in which particular criteria they demonstrated improving accuracy of self-assessment, that a composite analysis would almost certainly have failed to recognize any ability of these doctors to self-assess. In that instance, this study would have joined the ranks of many others saying that doctors do not self-assess well.

In fact, the use of intra individual analysis was able to demonstrate that these doctors can accurately self-assess in some areas, though less well in others.

Limitations

One limitation of this study was the small number of participants. However, using the intra individual approach meant that this was less important, as we were examining change within individuals over time.

Another limitation relates to the qualitative questionnaire used to assess participants' opinions regarding self-assessment. The medical culture in Nepal is one of great respect for seniors and a wish to say what the questioner wants to hear. To try and counter this, the questionnaire was anonymous and could be returned to the researcher's mailbox within the hospital.

It was also made very clear that an honest response, including any criticism, was important for the validity of the study.

This study was undertaken in a single institution, so the results may not be generalizable to other institutes in Nepal or Asia generally.

Conclusions

All three of the tools developed were demonstrated to be practically feasible for use in a hospital setting in Nepal. They had high content and face validity and all showed a high level of internal reliability.

The use of the “intra individual” approach, with multiple assessments over time, demonstrated that many doctors were able to accurately self-assess in different areas. The use of multiple tools for assessment and feedback from a tutor was important.

Even where self-assessment was not accurate, participants suggested that the process itself helped to develop awareness of key learning issues and enhance critical thinking.

In terms of future work regarding the development of self-assessment and life-long learning skills amongst Nepali doctors, these three tools do appear to be effective. The next stage would be to train a network of more senior doctors to act as mentors and tutors. Adequate training in the use of the criteria and the full range of the rating scales would be important for both tutors and new doctors, for the maximum benefit to be obtained. Orientation and training in how to give good, constructive feedback will also be important.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Notes on contributor

KATRINA BUTTERWORTH is a British general practitioner, working in Nepal as a medical missionary for the past 12 years. Her current work in Patan Hospital includes running the GP training programme and helping to develop a new medical school.

References

- Allen J, Velden RV. 2005. Role of self-assessment in measuring skills. Paper for transition in Youth Workshop 8–10 September 2005.
- Amery J, Lapwood S. 2004. A study into the educational needs of children's hospice doctors: A descriptive quantitative and qualitative survey. *Palliat Med* 18:727–733.
- Biernat K, Simpson D, Duthie Jr E, Gragg D, London R. 2003. Primary care residents self-assessment skills in dementia. *Adv Health Sci Educ Theory Pract* 8:105–110.
- Burgess H, Baldwin M, Dalrymple J, Thomas J. 1999. Developing self-assessment in social work education. *Soc Work Edu* 18:133–146.
- Calman K. 1999. A review of continuing professional development in general practice. London: HMSO.
- Challis M, Mathers NJ, Howe AC, Field NJ. 1997. Portfolio based learning: CME for GPs—a mid-point evaluation. *Med Edu Jan* 31:22–26.
- Cohen R, Amiel G, Tann M, Shechter A, Weingarten M, Reis S. 2002. Performance assessment of community-based physicians: Evaluating the reliability and validity of a tool for determining CME needs. *Acad Med* 77:1247–1254.
- Davis D. 1998. Does CME work? An analysis of the effect of educational activities on physician performance on health care outcomes. *Int J Psychiatry Med* 28:21–39.
- Davis D, Mazmanian PE, Fordis M, Harrison RV, Thorpe KE, Perrier L. 2006. Accuracy of physician self-assessment compared with observed measures of competence: A systemic review. *JAMA* 296: 1094–1102.
- Eva KW, Regehr G. 2005. Self assessment in health professions: A reformulation and research agenda. *Acad Med* 80:S46–S54.
- Evans A, Ali S, Singleton C, Nolan P, Bahrami J. 2002. The effectiveness of personal education plans in continuing professional development: An evaluation. *Med Teach* 24:79–84.
- General Medical Council. 2003. Tomorrow's doctors. London: GMC.
- Gordon MJ. 1991. Review of the validity and accuracy of self-assessments in health professions training. *Acad Med* 66:762–769.
- Hodges B, McIlroy JH. 2003. Analytical global OSCE ratings are sensitive to level of training. *Med Educ* 27:1012–1016.
- Hoftvedt BO, Mjell J. 1993. Referrals: Peer review as CME. *Teach Learn Med* 5:234–237.
- Holmboe E, Huot S, Chung J, Norcini J, Hawkins R. 2003. Construct validity of the mini clinical evaluation exercise. *Acad Med* 78:826–830.
- Laidlaw J, Harden RM, Morris AM. 1995. Needs assessment and the development of an educational programme on malignant melanoma for GPs. *Med Teach* 17: 79–87.
- Lindemann R, Jedrychowski J. 2002. Self-assessed clinical competence: A comparison between students in an advanced dental education elective and in the general clinic. *Eur J Dent Educ* 6:16–21.
- Little P, Hayes S. 2003. Continuing professional development: GPs perceptions of post-graduate education approved (PGEA) meetings and personal professional development plans (PDPs). *Fam Pract* 20:192–198.
- Mattheos N, Nattestad A, Falk-Nilsson E, Attstrom R. 2004. The interactive examination: Assessing students self-assessment ability. *Med Educ* 38:378–389.
- Norcini JJ. 2005. The mini clinical evaluation exercise (mini-CEX). *Clin Teach* 2:25–30.
- Norman GR, Shannon SI, Marrin ML. 2004. The need for needs assessment in continuing medical education. *BMJ* 328:999–1001.
- Orsmond P, Merry S, Callaghan A. 2004. Implementation of a formative assessment model incorporating peer and self-assessment. *Innov Educ Teach Int* 41: 273–290.
- Pinsky LE, Fryer-Edwards K. 2004. Diving for PERLS: Working and performance portfolios for evaluation and reflection on learning. *J Gen Intern Med* 19:582–587.
- Spivey BE. 2005. CME in the U.S.: Why it needs reform and how we propose to accomplish it. *J Contin Educ Health Prof* 25:6–15.
- Taras M. 2003. To feedback or not to feedback in student self-assessment. *Assess Eval High Educ* 28: 549–565.
- Ward M, Gruppen L, Regehr G. 2002. Measuring self-assessment: Current state of the art. *Adv Health Sci Educ* 7:63–80.

Appendix 1: Mini-clinical evaluation exercise: Self-assessment

Dr's name:Date

Complexity of case: Low ☐ Average ☐ High ☐

Clinical problem:

Complete questions using a cross in the box ☒

*If you are unable to comment please mark U/C

Please grade the following areas using the scale given	Below expectations		Borderline	Meets expectations	Above expectations		U/C*
<i>History taking</i>	1	2	3	4	5	6	7
Facilitates patient telling of story by listening and appropriate use of open and closed questions	☒	☒	☒	☒	☒	☒	☒
Responds appropriately to verbal and non-verbal cues	☒	☒	☒	☒	☒	☒	☒
<i>Physical examination skills</i>							
Efficient logical sequence, appropriate to clinical problem	☒	☒	☒	☒	☒	☒	☒
Sensitive to patient's comfort, modesty	☒	☒	☒	☒	☒	☒	☒
<i>Communication skills</i>							
Explores patient perspective, demonstrating empathy	☒	☒	☒	☒	☒	☒	☒
Clearly explains problem and shared plan of action with patient	☒	☒	☒	☒	☒	☒	☒
<i>Clinical judgment</i>							
Appropriate diagnosis and management plan	☒	☒	☒	☒	☒	☒	☒
Appropriate, selective use of investigations	☒	☒	☒	☒	☒	☒	☒
<i>Professionalism</i>							
Shows respect and compassion	☒	☒	☒	☒	☒	☒	☒
Aware of limitations and refers appropriately	☒	☒	☒	☒	☒	☒	☒
<i>Organisation/efficiency</i>							
Timely and prioritizes	☒	☒	☒	☒	☒	☒	☒
<i>Overall clinical care</i>							
	☒	☒	☒	☒	☒	☒	☒

Areas where you see need for improvement plus action plan:

Areas done well:

Appendix 2a: Clinical case: Reflective journal

Dr's Name:Date

1. Concise history and examination with appropriate positive and negative findings.
2. Differential Diagnosis and problem list
3. Clinical question that you want to research
4. Answer found in the literature (give literature source)
5. How might you change your practice in the light of what you have learned?

Appendix 2b: Self-assessment of clinical case reflection

Dr's name:Date

Please circle the number which best corresponds to your self-assessment of each area, where:

1 – disagree totally, 5 – strongly agree

*U/C Please mark this if you feel unable to comment

Please grade the following areas using the scale given		U/C*
Concise history and examination with appropriate positive and negative findings	1 2 3 4 5	Ⓢ
Appropriate differential diagnosis and problem list	1 2 3 4 5	Ⓢ
Relevant clinical question chosen, clearly defined	1 2 3 4 5	Ⓢ
Research answer to question using current best evidence in the literature and clear summary made.	1 2 3 4 5	Ⓢ
Shows can apply this knowledge to another situation	1 2 3 4 5	Ⓢ

Appendix 3a: Significant event analysis

Dr's Name:Date

1. What happened? Describe the event.
2. What essential factors contributed to the experience? What were the significant background factors?
3. Reflect. Why did you act in the way you did? What were the consequences for: i. you ii. the patient iii. other staff/the hospital?
4. What other choices did you have? What might the consequences of these have been?
5. Suggested steps to be taken to avoid similar events in the future. Consider personal factors and also system changes where appropriate.
6. What have you learnt from this event? If there are things you need to learn to stop this happening again, how might you learn these things?

Appendix 3b: Self-assessment of significant event analysis

Dr's name:Date

Please circle the number which best corresponds to your self-assessment of each area, where:

1 – disagree totally, 5 – strongly agree

*U/C Please mark this if you feel unable to comment

Please grade the following areas using the scale given		U/C*
Concise explanation of significant event	1 2 3 4 5	⓪
Excellent analysis of underlying issues contributing to the incident	1 2 3 4 5	⓪
Excellent reflection on own practice in a Constructive way	1 2 3 4 5	⓪
Makes a clear and realistic plan to address the identified issues in future practice	1 2 3 4 5	⓪
Shows awareness of strengths and weaknesses	1 2 3 4 5	⓪

Appendix 4. Questionnaire on usefulness of self-assessment

Thank you for all your help during the 6-month period of this study to help develop your self-assessment skills. One last task! Please would you complete this questionnaire about the study as honestly as possible. I won't be upset with criticism.

- (1) What do you think was most helpful for your learning during this study?
- (2) Was there any part of this study, or any aspect of it that you didn't like, or that you found difficult? Please explain.
- (3) Do you have any specific comments to make about each of the three tools? Please describe what you found most helpful and any particular difficulties.
 - (i) Mini-clinical evaluation exercise (mini-CEX) and self-assessment
 - (ii) Clinical case analysis and the self-assessment
 - (iii) Significant event analysis and the self-assessment
- (4) Which of the three tools did you find most useful in terms of learning. Please explain
- (5) How easy (or otherwise) did you find it to use the assessment criteria for self-assessment?
- (6) Please comment on the feedback you received, both verbal (after the mini-CEX) and written (after the clinical case and significant event analysis). Was it helpful or not? Any practical points you want to make.
- (7) Has taking part in this study changed the way you think about the way you learn and keep up to date? Please explain.
- (8) What do you think about the value of self-assessment?
- (9) What do you think about the value of assessment criteria to help you assess your own work?
- (10) Do you have any suggestions on how we could improve these tools, or any general comments?

Thank you again for your help. You can return this form anonymously to my post box, number 38 at Patan Hospital reception – Dr Katrina.