
Contents

Introduction	1
Project Aims	2
Project Milestones	3
Privacy notices	3
Piloting with colleges with adequate privacy notices	3
Implementation for all colleges and faculties	4
Reporting	4
Data Governance	5
Privacy Notice	5
Data Storage	5
Data Sharing Agreement	5
Data Uses	6
Information Commissioner Update	6
Reporting	7
Anonymous Reporting	7
Reporting Tiers	7
Reporting Requests	8
Data Submission – fields for collection	8
Candidate File	9
Returned Clean Candidate File	9
Exam Data	10
Research	10
Ethics and Research	10
Project Timeline	10
Implications for the certification process – possible future developments	10
Contacts and Information	11

Introduction

1. The GMC collects aggregated reports from medical royal colleges and faculties on pass rates for the examinations they run as part of the Annual Specialty Reports. Since the end of 2012 however, we have been working with colleges and faculties to collect and analyse identifiable exam data in the same way that we collect and report on the results of the Annual Review of Competence and Progression/ Record of in-Training Assessment (ARCP/RITA) processes.

2. There are several limitations to aggregated data: we can't test it against ARCP data to verify if the ARCP processes have been applied correctly, we can't validate the data to ensure consistency across datasets and little analysis could be performed on the data or conclusions drawn as a result.

3. This note summarises the information previously circulated in developmental briefing notes and provides and sets out the position we have reached to date. The intended audience of this note is mainly colleges, royal colleges and faculties though this information may also be useful for trainees/exam candidates as well as training providers.

4. We have worked with all colleges and faculties to discuss the project and address any queries they may have had. For a list of all queries raised by colleges, please see the project's Frequently Asked Questions (URL to be added).

Project Aims

5. The aim of the project is to improve how we analyse the progression patterns of doctors in training in accordance with our statutory responsibilities.

6. The project also importantly will make information about the outcomes of training more transparent to trainees, across the different specialty and GP training programmes and help those responsible for the quality of training to meet our standards in *The Trainee Doctor*.

"9.1 Organisations providing postgraduate training must demonstrate they are collecting and using information about the progression of trainees to improve the quality of training.

9.2 Trainees must have access to analysis of outcomes of assessments and exams for each programme and each location benchmarked against other programmes."¹

7. The project involves moving from collecting an aggregated summary of exam results to collecting raw data about individual candidates from colleges. This will enable us to:

a. Establish a clean and tested dataset of exam results for analysis by helping colleges and faculties to improve the accuracy and completeness of their candidate data.

b. Integrate exam results with ARCP data.

c. Aid in policy developments related to assessments by highlighting trends in the data that may shed light on issues in training. For example research we commissioned on medical school selection methods recommended that we should "*discuss, highlight and promote the importance*

¹, *The Trainee Doctor*, http://www.gmc-uk.org/Trainee_Doctor.pdf_39274940.pdf, Domain 9, (paragraph 10(c)), page 36

of further longitudinal work to follow up medical students through to trained post to enable the comparison of prior educational attainment".²

- d. Analyse exam data by UK medical school by tracking back using the GMC number to link to the register to provide medical schools with more information on the progression of their graduates across the different specialty and GP training programmes.
- e. Develop reports analysing exam data against data from ARCP, recruitment and the National Training Survey census to investigate potential associations between these markers of progression, including in relation to trainees' characteristics such as disability and ethnicity.
- f. Improve the efficiency of sharing data across interested stakeholders. This could include providing a single transparent point for trainees' exam outcomes to assist postgraduate deans, colleges and faculties in the management of training programmes and for the certification process.

Project Milestones

Privacy notices

8. Upon the presentation of the proposal, we sent colleges and faculties a note asking them to implement a Privacy Notice in order to notify candidates with a GMC number and candidates who will in the future register with the GMC that their exam data will be shared with us. As applications for exams due to take place in the Autumn of 2013 would start to be received by colleges and faculties months before, we requested Privacy Notices be implemented by 31 March 2013 to ensure that we could lawfully receive a full dataset for the 2013–14 training year.

9. In addition to the implementation of a Privacy Notice, we asked colleges and faculties to submit the actual data fields they use to record exam data in their various systems. These fields would be assessed for relevance with the project and then written into a Data Sharing and Confidentiality Agreement. The fields were requested to be received by 31st May 2013.

Piloting with colleges with adequate privacy notices

10. Some colleges had a Privacy Notice sufficient to be able to share exam data at an individual trainee level. These colleges kindly agreed to take part in a pilot for the project to assess whether collecting exam data for individual trainee doctors is ultimately useful; the results being reported to all colleges and faculties subsequently. The pilot aims to test:

- the process of implementing a privacy notice

² *Identifying best practice in the selection of medical students*, page 87 http://www.gmc-uk.org/Identifying_best_practice_in_the_selection_of_medical_students.pdf 51119804.pdf

- the efficiency of cleansing data
- the quality of data that can be achieved when combined with our other data sources
- the cost of time and resources for processing exam data from all colleges
- which analyses and reports are most useful to us and the colleges
- the capability of our systems to store and use the data

11. The reports for the pilot will be set up on an online reporting tool (password protected) where colleges and faculties will be able to access reports displaying their data in different ways (for example pass rates by deanery).

Implementation for all colleges and faculties

12. Once colleges and faculties close applications for their final exam sittings for the 1 August 2013 – July 2014 academic year the college can submit their full candidate file to the GMC for cleaning and matching with the register. Some colleges have indicated that they would prefer to share data following each diet which we can accommodate. However we will not be able to begin cleaning the data until June 2014, after the survey has closed and we have validated trainees' details with them.

13. From summer 2014, once we have returned the cleaned candidate files, colleges and faculties will be able to submit the exam results for 2013/14 academic year. Note:

- a. Exam files should only be submitted for exam sittings for which the privacy notice in place communicated the sharing of these data with the GMC.
- b. Exam results should only be submitted for candidates who are registered with the GMC.

14. There may be candidates who are not registered with the GMC when they sit the exam but subsequently register, some years later. If these candidates enter training programmes we would be able to identify them through analysis of ARCP results or the National Training Survey. If however candidates do not enter training when they subsequently register it may be difficult to identify them unless we keep the candidate files over time. We cannot currently assess the number of doctors who might fall into this latter category and we will need to monitor this in the first few years of the project.

Reporting

15. The development of reports will begin in winter 2013 but it will not be possible to deliver the reports until the full set of examination data are received. In practice, this is likely to be in the summer/early autumn of 2014.

Data Governance

Privacy Notice

16. The GMC has previously run projects to collect personal data including information related to ARCP³ outcomes. In that project, we suggested a suitable privacy notice for deaneries to implement. We sought legal advice in respect of this notice. Based on our previous experiences, we likewise sought Counsel's advice for this project.

17. Our proposed approach set out in our proposal was considered by Counsel to comply with the requirements of the Data Protection Act (DPA) as to the provision of fair processing of information. This is on the understanding that it will only apply prospectively to new candidates and will not apply retrospectively to candidates so that exam results taken prior to the implementation of the new privacy notices will not be shared with us.

18. The declaration to exam candidates that their personal and exam data will be shared with us is given as colleges and faculties display the proposed privacy notice during the application process. This would include adding the privacy notice to the exam application forms or added to websites should a college or faculty accept online applications.

19. For the avoidance of doubt the privacy notice should expressly mention exam data. The following wording was reviewed by Counsel and recommended to colleges and faculties:

If you are registered or anticipate being registered with the GMC then your personal data, including data about your exam results, will be passed to the GMC for quality assurance and research purposes and to facilitate the awarding of certificates of completion of training (CCTs).

Data Storage

20. Colleges and faculties will be asked to submit their exam data to us via our secure information sharing portal. Submissions are stored on the GMC's internal network in a directory to which only selected GMC Education staff will have access. All GMC staff undertake mandatory data protection and information security training as part of their staff induction. Data protection training is then renewed every two years, and information security training is renewed on an annual basis. The GMC is certified to the international information security standard, ISO – 27001. We are audited against this standard.

Data Sharing Agreement

21. To protect the demographic and exam data for individual candidates, we asked colleges and faculties to sign a data sharing and confidentiality agreement.

³ <http://www.gmc-uk.org/education/14105.asp>

After requesting and receiving comments on its form, the agreed document included what the data could and could not be used for as well as the specific data that would be shared between the both parties. Any additional use or additional data sharing would require a new data sharing agreement.

22. An example agreement can be found here: [Confidentiality and Data Sharing Agreement - Exam Data](#)

23. Upon entering into an agreement, exam data will be exempt from any request under the Freedom of Information Act (FOIA), Section 41, as the data was given to us in confidence. This exemption will apply unless we are legally compelled to disclose (e.g. by order of a court or competent body). The GMC would also engage Section 40 FOIA in relation to any request for personally identifiable information.

Data Uses

24. We have committed through the data sharing agreement not to use exam data to proactively open Fitness to Practise investigations. Furthermore, as per the working arrangements within the GMC, no directorate outside of Education will have access to these data.

25. With regard to the relationship between exam data and Fitness to Practise we have also identified the following:

- a. From time to time, examination results may be relevant when questions are raised about a doctor's fitness to practise.
- b. Equally, a doctor may seek to bring evidence of exam passes to demonstrate competence.
- c. It may be appropriate for a college examiner to refer a candidate to the GMC if, for example, there were questions raised about a doctor's probity such as fraud in the application or cheating in the examination processes.
- d. Exam data may be requested for a doctor who is being investigated for other reasons. However, the existing routes for the GMC to request information from colleges are appropriate and will remain as such in spite of the collection of this data.

Information Commissioner Update

26. Colleges and faculties are obliged to maintain a "notification" with the Information Commissioner's office. This outlines the nature of activities and the range of personal data processed by colleges and faculties, including a high level description of data recipients. We advise all colleges to make sure that their notification with the Information Commissioner is up to date. The following link explains giving notice:

http://www.ico.gov.uk/for_organisations/data_protection/notification.aspx

27. Each year, colleges and faculties will be asked to review and update their notification by the Information Commissioner. Colleges and faculties can also do this without any prompting from the Commissioner. It is our view that the majority of college notifications already provide sufficient detail to cover this proposed usage.

Reporting

Anonymous Reporting

28. The GMC will take all reasonable steps to prevent disclosure of confidential information, including:

- a. Exam data returns will be accessed by and visible only to the GMC Education and Standards Directorate staff responsible for delivery of education quality assurance.
- b. Identifiable candidate information will not be handled by or be visible to any other third parties, individuals or organisations. Anonymous raw data (i.e. with all identifiers removed) may be given to academic researchers upon application. Any such requests will be carefully considered and the Academy of Medical Royal Colleges will have representation on the group that will consider such applications.
- c. Confidential exam data will not be visible to or used in relation to any other areas of GMC business. Over time we may wish to work with colleges and faculties to use the data to support the certification process but this project is still in its infancy. We will not use this exam data for fitness to practice.
- d. Reports or publications based on exam data will be aggregated and anonymised. Individual respondents will not be identifiable in these reports. The only exception to this is the display of exam passes counting towards CCT in a doctor's GMC Online account.

Reporting Tiers

29. We will implement a tier system regarding data. The tier system suggested is as follows:

- a. Tier 1: Passes only reported on doctor's GMC Online account which will assist in the certification process.
- b. Tier 2: Only pass/fail published in aggregated form whereby no trainee is identifiable. For example pass rates for trainees in a particular deanery and specialty.
- c. Tier 3: Actual marks only used for academic research purposes. This is likely to be the mark relative to the pass mark, as the pass mark changes with each diet. We suggest that the actual mark or marks in relation to the pass

mark will only be available to academic researchers upon application. Colleges will have the opportunity to consider such applications. Any researcher would sign a confidentiality agreement and would not receive data containing identifiers such as GMC number. If the research data sets contain data from multiple sources this would already be linked using pseudonymised identifiers.

Reporting Requests

30. In order that the project benefits all interested stakeholders, we are open to requests for specific on-line reports from Colleges and Faculties, Dean's, Heads of Schools, etc. in addition to our reporting on progression.

Data Submission – fields for collection

31. From our conversations with colleges, we identified that colleges hold differing levels of data for exams, due mainly to the nature of exams themselves as well as system capacity and worth of data for future purposes. As such we agreed an individual data specification for each exam with each college and faculty. One of the benefits of this project is that colleges are freed from completing the current exam data template as part of the ASR process. As such, we only request exam data that colleges and faculties actually capture. However, if colleges and faculties develop their systems to include more data, we may also request this within an appropriate timescale.

32. In order to outline these specifications we requested colleges and faculties to send us their fields they use in their local systems. From this we could identify the fields we would use for reporting purposes, raising any queries if necessary and then highlighting those fields that we requested.

33. We will not request exam data for candidates who are not registered with the GMC at the time they take the exam. However, if these candidates subsequently become registered with the GMC we may ask for their exam data retrospectively providing they will have been presented with the appropriate Privacy Notice. We will never request exam data for candidates who never register with the GMC.

34. At the outset, one aspect of the data cleaning process which will affect the time it takes to process all datasets is whether candidates without registration with the GMC can sit an exam. In order for us to plan the amount of work the data cleaning may take, we asked colleges and faculties to inform us whether each exam that they run can have candidates sitting without registration with the GMC.

35. Based on previous work with candidate exam data, the most straight forward process for obtaining accurate data whereby colleges and faculties receive the benefit of clean datasets was to firstly ask for a simple candidate file which would consist of the following:

Candidate File

Field	Format
GMC Number	String of seven characters (where present)
Forenames	String
Surname	String
Gender	String
Date of Birth	DD/MM/YYYY
PMQ Year	String
PMQ Institute	String

36. From the above fields, we would be able to validate the GMC number, or in some cases add it where missing. Where there may be inconsistencies we will highlight the corrections and return the clean candidate file to the colleges and faculties. In the candidate file, we may be able to provide additional fields such as candidate's most current training level, specialty programme and post, current deanery/LETB, etc. from the deanery-provided census data depending on the timing of the exams and other data sources. The following example highlights additional fields that may be shared with colleges and faculties:

Returned Clean Candidate File

Field	Format
GMC Number	String of seven characters (corrected where needed/added where missing if possible)
Forenames	String
Surname	String
Gender	String
DOB	DD/MM/YYYY
PMQ Year	String
PMQ Institute	String
Training Level	String
Deanery/ LETB	String
Specialty Programme	String
Specialty Post	String

37. From the returned candidate file, colleges and faculties would then be able to extract exam data using the specification we agree with each college and faculty to return a final exam data file as follows:

Exam Data

Field	Format
GMC Number	String of 7 characters
First Name	String
Surname	String
+ DATA SPECIFICATION	Agreed Data Fields

38. As highlighted in the proposal and after receiving feedback, this exchange will be an overall benefit to colleges and faculties. In subsequent years to the initiation of this project, the cleaning process would get to a point where very little cleaning would be required, especially for candidates retaking exams.

Research

Ethics and Research

39. The analyses on the exam data that we are proposing to undertake to fulfil our statutory functions in relation to quality assurance of medical education and training does not require approval from a research or ethics committee. However, any research proposals that include the use of these data will be required to obtain ethics approval.

Project Timeline

- Beginning of June each training year Receive Candidate Files



- End of June each training year Return clean Candidate File



- End of July each training year Receive Exam Data File



- End of August each training year Reports published

Implications for the certification process – possible future developments

40. Colleges and faculties play an important role in scrutinising trainees' applications for a Certificate of Completion of Training (CCT). The current process for certification relies on the sharing of documents generated by the doctor in training between the postgraduate dean and the college or faculty. Following a recommendation from the trainees' local postgraduate dean that the trainee has

completed the requirements of their training programme the relevant college or faculty checks the application before recommending the award of a CCT to the GMC.

41. Some have inferred that early collection of exam data would remove the need for college or faculty scrutiny of applications for CCT. We do not agree: Our experience of quality assuring the CCT recommendation process confirms that the colleges play a vital role in assuring that only those who have successfully demonstrated ALL of the required competencies of the curriculum are recommended for a CCT. This is an important assurance mechanism and the GMC is not planning to remove the colleges' and faculties' role in scrutinising applications for a CCT.

42. Conversely, there may be opportunities to explore with colleges and faculties how the current arrangements can be improved to enable them to contribute to the quality assurance process throughout the training programme, not just at the end when reviewing documentation.

43. We would value greater college and faculty involvement through, for example, sampling educational supervisors' reports to help ensure that ARCPs are properly evidence based. We are also aware that there is a wealth of quality data within e-portfolios (and paper portfolios). We would value college and faculty scrutiny here too, with the results being reported through the Annual Specialty Report which we are working with the Academy of Medical Royal Colleges to enhance.

44. These ideas are emerging from discussion with colleges and faculties about the QA Review and from our collective experience of the certification process. We are still at a very early stage and we will seek over the next year to work the Academy to develop these initial thoughts into proposals. Any development work would therefore inevitably take several years.

Contacts and Information

45. If you would like to meet with us to discuss the project or for any general questions, please contact Anthony Rimmer at quality@gmc-uk.org.

All information related to the exam data project can be found here:

www.gmc-uk.org/education/examdata