

Annual Speciality Reports (ASRs) Exam Data Submission:

Proposal for exam data to be supplied at individual trainee level

Issue

1. Amending the collection of royal college examination data, collected as part of the Annual Specialty Report (ASR), to include results at trainee level and the GMC number of the trainee.

Further information

2. Anthony Rimmer 020 7189 5258 Arimmer@gmc-uk.org

Background to Annual Specialty Reports

3. The ASR is the mechanism by which medical royal colleges and faculties provide assurance to the GMC and public that they are driving the quality of training and the delivery of curricula in line with the GMC's *Standards for Curricula and Assessment Systems* and *The Trainee Doctor*.

4. ASRs are a concise summary of medical royal colleges'/faculties' information on education and training in a particular specialty, (from here on referred to as colleges). The ASR focuses on the intelligence and data collected from the colleges' role in improving the quality of training for doctors, for example exams pass rates, e-portfolio data on supervision and workplace-based assessments, and feedback from acting as external members to deanery processes. As experts and innovators in their specialties, colleges provide a vital perspective and the ASRs ensure that their views are embedded within the regulation of medical education and training.

5. Colleges submitted their fourth annual report in December 2011. The current ASR template has two elements:

- a. **The reporting template** – used for reporting exceptions in annual college activity and updating the GMC on existing issues.
- b. **The exam data template** – used for reporting aggregated exam results. A separate sheet is used for each examination.

6. The focus of this paper is the examinations data template. Currently, the exam data template collects aggregated data in four sections:

- a. The number of trainees who sat the examination from each deanery, aggregated separately by the following characteristics:
 - i. Candidate Gender
 - ii. Candidate Ethnicity
 - iii. Candidate region of primary medical qualification (PMQ) (UK, EEA, Rest of the world)
- b. The number of trainees who passed the examination from each deanery aggregated by the same characteristics listed above.
- c. The number of trainees who passed or failed the examination from each deanery, aggregated by training level.
- d. The number of trainees from each deanery who have taken the examination who have failed the same examination at least three times. This information is aggregated by the characteristics listed above.

Issues with the current Exam Data Template

7. The aggregated data provides a high level overview of passes and fails. It is not possible to check that the summaries from each college use the same underlying assumptions in compiling the report. For example, it was not possible to verify that the information provided only related to candidates who are training on a GMC approved training programme in the UK (including LAT and fixed term specialty trainees) or to verify how trainees who took the same exam twice in the same reporting period were treated.
8. It is also not possible to perform multivariate analysis of the aggregated datasets. As there are no linkages between the different categories it is impossible to determine whether trainees from a certain ethnic background are from a certain PMQ category or from a certain training grade.
9. We are unsure what methods colleges use to collect data for trainees by deanery i.e. how colleges know whether someone applying for an exam is actually in a training programme and which deanery the trainees are working in. If this information is candidate entered it will not be completely reliable.
10. With regard to the current breakdown of exam data by PMQ, it is not possible to determine at which UK medical school trainees studied. Exam data is important for our work with medical schools but we currently do not have this level of data for analysis purposes. This proposal will allow us to report against all demographic data held by the GMC, some of which is not held by colleges for all candidates, for example the GMC hold PMQ for all trainees, but this is missing for some candidates in the college returns.
11. As seen above with "Unknown" PMQ, there is also a gap in collecting data by trainee level. Again we are unsure how colleges collect data regarding trainee grades and how they maintain accuracy. With this proposal however, we can map exam candidates to the data we already hold on trainees to draw in the most accurate level of training to ensure clean data are used for reporting.
12. By collecting data by GMC number colleges are only required to provide the exam data. Colleges will not need to maintain and provide data on trainee demographics, deanery and so forth for this report as they can be obtained from existing data held by the GMC. This will remove issues of missing data that currently exist.
13. In summary, in order to produce coherent reports based on error free data, the current exam template and the data submitted do not provide enough reach or depth of data to provide an insightful view of the state of medical training as seen through the eyes of exam data.
14. For aggregated exam data we received for the period 2010/11, we produced and published an Exam Data Summary 2010/11 which highlights the extent to which we can analyse exam data. The report can be found here (http://www.gmc-uk.org/Exam_Data_Summary.pdf_49797108.pdf)

Data Collection

15. In view of the weaknesses identified in the previous section, the current form of the exam data would be replaced with a simple extraction of raw data regarding candidates sitting exams. We assume that the data we would request in this proposal will be collected as standard by the colleges when taking applications for exams. As such the data should not require manipulation post extraction. We would request data in an Excel spreadsheet where one line of data pertains to one candidate attempting the exam with the following data per candidate given:

Field	Format	Notes
GMC Number	String of 7 characters	
First Name	String	} Used to validate GMC number in the event of transcription errors
Surname	String	
Examination Name	String	
Date result awarded	MM/YYYY	
Examination Result	Pass/Fail	
Overall Mark achieved	Integer	
Marks for skill set– e.g. for practical exams where skills such as Physical Examination, Clinical Judgment and Communication are examined.	Integer	

16. As well as the above data, we would request information regarding exam formats, pass marks, etc. and necessary notifications of changes for exams e.g. if a marking scheme changes or curriculum updates that effect exams.

17. We request the GMC number as from this unique identifier, we will be able to link with other data we currently store on our own systems. The first name and surname would be used as a quality check to determine that the GMC number and the names match with the register. Based on further investigation with thanks to ICBSE and RCSENG however, we have identified a number of issues with data sharing, for example:

a. Candidates who may not have had a GMC number (oversees candidates or current medical students) but have taken the first part of a qualification and prior to taking subsequent parts have since registered with the GMC. Only candidates who take further exams with the college after obtaining a GMC number will have this unique identifier added to their record. Doctors who have no further contact will not have GMC numbers retrospectively added to their record.

b. Candidates who applied for an exam whilst at x training grade in y deanery, to have moved deanery and progressed to the next level of training when actually sitting the exam.

18. We collect, validate and maintain accurate trainee data via the National Training Survey (NTS), Annual Review of Competence Progression (ARCP) and recruitment data. As such we are offering to work with each college to “clean up” candidate demographic data in order that colleges can provide us with GMC numbers and candidate names and for their own purposes have accurate gender, PMQ, training level and deanery data. Since we currently do this with trainee data from deaneries, through this proposal we will be establishing a greater level of accuracy for trainee data across all stakeholders involved in medical training in the UK. This will also greatly enhance our quality assurance activity.

19. With the initial acceptance of this proposal, we will develop the process with colleges for cleaning up trainee demographic data.

20. The rest of the data as per the fields explained in the above table (15.) would include the lowest common denominator of data we expect colleges to collect when running an exam. When the proposal has been agreed and established, we will continue to review the system for enhancements.

21. **NB:** Prior to colleges being able to submit exam data, colleges will need to officially notify candidates that their data will be given to the GMC. If a privacy notice is not already in place with regard to sharing candidate data with us, we can work with colleges to put one in place. If privacy notices are in place, we will request retrospective data over the last three years. Please see Annex for details.

Reporting

22. The purpose of this proposal is to incorporate exam data in combination with ARCP and recruitment data to better understand training pathways as well as identify any patterns in trainee progression through reporting. Broadly the benefits to our reporting on exam data will be:

- a. More sophisticated statistical analyses, e.g. multivariate analyses of trainee demographics against exam success.
- b. Web-based reporting tool reports for stakeholders to see the results for their group of trainees relative to other groups.
- c. Clustering of reports so that exam data, ARCP and where applicable (i.e. uncoupled specialties) recruitment to higher training data can be presented side by side for an overall view of a programme.
- d. New reports such as exam success by medical school made possible by using GMC number to link to existing data.
- e. Anonymised version of linked data could be made available to academic researchers

23. We will not publish any reports identifying individual exam candidates. The results will only display aggregate data of three or more candidates, similar to the reporting of the National Training Survey (NTS) results.
24. We can build a picture of the trainee's situation using exam data in combination with ARCP outcomes and NTS comments. The combination of these sources of evidence will provide greater insight into the state of medical training in the UK while highlighting areas of concern and areas of good practice.
25. By holding data at the event level (exam, ARCP outcome, etc.), we can produce analyses looking at the interaction of all trainee variables and their association with outcome, for example to what extent do gender, PMQ country and ethnicity predict outcome. Feedback during the QA Review process has indicated appetite for the GMC to improve tracking of training programme outcomes at all stages and to share analysis with all those involved in training. This requires the use of a single identifier at all stages of training; the GMC number is the only constant reference from registration through a doctor's career.
26. Our intention with the exam data is to develop better reporting on a national level including a published annual summary of exam data including analyses of protected characteristics (age, gender, ethnicity group and disability) as well as by PMQ, training level, deanery and number of previous attempts.
27. Reports by, for example deanery or approved training programme would be available on a GMC web reporting platform such as that used for the ARCP data. It would be possible to review exam data and ARCP together for the same programme.
28. By holding exam data at the level of GMC number, we will be able to capture each trainee's progression in full and thereby investigate any patterns that may exist between exam passes and other progression points such ARCP and recruitment processes.
29. Exam data will be incorporated more effectively into our reporting for visits and monitoring of concerns. From PMQ data, we will be able to monitor more effectively trends across UK medical schools as well as deanery based analyses. Other research has shown that there is variation in pass rates by medical schools such as the McManus report (<http://www.biomedcentral.com/1741-7015/6/5/>). Exam data will provide us with a new perspective on progression from medical schools and will inform future policy developments that relate to undergraduate education. For example we will be able to report on the percentage of trainees passing their first attempt by medical school (across all exams taken).
30. With a growing volume of data, we will be able to perform longitudinal analyses of exam results to highlight improving or worsening performances over time while breaking down the analyses across place of training, ethnicity group, etc. Taking into account changes to exams, we can build trend rich reporting on various specialty specific issues and training related concerns. For smaller specialties we will also be able to report pass rates across several years (aggregating over time).

31. With our work on analysing recruitment data, combining exam data in the analysis will identify any influences that exams play for core trainees applying for higher specialty posts as well as foundation trainees looking to enter core training programmes. The analysis of foundation trainees will also help with our work with foundation schools. Identifying patterns where trainees are taking exams from more than one college is an area of research yet to be explored.

32. Exam data analysis will be used to inform the PLAB review. To be useful for this review data will need to include overall scores and scores by skill area.

Governance

33. We are committed to transparency and as such we intend to make the reports accessible over time.

34. The data itself is securely stored and only accessible to Education Staff. Access to the reports will be restricted to users through an online reporting tool that is password protected until we have fully understood the significance of the results and have had an opportunity to investigate any differences in outcomes. In a similar project, ARCP data are collected and reported by the same method with access only given where necessary.

35. One benefit as noted above is that exam data in combination with ARCP and recruitment data could be made available for research purposes. We would need to consider what mechanisms should be established to oversee this work.

36. From a general disclosure perspective, it is unlikely we would provide identifiable personal data in response to a public enquiry, for example, under the Freedom of Information Act. The latter contains exemptions which specifically relate to personal data. Our default position is that only aggregated data would be provided in response to an FOI request in accordance with the act.

37. Longer term, we are also exploring the potential for a data warehouse to facilitate the sharing of information with those involved in medical education and research while taking appropriate steps to ensure we continue to meet our statutory duties in respect of holding and processing data.

38. Our overall intention is to promote quality in medical education and training, not focus on individual performance.

Benefits for Postgraduate Medical Training

39. Collecting exam data by GMC number will allow us to subject the data from the colleges to greater scrutiny and validation. Using the current template, it is not possible to analyse the examination data against other data sources, or, as multiple diets are submitted together, report on the number of trainees who have taken the exam more than once during the reporting period. Through validation of the data, we can report more confidently with the enhanced reliability of data so that concerns can be acted upon more effectively and efficiently.

40. Incorporating exam data at this level of detail will add many benefits for developing a clearer picture of postgraduate medical training. This will encourage greater discussion between stakeholders as they view their data combined with other data sources to build a more complete picture of medical training and the part each stakeholder plays in it.

41. By combining exam data with ARCP and recruitment data, we can build a better picture of trainee progression. As access to information becomes more readily available, the collection of exam data at this level will also enhance the certification process when a lack of data delays awarding CCTs. All data will be available via log-in as a download.

42. On a practical level, the colleges will be freed from having to develop and write aggregated exam data summary points and will have access to benchmarking comparisons, where applicable.

Action Plan

43. We have had initial discussions with a few colleges to test the feasibility of a pilot which would help us identify data quality issues and resource requirements for this work. We would welcome interest from any colleges willing to take part in the pilot.

44. In order to comply with the Data Protection Act 1998 (DPA), organisations are generally required to provide a privacy statement or collection notice when “data subjects” supply their personal data. This sets out the proposed use for an individual’s personal information. In this context, a privacy statement should set out how colleges and faculties will use trainee data, as well as specifically refer to the transfer of their data to organisations such as the GMC. Contextual information about the data sharing work is likely to be useful for applicants.

45. In order to prepare for the 2013 ASR report and support the transfer of exam data, colleges will need to update their collection notices so that all applicants are aware that their personal data will be passed to the GMC. This should explain why their data is being shared. If organisations do not already have a privacy statement, we recommend they seek information governance advice if this is required. For those organisations with an existing privacy statement, we have provided an example of a statement that may be used to notify candidates about sharing their data, highlighting the purposes explained in this business case, e.g.

“If you are registered or anticipate being registered with the GMC, your personal data will be passed to the GMC for quality assurance purposes and to facilitate the awarding of CCTs.”

46. Based on the current position, we will start to collect data for the pilot as soon as privacy notices have been updated or established. We anticipate that Privacy Notices will be incorporated by Easter 2013. We will produce a number of reports as explained above using this data for review by the colleges. From September 2013

we expect all colleges to be compliant and ready for us to collect exam data as described in this proposal.

47. For colleges not in a position to share exam data until their privacy notices have been updated, we will seek to audit and clean demographic data to ensure that once we are in a position to share data they are clean. For example we can supply missing GMC numbers if enough demographic information is available and validate existing GMC numbers.

48. The accompanying annex contains links and information regarding privacy notices and details recorded by the Information Commissioner as they appear on college websites and exam application forms.