Author’s response to reviews

Title: Comparative Attainment of 5-Year Undergraduate and 4-Year Graduate Entry Medical Students moving into Foundation Training

Authors:

Gillian Manning (gillian.manning@nottingham.ac.uk)
Paul Garrud (paul.garrud@nottingham.ac.uk)

Version: 2 Date: 6 December 2009

Author’s response to reviews: see over
Anastasios Koutsos, PhD.
Scientific Editor
BMC-series Journals
BioMed Central
Floor 6, 236 Gray's Inn Road
London, WC1X 8HL

7\textsuperscript{th} December 2009

Dear Anastasios,

Re: Comparative Attainment of 5-Year Undergraduate and 4-year Graduate Entry Medical Students Moving into Foundation Training - MS: 109843664315158

Thank you for your email of the 20\textsuperscript{th} November 2009, we are delighted to have the opportunity to respond to the helpful comments and feedback from the referees.

Referee 1

Comment 1: No, the core content and learning outcomes are aligned between the GEM and UG programmes; it is the delivery (PBL in systems-based modules, and systems-based and discipline-based modules, respectively) that varies between the 4 and 5 year medicine programmes. The only other difference, that is already spelled out in the text (Background, 3\textsuperscript{rd} para), is a 5\textsuperscript{th} semester research project and dissertation that leads to a Bachelor of Medical Sciences degree, awarded at the end of year 3 in the 5-year programme.

Comment 2: The reviewer is correct that 14 separate Chi-square analyses of the numbers failing each assessment are likely to inflate the chances of a type 1 error. The Bonferroni correction, though, cannot be applied to a series of non-parametric tests like this. Tukey suggested, many years ago, an approach to adjusting the statistical criterion used (alpha) for multiple comparisons:--

\[ \text{alpha (\%)} = \frac{10}{(k(k-1))} \text{ where } k=\text{number of groups/samples being compared with each other.} \]

Applying this would suggest using \( p=0.005 \) as the statistical criterion. The only statistically significant difference reported (Table 3) has \( p=0.04 \) – marginal even at the \( p=0.05 \) level. We think readers of BMC Medical Education will be able to make an appropriate appraisal of this result.

Comment 3: reliable has been changed to significant

Comment 4: Chi square has been replaced with \( \chi^2 \) in the text.
Referee 2

Comment 4: We have commented on one pattern in our results (see response and re-written section of discussion set out below under Referee 3) – namely the divergence between GEM and UG groups that develops during the clinical phases. We do not think that the nature of the assessments reported here (MCQs, OSCE & OSLER) favour a PBL over a non-PBL approach, or vice versa. Both GEM and UG experience MCQ and OSCE assessments in the first part of their courses (years 1 & 2 for UG, year 1 and 1st semester year 2 for GEM). The only difference that suggests itself is that learning that is case-based (PBL) may, perhaps, enable GEM students to adapt more quickly to the case-based (actual patients) teaching encountered in the clinical phases and contribute to their superior performance on the MCQ exam at that point. This is pure speculation, however, and we do not know of any systematic evidence to support the idea.

Comment 5: In fact there are not lower marks on the OSLER, though there was a marginal (p=0.04) difference in the number of students failing this assessment at first attempt: both figures given in Tables 2 & 3 respectively. Given comment 2 of referee 1 (above), perhaps the likeliest explanation is that this may be a type 1 error and not a genuine difference.

The results of the Community Follow up are not given in the tables and only the text reports that a higher proportion of GEM than UG failed this assessment. There was no significant difference in mean z-score transformed marks on this assessment either. We are unsure what the best approach is on this assessment. It is an unusual assessment that is unlike any other in the shared clinical attachments of the Nottingham courses, comprising a patient-based case study, assessed via a presentation and a written report. Our informal view, based on unsystematic feedback from students in the two cohorts studied for this paper, is that some GEM students, having had three previous pieces of coursework that were similar (2 patient narratives in year 1, and a disability case study in year 2), approached this piece of work carelessly and in particular did not ensure they met the somewhat different assessment criteria. Factor analysis of the clinical phase assessments has shown that all the knowledge-based assessments (MCQs) load heavily on a single factor, the skills-based assessments (OSCE & OSLER) load on both that factor and on a separate factor, but the community follow-up assessment loads on neither factor, but uniquely on a third that accounts for a relatively small amount of variance.

We have therefore omitted the reference to the community follow-up from the text entirely, so that the paper focuses on the core issues of knowledge and clinical competence.

Referee 3

This reviewer argues that we have misinterpreted the results of the study and that the discussion needs to be re-written.

The basic results are:

- significantly more GEMs complete in standard time (p=0.04);
- no significant overall difference in assessment failures (p>0.05);
- a significant overall difference in mean assessment performance (p<0.001)

We have taken the view that the two statistically reliable differences warrant some further exploration and discussion. The first is discussed straightforwardly in the first paragraph of the Discussion. The second led to a series of univariate analyses (as reported in the results) and a pattern of differences between GEM and UG students on the MCQ knowledge-based assessments. The pattern is that
of two groups diverging over time, as they progress through the shared clinical phases. Thus GEM are better than UG in clinical phase 1, then deteriorate and are worse than UG through most of clinical phase 2 and then clinical phase 3 (mean z-scores: 0.17, -0.10, -0.14, -0.15, -0.17, -0.19, -0.19). It is this pattern of developing poorer performance on the knowledge-based assessments that we then discuss (paras 4 and 5 of the Discussion). One possibility is that these differences (GEM getting worse on the knowledge assessments) show up partly as an artefact of the lower drop out from the course – perhaps a slightly higher proportion of GEM students being ‘strugglers’ who generally perform weakly. We do not have any systematic data to confirm or infirm such a suggestion. Stewart Peterson (personal communication) has found that the Leicester-Warwick graduate entry students are underrepresented both in the highest 10% and the lowest 10% of the combined class during their clinical phases – fewer strugglers and fewer high fliers: that might suggest this isn’t happening on the Nottingham course, but we don’t really know at this point. Because the admissions policy for GEM is deliberately broad (eligibility being a 2.2 degree or better), it does seem to be an important research question to see if there is any suggestion that poorer prior educational achievement is related to poorer performance on the medicine programme. In sum, we believe our discussion is warranted and does not misrepresent the results. We have re-written the discussion, however, to make the summary of results and the purpose of the discussion clearer.

Comment 1: The reviewer’s point relates to the pattern of results discussed above. The discussion has been amended at this point to make it clearer.

Comment 2: A reference to “first cohort” effects is now included at this point.

Abstract: Aim is now a separate section of the abstract.

Results para 3: The sentence reports a result from the multivariate analysis of variance. What it means is that the difference between GEM and UG varied between cohorts (ie graduating years). This interaction was clearly produced by the difference reported in the penultimate sentence of that paragraph – that UG performed better than GEM on the OSLER and OSCE at the end of clinical phase 3 in 2007, but it was reversed in 2008 (GEM performed better than UG on those same assessments the following year). A clearer link has been made between the two sentences in that paragraph of the results.

We trust that we have give due consideration to all of the comments from the referees and submit our revised manuscript which we feel is much approved as a result of the review process.

We look forward to seeing our manuscript published.

Yours sincerely,

Dr Paul Garrud
Assistant Director of Medical Education

Dr Gillian Manning
Associate Professor & Director of Postgraduate Research Education