Do Quality Indicators for GP teaching practices predict good outcomes for students?

M Bartlett, J Potts, R.K. McKinley

Abstract (200 words)

Introduction

Keele medical students spend 113 days in general practices over our five-year programme. We collect practice data thought to indicate good quality teaching.

We explored the relationships between these data and two outcomes for students; OSCE scores and feedback regarding the placements. Though both are surrogate markers of good teaching, they are widely used.

Method:

We collated practice and outcome data for one academic year. Two separate statistical analyses were carried out: 1) to determine how much of the variation seen in the OSCE scores was due to the effect of the practice and how much to the individual student. 2) to identify practice characteristics with a relationship to student feedback scores.

Results:

1) OSCE performance: 268 students in 90 practices: six quality indicators independently influenced the OSCE score, though without linear relationships and not to statistical significance. 2) Student satisfaction: 144 students in 69 practices: student feedback scores are not influenced by practice characteristics.

Conclusions:

The relationships between the quality indicators we collect for practices and outcomes for students are not clear. It may be that neither the quality indicators nor the outcome measures are reliable enough to inform decisions about practices' suitability for teaching.

Main Paper (2911 words)

INTRODUCTION

Good quality teaching is assumed to lead to better outcomes for learners, however, an understanding of the meaning of 'quality' in educational settings is based on abstract constructs or value systems, and as such is very difficult to define and measure. [1-3] There are no straightforward directly measurable indicators that reflect all aspects of learning, so surrogate measures are needed. [4, 5] These may give an indication of quality, but will not reflect many important elements some of which may be unmeasurable, while undue weight may be given to those which are measurable but contribute little, or which may be misleading. Using examination performance as a measure of success demonstrates these limitations; it is affected by many factors, not all of which are related to the 'teaching' that has been offered, nor to the knowledge which learners construct as a result. Learners may compensate for poor learning opportunities, and, as a result, perform well in examinations. [6, 7] However, OSCE performance is an established proxy outcome measure for the effectiveness of community placements. [4, 8, 9]

Learners' perceptions of their experiences are commonly used as outcome measures, [10] but what learners value may be very different from what faculty values, and the usefulness of their feedback as a measure of quality depends on the nature of the questions.

An important determinant of the effectiveness of teaching is the relationship between learners and clinical teachers; the social aspects of the learning environment. [11, 12] In attempting to measure quality, these difficult to measure factors must be taken into account.

Cotton et al [13] published a common set of quality criteria for community based education developed though national discussion with stakeholders. These included practice factors such as the physical and educational environments, and characteristics of the teachers themselves. The criteria

were further categorised into essential and desirable characteristics for occasional or intensive teaching. Although the consensus is that these are linked to quality (and by implication, good outcomes for learners), we are not aware that this has ever been tested empirically.

At Keele, 24% of the clinical curriculum across the five-year programme is delivered in primary care, which is high by UK standards, [14] in around 100 local practices. [15] In the academic year 2012-13, we delivered 3208 placement weeks in general practice. Quality assurance of primary care teaching is therefore of great importance to our School and we monitor a variety of indicators including students' feedback about their experiences and the relationships they have had with people in their practices. [11, 12] All of our indicators are included in Cotton et al's consensus list [13] and the majority are in the 'essential' category for intensive teaching

This study aimed to explore the relationships between a selected group of quality indicators and two outcomes for students; their scores in OSCEs and their experiences of learning in general practice. Our rationale for using the OSCE scores was that it is a clinical examination designed to test students' learning as a result of their clinical placements. Both year 4 and year 5 students had spent time in primary care placements; Year 4 students having had nine weeks over two years and the year 5 students having had a further fifteen weeks at the time of their OSCE. [16] The mix of stations in the year 5 OSCE reflects the equal proportion of learning in primary and secondary care in the final year of the programme. We expected, therefore, that placement-based learning would influence their performance, particularly in year 5, but needed to be cautious about how much weight we gave to the findings relating to this outcome.

Clear relationships between particular quality indicators and outcomes for students would mean that the indicators would be useful in informing practice recruitment decisions, designing tutor development activities, enhancing quality assurance processes, and managing the educational risks associated with students' spending prolonged periods learning in sites remote from the School and its associated teaching hospitals.

We included a measure of rurality, which is not listed as an indicator of quality by Cotton et al, [13] but as a number of rural practices are involved in teaching our final year students, we were interested to know if, and how, such rural placements might affect the outcomes.

Method

Ethical approval was granted by the School of Medicine's Ethics Committee in August 2010.
Sample: data relating to all students placed in primary care in years 4 and 5 (n=268) and all the teaching practices in which they were placed during the academic year 2012-13 (n=90) were included.

Data collection: this was in two parts. Firstly, the data relating to practices was extracted from our existing records. These data are collected routinely during practice recruitment and teaching review visits and from the annual reports submitted by teaching practices. Secondly, data relating to students' performance in clinical examinations was used in accordance with the School's policy on the use of assessment data in research, and 'end of placement' student survey data, which is routinely scored (see appendix 1) and fed back to teaching practices, were collated by practice.

The data relating to the quality indicators were assigned numerical values as described below. The OSCE scores and the student feedback scores were recorded as numerical values.

All data was collated by practice or student number for ease of analysis and to maintain confidentiality for both.

The 'quality indicators'

1. The number of tutor development sessions attended in the most recent academic year. Our terms of agreement with the practices require that a practice teacher attends a minimum of two half

day sessions per year per year group taught; some attend more, some fewer and others, in some academic years, none. We consider this to be a measure of the GP tutors' commitment to teaching.

2. The number of years the practice has taught.

3. The practices' physical environment score. This score was constructed by assigning one point to each of the following

- Space in consulting rooms to accommodate students when consulting with patients alongside their tutor.
- A room in which a student can consult alone with patients
- A practice library
- A place for quiet study or rest
- A conference or meeting room

The maximum score was thus five.

4. The list size to whole time equivalent (WTE) ratio for the practices. The ratio of the number of registered patients to the number of doctor sessions available is an accepted measure of the 'busyness' of practices in the United Kingdom. [17] For recruitment purposes, a ratio in the range of 1500-2000 is considered desirable. [13] We consider that this range means that doctors are likely to have time to teach.

5. The enthusiasm of the tutors. At Keele, a clinical lecturer is responsible for one of four geographical 'patches', and works with the practices in their patch to develop and support their teaching. Each was asked to consider the enthusiasm of each of their practices for teaching and assign them a score of between one (low) and five (high).

6. Postgraduate teaching in the practice. If the practice was involved in postgraduate teaching in the foundation or specialty training years, it was assigned a value of one, if not it was assigned a value of zero.

7. The rurality of the practices. Patch lecturers were asked to categorise their practices as urban, mixed or rural. The accepted definition of rurality in England is a community of fewer than 10,000 people, [18] however, market towns with populations close to 10,000 were included in the mixed rather than the urban or rural categories as we considered this to be a more accurate reflection of their populations' characteristics from a general practice perspective.

8. The quality of the written summaries of workplace based assessments. We use a validated consultation skills assessment tool [19] in our workplace based assessments. Tutors provide written summaries of the verbal feedback they provide for students. We have developed a scale for scoring the quality of this feedback and the scores are fed back to practices as part of our tutor development and quality assurance processes. [20] The average score per practice on the seven-point scale was recorded as a number.

The outcomes for students

1. Year 4 and 5 OSCE scores: The students' overall scores were recorded as numerical values and collated by practice.

2. The students' feedback about their experience in the practices: At the end of each placement, students are asked to give feedback about their experiences, in four domains. These are scored by means of an in-house system (see appendix 1):

- the practice as a place to learn
- the GP as teacher
- the clinical exposure

- assessments and feedback

The total scores were calculated for each student and collated by practice.

Statistical Analysis

Two separate analyses were conducted, one relating to each of the outcomes. This was to avoid bias, as we had OSCE data relating to all students and therefore all practices but this was not the case for the student feedback scores; giving feedback is voluntary and not all students chose to give it. For the OSCE scores, multilevel modelling was conducted to determine how much of the variation seen in the scores was due to the effect of the practice and how much to the individual student. For the students' perceptions of the placements, a multivariate linear model was fitted to identify individual practice covariates with a relationship to student feedback.

Further details of the statistical analysis can be found in Appendix 2.

RESULTS

1. The OSCE arm

Study Characteristics

268 undergraduate medical students in their 4th and 5th years were placed in 90 different practices (range 1-10 students per practice). A summary of the statistics of the routinely collected data for these practices is given in Table 1. Of 268 students, 138 were in their 4th year of study and 130 in their 5th year. Over 50% of the practices had been teaching students for 5 years (the maximum number). 46% of practices taught postgraduates. The enthusiasm of the tutor and the quality of the summaries of workplace based assessments were evenly spread over the range of possible scores. 70% of the practices were in an urban setting and over 90% had the highest possible physical environment score.

Insert table 1 here

Model Results for part 1 – the OSCE scores

The final statistical model had a total of six practice characteristics which appeared to influence the variation seen in in the OSCE scores. The between practice variance was 7.58% (95%CI 2.25, 25.58) with a between student variance of 30.12% (95% CI 23.48, 38.64). This means that 7.58% of the variation in students' OSCE scores is attributable to the practices in which they were placed. It was found that at a practice level, the number of tutor development sessions attended, the number of teaching years, practice's physical environment, quality of the written summaries of workplace based assessments, whole time equivalent to list size ratio, and enthusiasm of the tutors independently influenced students' overall OSCE score, as did the students' year group. It was found that on average a student in their 5th year of study had an overall OSCE score 4.47 (95% CI 2.25, 6.69) marks higher than a student in their 4th year. The student year of study was the only characteristic which had a statistically significant influence on the OSCE scores. The List size:WTE ratio covariate had borderline significance as the confidence interval includes zero but only just. Each unit increase in the List size:WTE ratio of the practice a student was placed in increased the average OSCE score by 0.002 (95% CI 0.00, 0.01).

The full results can be seen in Table 2

Insert table 2 here

2. The Student Feedback Arm

Study Characteristics

There were 69 practices included in the analysis. A summary of the statistics of the routinely collected data for these practices is given in Table 3. Over 53% of the practices had been teaching students for 5 years (the largest number of teaching years); however, over 20% of practices were in their first year of teaching. 48% of practices taught postgraduates as well as undergraduates. The

enthusiasm of the tutor received equal scores across the practices. 74% of the practices were based in urban settings and the remaining 26% were split equally between rural and mixed environments. 38% of practices had attended no tutor development sessions and 5% had attended more than 5 sessions, in the previous academic year.

Insert table 3 here

Model Results part 2- students' feedback scores

No practice characteristics were statistically significant in the model, suggesting that the routinely collected practice data are not able to predict the average feedback score given to the practice (see table 4).

Insert table 4 here

DISCUSSION

The analysis suggests that there are no strong relationships between the practice characteristics which had been assumed to indicate quality of teaching. Practice characteristics are linked to students' OSCE scores, being responsible for 7.58% of the variation in scores, but with non- linear relationships which do not reach statistical significance (apart from the list size:WTE ratio which had borderline significance). This suggests that some of the variation seen in OSCE scores may be due to the characteristics of practices selected in this study, but other unknown factors are also influencing the scores. A large proportion of the variation in scores is attributable to student characteristics (30.12%), which is to be expected. The relationship between practice characteristics and students' OSCE scores is likely to be complex because it will be affected by student characteristics such as interest, ability, drive and ambition. The only significant covariate in the final model for part one is the year of study of the student. What is more difficult to understand is the finding that practice characteristics do not predict students' self-reported experiences in their practices.

The 'busyness' of the practice as measured by the list size:WTE ratio has an influence on OSCE scores which is bordering on significant, and suggests that a higher ratio leads to a higher OSCE score. This is difficult to explain; it is easy to assume that higher ratios mean that the GP tutors have less time for teaching. The ratio does not take into account the overall size of the practice, the characteristics of the patient population nor the variety of clinical services provided.

While excellent teaching and learning can occur in poor facilities, one would expect that that it would improve with better facilities. [21] The physical environment score is made up of a number of factors that were agreed in Cotton et al's work [13] to be important indicators of quality. The absence of a significant relationship between student OSCE performance and faculty perception of the practices' enthusiasm for teaching is interesting; the characteristics of practice teachers, [3, 13, 22,23,24] especially their enthusiasm for teaching [3,13] have been seen as important in influencing outcomes for students. However, the reliability of our assessments of the enthusiasm of practice teachers is unknown and may be influenced by factors such as relationships within a community of practice [11, 12] and local opinion or hearsay.

Neither OSCE performance nor student experience were related to postgraduate training status or rurality. Educators in postgraduate training practices generally have more educational training and experience, and more opportunities for near peer teaching are provided, so we had expected positive relationships. It is possible that the presence of multiple learners means that the needs of each are less easily met. There is much in the literature about the benefits for students from learning in rural practices [25] but this work has not demonstrated such impacts. This is likely to be because students are following the same curriculum regardless of where they are placed.

Strengths and weaknesses. The analysis has been conducted by an independent statistician (JP) who has conducted a multilevel analysis which accounts for the clustering of students within practices. We have OSCE outcome data but not experience data for all students which is a weakness. Two

variables (practice enthusiasm and to a lesser extent, physical environment) are dependent on subjective faculty assessment.

Though we used data relating to two entire year groups, numbers are modest.

Implications for practice

This study has cast doubt on whether the data we routinely collect should be used to inform recruitment and retention decisions about teaching practices and therefore the utility of collecting it. We may be erecting unnecessary barriers to teaching or imposing an unnecessary burden on practices in terms of reaching or maintaining standards. [13] There are no clear indications of how we could change what we do as a result of this work, except to be more cautious about how much weight we place on these data when making decisions about practices.

We need to explore the relationships between 'quality indicators' and outcomes for students further, and many aspects of this exploration require qualitative methods. In particular, we want to explore the relevance and influence of the number of years of teaching experience, the issue of enthusiasm and how this affects students' learning and is perceived and quantified by faculty, and how the physical environment affects learning.

Conclusion

In this study, the relationships between the quality indicators recorded for practices and outcomes for students are not clear. It may be that neither the data items nor the outcome measures are reliable enough to act as indicators of the quality of teaching to inform decisions about practices' suitability for teaching. Though there are some generally accepted indicators of quality in

community based teaching for medical students, there is perhaps a need for large scale testing of their utility.

References

[1] Suwanwela C. A Vision of quality in medical education. Acad Med. 1995; 70 (no 7 supplement) S32-S37.

[2] Vroeijenstijn AI. Quality assurance in medical education. Acad Med.1995; 70 (no 7 supplement) S59-S67

[3] Bowen JL, Stearns JA, Dohner C, et al. Defining and evaluating quality for ambulatory care educational programs. Acad Med.1997; 72 (6), 506-510.

[4] Esterman A, Prideaux D, Worley P. Cohort study of examination performance of undergraduate medical students learning in community settings. BMJ. 2004; 328. 207-209

[5] Richardson JTE. Instruments for obtaining student feedback: a review of the Literature. Assess Eval High Educ. 2005; 30 (4) 387-415

[6] Ten Cate O. What happens to the student? The neglected variable in educational outcomes research. Adv Health Sci Educ Theory Pract. 2001;6(1):81-88

[7] Dornan T, Muitjens A, Hadfield J, et al. Student evaluation of the clinical 'curriculum in action'. Med Educ. 2006; 40:667-674

[8] Murray E, Jolly B and Modell M. A Comparison of the Educational Opportunities on Junior Medical Attachments in General Practice and in a Teaching Hospital: A Questionnaire Survey. Med Educ. 1999;33, 170-176

[9] Worley P, Silagy C, Prideaux D et al. The parallel rural community curriculum: an integrated clinical curriculum based in rural general practice. Med Educ.2000; 34 558-565

[10] Kember D, Leung DYP, Kwan KP. Does the use of student feedback questionnaires improve the overall quality of teaching? Assess Eval High Educ. 2002;27 (5) :412-425

[11 Worley P, Prideaux D, Strasser R, et al. Empirical evidence for symbiotic medical education: a comparative analysis of community and tertiary-based programmes. Med Educ.2006;40: 109-116

[12] Wenger E. Communities of practice: learning, meaning, and identity. New York: Cambridge University Press; 1998.

[13] Cotton P, Sharp D, Howe A, et al. Developing a set of quality criteria for community based education in the UK. Educ Prim Care. 2009;20, 143-51.

[14] Harding A, Rosenthal J, Al-Seaidy M, et al. Provision of medical student teaching in UK general practices. BJGP.2015; DOI:10.3399/bjgp15X685321

[15] Bartlett M, Webb M, McKinley RK. The Investment of a new medical school in its local primary care community. Educ Prim Care. 2015;26: 89-94

[16] McKinley RK, Bartlett M, Coventry P, et al. The systematic development of a novel integrated spiral undergraduate course in general practice. Educ Prim Care. 2015; 26:189-200

[17] NHS England Improving general practice- a call to action evidence pack 2014 [internet]. London [cited 2015 Mar 7]. Available from: http://www.england.nhs.uk/wp-content/uploads/2013/09/igp-cta-evid.pdf

[18] Department for the Environment, Food and Rural Affairs (UK Government).2016. Defining rurality. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/495639/Defining_rural_areas__Jan_2016_.pdf (accessed 10.2.2016)

[19] Lefroy J, Gay S, Gibson S, et al. Development and face validation of an instrument to assess and improve clinical consultation skills. IJOCS. 2012; 5 (2) 77-87.

[20] Bartlett M. Improving the quality of written feedback from workplace based assessments: the utility index in action. Paper presented at the 16th Ottawa Conference on the assessment of competence in medicine and the healthcare professions; 2014 April 25-29; Ottawa, Canada.

[21] Maslow AH. A theory of human motivation. Psych rev. 1943; 50(4); 370-396

[22] Shores JH, Clearfield M, Alexander J. An index of students' satisfaction with instruction. Acad Med. 2000; 75 (10), S106-108.

[23] Lawrence SL, Lindmann JC, Gottlieb M. (1999) What students value: learning outcomes in a required third-year ambulatory primary care clerkship. Acad Med. 1999; 74(6), 715-717

[24] Robins LS, Gruppen LD, Alexander GL, et al. (1997). A predictive model of student satisfaction with the medical school learning environment. Acad Med. 1997; 72 ,134-139

[25] Barrett FA, Lipsky MS, Lutfiyya MN. The impact of rural training experiences on medical students: A critical review. Acad Med. 2011;86 (2), 259-263.

Table 1: Summary statistics of routinely collected practice data for part 1 - the OSCE scores

	n=90	%
No of previous teaching years		
less than 1	19	21.11
2	9	10
3	9	10
4	5	5.56
5	48	53.33
Teach postgraduate students		
Yes	41	45.56
No	49	54.44
Rurality		
Rural	14	15.56
Mixed	13	14.44
Urban	63	70
Physical Environment score(n=84)		
3	2	2.38
4	6	7.14
5	76	90.48
Enthusiasm of tutors		
2	10	11.11
3	26	28.89
4	33	36.67
5	21	23.33
No of tutor development sessions attended		
0	35	38.89
1	15	16.67
2	16	17.78
3	11	12.22
4	6	6.67
5 or more	7	7.77
Summary score of workplace assessments (n=87)		
less than 3	12	13.79
4	15	18.39
5	21	24.14
6	24	27.59
7	14	16.09
Whole Time Equivalent to list size ratio mean (SD, range) (n=79)	1864.9 (351.0, 928-2800)	

Table 2: Estimates and 95% CI of final model for part 1- the OSCE scores

	Parameter Estimate	95% Confidence Interval
Constant	67.69	(57.86, 77.52)
WTE Ratio	0.002	(0.00, 0.01)
Year 4 student	--	
Year 5 student	4.47	(2.25, 6.69)
Number of previous teaching years		
Less than 1	--	
2	-3.72	(-8.97, 1.53)
3	-3.18	(-7.50, 1.15)
4	4.57	(-1.56. 10.70)
5	0.16	(-2.99, 3.32)
Physical environment score		
3	--	
4	0.44	(-7.73,8.60)
5	-2.95	(-9.82,3.92)
Enthusiasm of tutors score		
2	--	
3	-3.67	(-8.01, 0.68)
4	-3.78	(-8.02, 0.47)
5	-3.16	(-7.54, 1.21)
No of Tutor Development Sessions attended		
0	--	
1	1.06	(-2.28, 4.40)
2	-0.06	(-3.20, 3.08)
3	2.71	(-1.13, 6.54)
4	3.18	(-1.93, 8.30)
5 or more	-1.44	(-6.28,3.39)
Average quality of work based assessments		
Less than 3	--	
4	0.42	(-3.84, 4.67)
5	0.25	(-4.06, 4.55)
6	0.22	(-3.79, 4.23)
7	-0.47	(-5.18, 4.24)

Table 3: Summary statistics of routinely collected practice data for part 2 - the Student Feedback scores

	n=69	%
No of previous teaching years		
less than 1	14	20.29
2	8	11.59
3	7	10.14
4	3	4.35
5	37	53.63
Teach postgraduate students		
Yes	36	52.17
No	33	47.83
Rurality		
R	9	13.04
M	9	13.04
U	51	73.92
Enthusiasm of tutors		
2	10	14.49
3	20	28.99
4	22	31.88
5	17	24.64
No of tutor development sessions attended		
0	26	37.68
1	16	23.19
2	8	11.59
3	10	14.49
4	5	7.25
5 or more	4	5.8
Whole Time Equivalent to list size ratio mean mean (SD, range) n=55	1864.90(376.4, 928-2800)	

Table 4:Estimates and 95% CI of final model for part 2- the Student Feedback Scores

		Parameter Estimate	95% Confidence Interval
WTE ratio		-0.01	(-0.02, 0.01)
Rurality			
Mixed	--		
Rural		-12.12	(-28.82, 4.58)
Urban		-9.32	(-21.24, 2.60)
No of year teaching			
1	--		
2		-1.75	(-15.73, 12.23)
3		7.39	(-6.20, 20.99)
4		9.37	(-11.23, 29.98)
5		0.68	(-10.19, 11.55)
No of Tutor Development Sessions attended			
0	--		
1		-5.8	(-16.21, 4.61)
2		-0.44	(-12.64, 11.77)
3		-0.2	(-13.12, 12.72)
4		0.83	(-13.80, 15.47)
5		3.93	(-13.72, 21.58)
Enthusiasm			
2	--		
3		7.45	(-5.17, 20.08)
4		8.64	(-4.36, 21.65)
5		10.65	(-1.35, 22.66)
Teach postgraduates			
No	--		
Yes		-5.74	(-13.67 , 2.19)
Constant		89.53	(65.38, 113.67)

Appendix 1 – students' feedback scoring system

Appendix 2 – statistical methods