PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The relationship between school type and academic performance at medical school: a national, multi-cohort study.
AUTHORS	Kumwenda, Ben; Cleland, Jennifer; Walker, Kim; Lee, Amanda; Greatrix, Rachel

VERSION 1 - REVIEW

REVIEWER	Paul Tiffin University of York, UK I am co-author of a draft paper soon to be submitted addressing a similar question, though using different predictor and outcome variables.
REVIEW RETURNED	21-Feb-2017

GENERAL COMMENTS	This is a well written paper which makes a contribution to the literature on medical selection. I found the background concise but covering the relevant key previous papers. I have a number of points, which if addressed, feel would improve the manuscript. 1. The distinction between selective/independent and state/non-selective needs to be clarified. For example, I attended a state funded, but selective, grammar school. Schools could be classified using either selective/non-selective or according to their funding. There are advantages and disadvantages to both, but this should be made clear. Also, in my experience (and there is some research to support this) selective state schools often behave similarly to privately funded schools in an educational epidemiology context. For this reason I have classified schools according to selection status rather than funding in my previous research. 2. There is a considerable overlap in performance between state and independent schools. Figure 3 in the 2013 UKCAT12 paper by McManus et al. illustrates this nicely for selective vs non-selective schools. This should be mentioned and perhaps referenced, as it highlights the potential limitation of dichotomizing school types. It also has implications for policy. 3. I think the paper would benefit from a data flow chart so it is clear what the study frame is and where missing data are present. This is usually required for observational studies published in the BMJ but I'm not sure about BMJOpen. 4. The authors correctly highlight the issues with comparing the EPM across cohorts. There is no ideal solution to this but it is a limitation

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	worth noting. The use of the EGA addressed this to some extent though. 5. I wasn't clear how the UCAS tariff points were handled. Previously I have created a metric based on a standardized version of the maximum UCAS points available for an application (e.g. currently 420 for an applicant from England doing A-levels) which allows a crude equating across the UK nations (see Tiffin et al 2012). Other researchers have similar adopted this approach, though it is imperfect and there have been discussions about the 'best' metric of Prior Educational Attainment (PEA). 6. I prefer the term 'multivariable' rather than 'multivariate' when referring to multiple predictors to avoid confusion. In the behavioural sciences 'multivariate' usually is reserved for modelling with multiple dependent variables (e.g. ordinal factor analysis). 7. EPM is a local measure and therefore any clustering effects relating to medical schools may have been mitigated. However, the reason why a statistical method which would accommodate such dependency of observations (e.g. MLM, GEE) within medical schools has not been used (i.e. a single level analysis) should be justified. This could be simply citing the (presumably low) intraclass correlation derived from a variance components analysis. In this case it would refer to the correlation of the latent variable conceptualized as underlying the binary responses modelled by the logistic regression. 8. The discussion rightly highlights some implications for admissions policy, but again should refer to the overlap in school level performance in different school types. The authors should also highlight that any adjustment of advance qualification grade should be therefore made on more granular metrics of a secondary school's performance (such as the average grade that each student attending obtained for A-levels, Highers etc). This of course brings its own challenges, as different nations within the UK will report this differently. It would also be worth noting in the discussion that the advent of more available national performance metrics (e.g. postgrad Royal College exams in UKMED and the UKMLA) will make such modelling easier, as hopefully there will be less reliance on local measures such as the EPM as an outcome. 9. There is a typo in the abstract- I think 'per-entry' should be 'pre-entry' 10. On pg 4 (HFCE) should be (HEFCE) Also- I have not had sight of the STROBE checklist for this draft paper (hence I did not answer 'yes' to this item on the above checklist). I hope these comments and suggestions are helpful.

REVIEWER	Chris McManus University College London, UK
REVIEW RETURNED	01-Mar-2017

GENERAL COMMENTS	This paper uses data from the newly available UKMED, a database

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	of administrative data on medical training, to assess the extent to which type of secondary schooling is related to outcome in final year medical school exams. In so doing it continues a line which a number of other studies have done previously, albeit with a more complete sample of medical schools. The main claim is that, "students from state funded schools are likely to outperform students from independent schools". That is probably correct in part, but whether the study really shows it is unclear at present. Major points. 1. The study looks at different types of schools, and there is immediately an issue of nomenclature. Independent (private) schools are mostly selective, sometimes in terms of academic ability, almost always in terms of financial ability, and state ('public' in the sense of being available to all members of the public) are mostly not selective. There is however a large overlap (and see figure 3 of the UKCAT-12 paper for an illustration of that). The HEFCE 2003 report separates out high and low achieving schools from different types of schools, and there it is clear that independent schools appear to achieve less well, but they nevertheless have higher A-level grades overall (see table 2 of the report). That makes teasing apart selectivity of school and type of school quite difficult. 2. p.3: "But are stronger pre-entry grades an indicator of potential at medical school?". A key issue, and it is said that "the wider literature is conflicting". The only basis for the latter claim is the paper by Thiele et al at the University of Liverpool, which, despite the present authors claiming that, "most of the attainment differences observed between students … either decreased or disappeared by year four of a five-year programme", (p.4), the Thiele et al paper actually says that, " In order to make comparisons between degree programmes and students as fair as possible, students registered on four- and five-year programmes including Veterinary Science, Medicine and Dentistry were excluded from the data set." Something is not right in the paper, and clearly the current description of the study is wrong. 3. The question of how well academic qualifications predict medical school performance is complex, and the authors should perhaps have read the two BMC Med papers accompanying the UKCAT-12 study. The first shows that entry qualifications continue to predict through medical school and on to postgraduate training, albeit with the diminishing effects expected with a 'simplex' data pattern. The second, on construct validity, also takes into account the unreliability and range restriction of the selection measures and the outcome measures, as well as the fact that the least well qualified of entrants never entered medical school, but nevertheless suggests that had they entered they would have underattained (which is why medical schools select). 4. The present study has data from 33 medical schools, which have been analysed in a single analysis. That is probably not an appropriate way to analyse the data since the outcome measure is EPM, which has quartiles/deciles within individual medical schools. Since medical schools differ both in entry qualifications and in demography there is a major risk here of biases. The appropriate method is either multi-level modelling, so that the schools are separate, or separate analyses for each medical school followed by a meta-analysis. Both methods allow answers to a key question of whether effects differ between medical schools (and that may well

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	be the case for multiple reasons). Ironically, Thiele et al particularly emphasise that their data, “exemplify how patterns observed nationally may differ between universities”, while the present paper ignores such issues. 5. The problem for EPM is not present for the UKFPO SJT, and it would probably make sense also to analyse it. Quite what the SJT is measuring is still controversial but it is clear that it correlates highly with the EPM, but there are also substantial differences between medical schools. Whether it is a measure of academic attainment is another matter (or perhaps some intrinsic sense of ‘moral reasoning’ as has been argued), but it in some ways is acting as a proxy for an outcome measure common to all schools and on the same scale. 6. The predictor variables fall into two categories, demographic and academic. The former are not problematic but the latter are. UCAS tariff points are a very dubious measure. It should be made clear that these are ‘old tariff points’ and were recorded before A* grades appeared at A-level. Tariff points are a mish-mash, and are almost totally uninterpretable in students at the top end of the range. For instance, the mean score here is about 480 (SD about 100), with a 95% range of about 280 to 680. An A-level grade A got 120 points back then, so that this is equivalent to an average of four grade As, with a range from 2.3 A grades to 5.7 A grades. Most schools select on three best A-levels, and therefore there is something awry here. More problematic is that General Studies is counted by some schools and not others, etc, etc.. My guess is that the 5+ A-level grades mostly reflect maths and its variants, or perhaps a couple of AS-levels (also in the formula if not taken at A-level). For all such reasons the UKCAT-12 paper carefully avoided tariff scores, and instead calculated a range of special purpose measures (such as three best A-levels, etc). All of that data is in UKMED, and can be extracted from the HESA data, and Daniel Smith tells me that the SPSS syntax is available in the safe haven. I can see no reason at all therefore for using tariff scores. 7. Perhaps the saving grace for the tariff scores is that, despite implications otherwise, actually they are *not* used in the present study, which is a shame since the main claim is that, “all things being equal [presumably including academic attainment measures] medical students from state (non-selective) schools do better at medical school …”. However the analysis of table 3 suggests that only demographic and not attainment measures have been included in the analysis. Clearly something is not right here, as I presume what is required is an analysis which is the equivalent of figure 3 in the UKCAT-12 paper. As it is, it is difficult to interpret anything here. 8. To summarise, there is no evidence that academic attainment predicts outcomes because no multivariate analyses have been carried out of the relationship between entry measures and demographics in relation to outcome measures. That is hardly evidence of a lack of a relationship. The only evidence is very indirect, based on private sector students having a) higher performance on UKCAT and b) lower performance on EPM. That may be correct but it is hardly compelling. Minor points. 9. p.3: “The medical profession has always been socially exclusive”. Hmm – in ancient Greece, in 13th century England, or wherever

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	else? It has, I agree, for the past half century (most clearly since the Todd Report data of 1968) been clear that a majority of entrants to medical school come from higher socio-economic groups, but that does not necessarily mean the institutions are exclusive in the sense of a Swiss Finishing School. Mental ability on tests correlates well with social status (perhaps not surprisingly since high ability results in upward social mobility), and in almost all studies SEG and school performance are correlated highly (and see McManus, 1982, Med Ed, 16: 72) for a discussion of the extent to the social class gradient is explained by such differences. It may be that the authors don't want to consider other explanations of social class gradients, which are very common, but that doesn't mean they should ignore those possible explanations. 10. p.3: "Only 7% of the UK population goes to independent schools yet across the years a pattern of more than 20% of medical students coming from such schools"… If the population were allocated randomly to state and independent schools then that would be compelling evidence for something. However, as stated, students in independent schools do better on school leaving examinations (and that may not be due to better education but to different ability at intake). The arguments here and elsewhere need to be more clearly presented, rather than merely asserted as obvious. 11. The references to official reports on the web should have URLs (e.g. refs 19, 20, 22, etc).. 12. p.5. "an analysis of anonymised secondary data". Presumably this should be, "a secondary analysis of anonymised data". 13. I wasn't convinced by the argument for using just the extreme 20/25% and comparing them. At the least there should be three groups (0-20, 20-80 and 80-100) but really it makes little sense to have a fine-grained measure (EPM deciles) and then throw away 50% or 60% of the data. That also removes any possibility of looking for non-linearities in the data. Personally I would convert the deciles/quartiles to normal scores (SPSS does that easily) and then treat those as normal and use regression in a multi-level model. 14. p.7: "p>0.1" – presumably <.1. 15. p.11: "An unrepresentative proportion of medical students come from independent schools". What would be a representative proportion and how would it be calculated? Perhaps the authors could give some illustrative figures. Overall, the present dataset has the potential to answer some important questions, even given the general inadequacies of the EPM which is a within-school measure, but at present the use of the UCAS tariff and the lack of proper multivariate analyses, including multi-level modelling, make it hard to know what is going on here. Some of those same problems are lurking within the UKCAT-12 paper (where again the outcome measures are within school) but if that analysis were repeated now it would almost certainly use multi-level modelling. I think I can only ask for a major revision.

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

VERSION 1 – AUTHOR RESPONSE

Manuscript ID bmjopen-2017-016291 entitled "The relationship between school type and academic performance at medical school: a national, multi-cohort study."		
OVERALL EVALUATION AND GENERAL COMMENTS		
Reviewer: 1	**Reviewer 2**	
This is a well written paper which makes a contribution to the literature on medical selection. I found the background concise but covering the relevant key previous papers. Institution and Country: University of York, UK Competing Interests: I am co-author of a draft paper soon to be submitted addressing a similar question, though using different predictor and outcome variables.	This paper uses data from the newly available UKMED, a database of administrative data on medical training, to assess the extent to which type of secondary schooling is related to outcome in final year medical school exams. In so doing it continues a line which a number of other studies have done previously, albeit with a more complete sample of medical schools. The main claim is that, "students from state funded schools are likely to outperform students from independent schools". That is probably correct in part, but whether the study really shows it is unclear at present.	General comment. No action required.
TITLE AND ABSTRACT		
1.9. There is a typo in the abstract- I think 'per-entry' should be 'pre-entry'		Typo corrected. The word now reads 'pre-entry'
INTRODUCTION AND CONCEPTUAL FRAMEWORK (Q3)		
1.1. The distinction between selective/independent and state/non-selective needs to be clarified. For example, I attended a state funded, but selective, grammar school. Schools could be classified using either selective/non-selective or according to their funding. There are advantages and disadvantages to both, but this should be made clear. Also, in my experience (and there is some research to support this) selective state schools often behave similarly to privately	**2.1**. The study looks at different types of schools, and there is immediately an issue of nomenclature. Independent (private) schools are mostly selective, sometimes in terms of academic ability, almost always in terms of financial ability, and state ('public' in the sense of being available to all members of the public) are mostly not selective. There is however a large overlap (and see figure 3 of the UKCAT-12 paper for an illustration of that). The HEFCE 2003 report separates out high and low achieving schools from different types of schools, and there it is clear that independent	We have added a new paragraph to clarify the difference: For purposes of this study, the type of school attended was defined according to funding criteria, whether state-funded or privately funded. Here in the UK, state-funded education is free and mostly non-selective, while privately funded schools are selective, they require pupils to pay tuition fees, and have a

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

funded schools in an educational epidemiology context. For this reason I have classified schools according to selection status rather than funding in my previous research.	schools appear to achieve less well, but they nevertheless have higher A-level grades overall (see table 2 of the report). That makes teasing apart selectivity of school and type of school quite difficult.	greater proportion of pupils from affluent backgrounds. For this reason, we have classified privately funded schools as independent. We acknowledge the potential limitation of dichotomising schools by funding source as there is an overlap of some schools, for example grammar schools in England, which are both state-funded and selective. Such schools may even share the academic characteristics of independent schools[22] but this categorisation was deemed appropriate for analysis on the basis of our knowledge of how this has been approached in studies from the wider field of education.[pg. 6] ---------- Also, a careful consideration has been made not to generalise that all state funded schools are non-selective. We have acknowledged that state schools also include grammar schools in England (which are selective)
1.2. There is a considerable overlap in performance between state and independent schools. Figure 3 in the 2013 UKCAT12 paper by McManus et al. illustrates this nicely for selective vs non-selective schools. This should be mentioned and perhaps referenced, as it highlights the potential limitation of dichotomizing school types. It also has implications for policy.		Please see the response above.

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	2.2. p.3: “But are stronger pre-entry grades an indicator of potential at medical school?”. A key issue, and it is said that “the wider literature is conflicting”. The only basis for the latter claim is the paper by Thiele et al at the University of Liverpool, which, despite the present authors claiming that, “most of the attainment differences observed between students … either decreased or disappeared by year four of a five-year programme”, (p.4), the Thiele et al paper actually says that, “ In order to make comparisons between degree programmes and students as fair as possible, students registered on four- and five-year programmes including Veterinary Science, Medicine and Dentistry were excluded from the data set.” Something is not right in the paper, and clearly the current description of the study is wrong.	We had Thiele’s reference mixed up with her other article that was published around the same period. The correct reference reads: 23. Thiele T, Pope D, Singleton A, Stanistreet D. Role of students' context in predicting academic performance at a medical school: a retrospective cohort study. BMJ Open. 2016;6(3):11. The authors report that… ‘most attainment differences observed between students either decreased or disappeared during university.’ Our work extends this conversation by looking at a nationally recognised indicator of overall performance, and not just the 4th year result.
	2.3. The question of how well academic qualifications predict medical school performance is complex, and the authors should perhaps have read the two BMC Med papers accompanying the UKCAT-12 study. The first shows that entry qualifications continue to predict through medical school and on to postgraduate training, albeit with the diminishing effects expected with a ‘simplex’ data pattern. The second, on construct validity, also takes into account the unreliability and range restriction of the selection measures and the outcome measures, as well as the fact that the least well qualified of entrants never entered medical school, but nevertheless suggests that had they entered they would have underattained (which is	We have read these papers and thank the reviewer for this guidance.

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	why medical schools select).	
	2.9. p.3: "The medical profession has always been socially exclusive". Hmm – in ancient Greece, in 13th century England, or wherever else? It has, I agree, for the past half century (most clearly since the Todd Report data of 1968) been clear that a majority of entrants to medical school come from higher socio-economic groups, but that does not necessarily mean the institutions are exclusive in the sense of a Swiss Finishing School. Mental ability on tests correlates well with social status (perhaps not surprisingly since high ability results in upward social mobility), and in almost all studies SEG and school performance are correlated highly (and see McManus, 1982, Med Ed, 16: 72) for a discussion of the extent to the social class gradient is explained by such differences. It may be that the authors don't want to consider other explanations of social class gradients, which are very common, but that doesn't mean they should ignore those possible explanations.	We have reframed the opening paragraph of the introduction to acknowledge that the issue is more complicated than school performance alone: 'While the demand for access to the profession remains high, this unbalanced pattern is often attributed to the differences in prior academic achievement. For example, students from disadvantaged backgrounds are reportedly more likely to obtain lower school grades than students from more affluent backgrounds.[1-3] However, the issue of social origins and access to the professions is more complex than the differences in academic attainment alone.[4,5] [Pg3].
	2.10. p.3: "Only 7% of the UK population goes to independent schools yet across the years a pattern of more than 20% of medical students coming from such schools"… If the population were allocated randomly to state and independent schools then that would be compelling evidence for something. However, as stated, students in independent schools do better on school leaving examinations (and that may not be due to better education but to different ability at intake). The arguments here and elsewhere need to be more clearly presented, rather than merely asserted as obvious.	Thanks for this comment. As indicated in the qualifying statement we have stated that students from independent schools do better on school leaving examinations, which we believe is sufficient explanation at this point in the paper.
METHODICAL RIGOUR		

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

1.4. The authors correctly highlight the issues with comparing the EPM across cohorts. There is no ideal solution to this but it is a limitation worth noting. The use of the EGA addressed this to some extent though. --------- **1.7.** EPM is a local measure and therefore any clustering effects relating to medical schools may have been mitigated. However, the reason why a statistical method which would accommodate such dependency of observations (e.g. MLM, GEE) within medical schools has not been used (i.e. a single level analysis) should be justified. This could be simply citing the (presumably low) intraclass correlation derived from a variance components analysis. In this case it would refer to the correlation of the latent variable conceptualized as underlying the binary responses modelled by the logistic regression.	**2.13**. I wasn't convinced by the argument for using just the extreme 20/25% and comparing them. At the least there should be three groups (0-20, 20-80 and 80-100) but really it makes little sense to have a fine-grained measure (EPM deciles) and then throw away 50% or 60% of the data. That also removes any possibility of looking for non-linearities in the data. Personally I would convert the deciles/quartiles to normal scores (SPSS does that easily) and then treat those as normal and use regression in a multi-level model.	Instead of a simple binary logistics regression we have now instituted a mixed-effect logistics model with a random intercept for medical school [pg 8-9]. ---------- We re-fitted the models with all the EPM quartiles (2012) and EMP quintile (2013) scores treating them as ordinal categorical outcomes. We can confirm that the new finding complements the earlier findings. Detailed results of mixed-effect ordinal logistic regression are presented as supplementary tables. However, for purposes of this article, we stand by our contention to report the 'Extreme Group', on the basis that this approach helps contextualise the groups that are of interest to our research. Besides, the approach has not compromised the statistical power. That said, we agree with reviewer 1's assertion that there is no ideal solution to this but it is a limitation worth noting (which we have done in the discussion section, pg13). --------- The mixed-effects ordinal logistic analyses confirm everything that we found before. The odds ratios appear closer to 1, whether they were above 1 or below 1 in the high group vs low group analyses. Hence the advantage for some groups (female, state-funded) appear smaller, and the disadvantage of other groups (non-white)

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

		appear smaller too, but the differences remain significant.
	2.4. The present study has data from 33 medical schools, which have been analysed in a single analysis. That is probably not an appropriate way to analyse the data since the outcome measure is EPM, which has quartiles/deciles within individual medical schools. Since medical schools differ both in entry qualifications and in demography there is a major risk here of biases. The appropriate method is either multi-level modelling, so that the schools are separate, or separate analyses for each medical school followed by a meta-analysis. Both methods allow answers to a key question of whether effects differ between medical schools (and that may well be the case for multiple reasons). Ironically, Thiele et al particularly emphasise that their data, "exemplify how patterns observed nationally may differ between universities", while the present paper ignores such issues.	We have taken this comment into account and have since used a mixed-effects logistic regression model with a random intercept for medical school. This helped to control for the clustering effect of individual students within medical schools.
	2.5. The problem for EPM is not present for the UKFPO SJT, and it would probably make sense also to analyse it. Quite what the SJT is measuring is still controversial but it is clear that it correlates highly with the EPM, but there are also substantial differences between medical schools. Whether it is a measure of academic attainment is another matter (or perhaps some intrinsic sense of 'moral reasoning' as has been argued), but it in some ways is acting as a proxy for an outcome measure common to all	While the EPM is used in conjunction with SJTs in postgraduate selection, and despite some level of interaction between the two outcomes, we do not wish to create the impression that one is a surrogate of the other. We agree that examining the same question for the SJT is important, but this is a different study.

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

<table>
<tr><td></td><td>schools and on the same scale.</td><td></td></tr>
<tr><td>1.5. I wasn’t clear how the UCAS tariff points were handled. Previously I have created a metric based on a standardized version of the maximum UCAS points available for an application (e.g. currently 420 for an applicant from England doing A-levels) which allows a crude equating across the UK nations (see Tiffin et al 2012). Other researchers have similar adopted this approach, though it is imperfect and there have been discussions about the ‘best’ metric of Prior Educational Attainment (PEA).</td><td>2.6. The predictor variables fall into two categories, demographic and academic. The former are not problematic but the latter are. UCAS tariff points are a very dubious measure. It should be made clear that these are ‘old tariff points’ and were recorded before A* grades appeared at A-level. Tariff points are a mish-mash, and are almost totally uninterpretable in students at the top end of the range. For instance, the mean score here is about 480 (SD about 100), with a 95% range of about 280 to 680. An A-level grade A got 120 points back then, so that this is equivalent to an average of four grade As, with a range from 2.3 A grades to 5.7 A grades. Most schools select on three best A-levels, and therefore there is something awry here. More problematic is that General Studies is counted by some schools and not others, etc, etc.. My guess is that the 5+ A-level grades mostly reflect maths and its variants, or perhaps a couple of AS-levels (also in the formula if not taken at A-level). For all such reasons the UKCAT-12 paper carefully avoided tariff scores, and instead calculated a range of special purpose measures (such as three best A-levels, etc). All of that data is in UKMED, and can be extracted from the HESA data, and Daniel Smith tells me that the SPSS syntax is available in the safe haven. I can see no reason at all therefore for using tariff scores.</td><td>This is the first study to use the UKMED and communication from the data supplier indicates that the UCAS tariff score was indeed calculated differently by UCAS (prior to 2007/8) and HESA (2008 onwards). We could not access the individual subject grades in time to allow us manipulate the scores to the specified level of standardization. However, to address the reviewer’s concern, we have carried out a sub-group analysis using the DoE average point score per A level student. This is a standardised tariff available for schools in England. As expected, the trend for pre-entry grades remain the same, with state schooled pupils coming into medical school with substantially lower school grades than privately educated pupils.

To control for the school performance effect, we capped the UCAS tariff points into an average point score per A level student (relative to the average for their school in the exam year).

<table><tr><td>Grade</td><td>DoE (DFE)</td><td>UCAS</td></tr></table></td></tr>
</table>

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

<table>
<tr>
<td></td>
<td></td>
<td>
<table>
<tr><td></td><td>points</td><td>Tariff</td></tr>
<tr><td>A*</td><td>300</td><td>140</td></tr>
<tr><td>A</td><td>270</td><td>120</td></tr>
<tr><td>B</td><td>240</td><td>100</td></tr>
<tr><td>C</td><td>210</td><td>80</td></tr>
<tr><td>D</td><td>180</td><td>60</td></tr>
<tr><td>E</td><td>150</td><td>40</td></tr>
<tr><td>U</td><td>0</td><td>0</td></tr>
</table>
The adjusted UCAS tariff was thus calculated as:

=Average point score per A level student*(2/3))-60.
</td>
</tr>
<tr>
<td></td>
<td>2.7. Perhaps the saving grace for the tariff scores is that, despite implications otherwise, actually they are *not* used in the present study, which is a shame since the main claim is that, “all things being equal [presumably including academic attainment measures] medical students from state (non-selective) schools do better at medical school …”. However the analysis of table 3 suggests that only demographic and not attainment measures have been included in the analysis. Clearly something is not right here, as I presume what is required is an analysis which is the equivalent of figure 3 in the UKCAT-12 paper. As it is, it is difficult to interpret anything here.</td>
<td>Our inclusion of the UCAS tariff score (alongside the UKCAT and GAMSAT) is simply to illustrate that students from state schools come into medical school with substantially lower school grades.</td>
</tr>
<tr>
<td colspan="2">RESULTS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.14. p.7: “p>0.1” – presumably <.1.</td>
<td>Conservative p-value corrected to ‘p<0.1’.</td>
</tr>
<tr>
<td colspan="2">DISCUSSION AND CONCLUSIONS (Q6)</td>
<td></td>
</tr>
</table>

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

1.8. The discussion rightly highlights some implications for admissions policy, but again should refer to the overlap in school level performance in different school types. The authors should also highlight that any adjustment of advance qualification grade should be therefore made on more granular metrics of a secondary school's performance (such as the average grade that each student attending obtained for A-levels, Highers etc). This of course brings its own challenges, as different nations within the UK will report this differently. It would also be worth noting in the discussion that the advent of more available national performance metrics (e.g. postgrad Royal College exams in UKMED and the UKMLA) will make such modelling easier, as hopefully there will be less reliance on local measures such as the EPM as an outcome.		We decided to highlight the potential limitation of classifying school by funding type in the data description section on page 6. --- The use of a range of special purpose measures such as DfE average point score per A level student We have added text in the discussion section that reads: 'Furthermore, the localised adjustment of some of the cognitive competencies has implications for more sophisticated statistical models and this may influence the generalisability of the results. For example, it was not possible to extrapolate the average point score per A level student because A levels are (mostly) restricted to English schools. As noted earlier, the EPM is directly measured by the individual medical schools, although we tried to minimise this variability by focusing on the extreme groups and deploying a mixed-effects logistic regression technique, there is no agreed approach for doing so. Future research should be less reliant on these local measures and instead focus on more national performance metrics like the Royal College postgraduate examinations or the proposed UK Medical Licensing Assessment (http://www.gmc-uk.org/education/29000.asp).[Pg 12]

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	2.8. To summarise, there is no evidence that academic attainment predicts outcomes because no multivariate analyses have been carried out of the relationship between entry measures and demographics in relation to outcome measures. That is hardly evidence of a lack of a relationship. The only evidence is very indirect, based on private sector students having a) higher performance on UKCAT and b) lower performance on EPM. That may be correct but it is hardly compelling.	The decision to use entry measures (Table 1) is mainly to help confirm that the sample is a true reflection of the student population, and that students from state schools do indeed come to medical school with lower school grades and perform less well at the UKCAT. As indicated in the introduction section, our aim is to evaluate the predictive validity of a range of socio-demographic factors, including school type, in relation to EPM.
	2.15. p.11: "An unrepresentative proportion of medical students come from independent schools". What would be a representative proportion and how would it be calculated? Perhaps the authors could give some illustrative figures.	We have edited the sentence to further clarify the proportions, and it now reads: 'Around 80% of applications come from independent and selective schools[53], yet medical students from state schools do better than those from independent schools at medical school.
	Overall, the present dataset has the potential to answer some important questions, even given the general inadequacies of the EPM which is a within-school measure, but at present the use of the UCAS tariff and the lack of proper multivariate analyses, including multi-level modelling, make it hard to know what is going on here. Some of those same problems are lurking within the UKCAT-12 paper (where again the outcome measures are within school) but if that analysis were repeated now it would almost certainly use multi-level modelling. I think I can only ask for a major revision.	We have duly revised the statistical analysis and results sections to incorporate the recommended changes
CLARITY		

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

1.3. I think the paper would benefit from a data flow chart so it is clear what the study frame is and where missing data are present. This is usually required for observational studies published in the BMJ but I'm not sure about BMJOpen.		We have added a flow chart to the supplementary section [pg 24]. The diagram shows the number of cases used at each stage of the analysis.
1.6. I prefer the term 'multivariable' rather than 'multivariate' when referring to multiple predictors to avoid confusion. In the behavioural sciences 'multivariate' usually is reserved for modelling with multiple dependent variables (e.g. ordinal factor analysis).		We have now deleted the term 'multivariate' and replaced it with 'multivariable' – pg7.
1.10. On pg 4 (HFCE) should be (HEFCE)		Correction done
1.11. Also- I have not had sight of the STROBE checklist for this draft paper (hence I did not answer 'yes' to this item on the above checklist).		STROBE checklist now completed.
	2.11. The references to official reports on the web should have URLs (e.g. refs 19, 20, 22, etc)..	The references have been edited
	2.12. p.5. "an analysis of anonymised secondary data". Presumably this should be, "a secondary analysis of anonymised data".	We have done the requested corrected.

VERSION 2 – REVIEW

REVIEWER	Paul Tiffin University of York I have a similar study under review at present with another journal.
REVIEW RETURNED	12-May-2017

GENERAL COMMENTS	The paper has been significantly improved by the revisions. I have a few further suggested minor corrections however: 1. On page 3 'weakness' should perhaps be 'weaknesses' 2. page 6. "Data on age, gender, ethnicity and whether the individual was a school leaver or a graduate at the time of entry to medical school was also included. "should be "Data on age, gender, ethnicity

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

	and whether the individual was a school leaver or a graduate at the time of entry to medical school were also included." as 'data' are plural. 3. "but this categorisation was deemed appropriate for analysis on the basis of our knowledge of how this has been approached in studies from the wider field of education. " page 6. Actually previous educational research highlights that selective state schools (e.g. grammars) behave as predictors quite often like independent schools. Thus, this is not a good justification. It may be better to justify it on the basis of potential implications for admissions policy (where the state private distinction is used commonly)? 4. My understanding is that the cohorts with EPM quintiles and quartiles were not directly comparable when EPM- as an outcome- was used this way. Thus, ordinal logistic regression could have been used separately for each cohort but not for the pooled sample. I believe this to be the case as the change in odds for a change in each step in the scales cannot be assumed to be equal (e.g. the step from 4th to 5th quintile may not relate to the step between the 3rd to 4th quartile in the other cohort). Also- the interpretation of the ORs from the ordinal regression could be made clearer- the resulting ORs are the odds ratio for being in a higher versus a lower category (quantile in this case) (e.g. 3 vs 2 or 2 vs 1) given the value of the predictor variable. Thus, naturally the ORs will be nearer to 1 compared to those recovered from a binary logit model as usually the outcome metric is smaller. However, the SEs should also be smaller in most cases because there is information contained in the ranked order of an ordinal outcome. I also noted the other reviewer's comments that results that make some adjustment for previous educational performance should be included. I tend to agree with this, as I would anticipate it may highlight the differences in performance between state and private students (i.e. high grades obtained from a high performing (often private) school are 'worth' less than those gained in a more challenging secondary school environment. I hope these comments are helpful in producing what I hope would be a finalised version of this interesting paper.

VERSION 2 – AUTHOR RESPONSE

Reviewer's comment	**Author response**
1. On page 3 'weakness' should perhaps be 'weaknesses'	We have corrected this; the text in the manuscript now reads 'weaknesses'
2. page 6. "Data on age, gender, ethnicity and whether the individual was a school leaver or a graduate at the time of entry to medical school was also included. "should be "Data on age, gender, ethnicity and whether the individual was a school leaver or a graduate at the time of entry to medical school were also included." as 'data' are plural.	We have corrected this; the text in the manuscript now reads 'were'
3. "but this categorisation was deemed appropriate for analysis on the basis of our knowledge of how this has been approached in studies from the wider field of education. " page 6. Actually previous educational research highlights that selective state	We have corrected this; the text in the manuscript now reads: 'such schools may even share the academic characteristics of independent schools[22] but this categorisation

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

schools (e.g. grammars) behave as predictors quite often like independent schools. Thus, this is not a good justification. It may be better to justify it on the basis of potential implications for admissions policy (where the state private distinction is used commonly)?	was deemed appropriate for the analysis because it is a commonly used contextual factor in admissions practice and thus has potential implications for admissions policy. (pg6)
4. My understanding is that the cohorts with EPM quintiles and quartiles were not directly comparable when EPM- as an outcome- was used this way. Thus, ordinal logistic regression could have been used separately for each cohort but not for the pooled sample. I believe this to be the case as the change in odds for a change in each step in the scales cannot be assumed to be equal (e.g. the step from 4th to 5th quintile may not relate to the step between the 3rd to 4th quartile in the other cohort). Also- the interpretation of the ORs from the ordinal regression could be made clearer- the resulting ORs are the odds ratio for being in a higher versus a lower category (quantile in this case) (e.g. 3 vs 2 or 2 vs 1) given the value of the predictor variable. Thus, naturally the ORs will be nearer to 1 compared to those recovered from a binary logit model as usually the outcome metric is smaller. However, the SEs should also be smaller in most cases because there is information contained in the ranked order of an ordinal outcome.	Yes, these points about separating the cohorts are correct and were presented in the paper. The EPM quintiles (2012) and quartiles (2013) are not directly comparable and these were analysed separately (see table 3, pg16) using a mixed effect binary logistic regression. We also deployed a mixed-effects ordinal logistic regression (with the 2012 and 2013 cohorts separated) which also confirmed the earlier findings. We have updated the text in on pg 10 to help clarify this. As the reviewer states, the odds ratios were closer to 1 for the ordinal logistic regression (see supplementary file, table 1 and supplementary file, table 2) than for the binary logistic regression and the standard errors were generally smaller. We presented the results in the main article on page 10, but the full tables of the results were presented in the supplementary file.
I also noted the other reviewer's comments that results that make some adjustment for previous educational performance should be included. I tend to agree with this, as I would anticipate it may highlight the differences in performance between state and private students (i.e. high grades obtained from a high performing (often private) school are 'worth' less than those gained in a more challenging secondary school environment.	The decision to use previous educational performance (UCAS points) was simply to illustrate the point that students from state-funded schools come into medical school with relatively weaker grades as compared to peers from privately funded schools. However, we did not include this in the regression model due to lack of comparability of UCAS tariff points across countries. Specific subject qualification (with exam dates) were not available to the research team in time for a more nuanced calculation of education performance. We have added a short section in the discussion (pg12) to reflect upon reviewer's point.

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

VERSION 3 – REVIEW

REVIEWER	Paul Tiffin University of York UK
REVIEW RETURNED	07-Jul-2017

GENERAL COMMENTS	The further revisions seem to have addressed previous issues raised by reviewers.

BMJ Open: first published as 10.1136/bmjopen-2017-016291 on 31 August 2017. Downloaded from http://bmjopen.bmj.com/ on November 18, 2019 by guest. Protected by copyright.

Correction: *The relationship between school type and academic performance at medical school: a national, multi-cohort study*

Kumwenda B, Cleland JA, Walker K, *et al*. The relationship between school type and academic performance at medical school: a national, multi-cohort study. *BMJ Open* 2017;7:e016291. doi: 10.1136/bmjopen-2017-016291

The paper should have the following Acknowledgements section:

Acknowledgements Data Source: - UK Medical Education Database ("UKMED") UKMEDP26 extract generated on 12/08/2016. Approved for publication on 27/03/2017. We are grateful to UKMED for the use of these data. However, UKMED bears no responsibility for their analysis or interpretation. The data includes information derived from that collected by the Higher Education Statistics Agency Limited ("HESA") and provided to the GMC ("HESA Data"). Source: HESA Student Record 2007/2008 and 2008/2009 Copyright Higher Education Statistics Agency Limited. The Higher Education Statistics Agency Limited makes no warranty as to the accuracy of the HESA Data, cannot accept responsibility for any inferences or conclusions derived by third parties from data or other information supplied by it.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See:http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

BMJ Open 2017;**8**:e016291corr1. doi:10.1136/bmjopen-2017-016291corr1