

Meeting: 1st - 2nd June Joint COPMeD COGPED Directors Meeting	Date: 27 May 2016
Report Title: UKMED update	
Submitted by: Daniel Smith, Kirsty White	
Report summary: COPMeD, the four countries, Medical Schools Council, The UKMED website has been launched: http://www.ukmed.ac.uk/ Three research projects are now underway; with a further four due to commence over the summer. We have begun to consider options and priorities for Phase 2 and are inviting feedback on the research questions that UKMED should be seeking to answer in Phase 2 in order to promote excellence in medical education and increase understanding of training pathways.	
Main Issues: We are interested in views on options and priorities for Phase 2 of UKMED.	
Questions for COPMED to consider: <ol style="list-style-type: none"> 1. What research questions should UKMED help us answer about medical education and training pathways? About other areas? 2. Therefore which datasets in table 2 should be prioritised for inclusion in phase 2? 3. Are there are any datasets not listed that should be considered for inclusion? 	

UKMED update June 2016

Background

- 1** The UK Medical Education Database (UKMED) project brings together undergraduate and postgraduate data into a single repository of data describing doctors' progression from entry to medical school through to the first few years of training and practice.
- 2** These data have not been linked systematically before and no single stakeholder can currently link all data on their own. The current phase of the UKMED project is therefore a pathfinder pilot focused on the first two phases of training (medical school and foundation) as a way of testing the quality of the data, the systems, processes and, the associated costs of establishing and maintaining the database.
- 3** Higher Education Statistics Agency data (HESA) are used to define cases for inclusion as all those who start a degree that leads to provisional registration. For phase 1 these have been restricted to those commencing their studies in 2007 and 2008 (two admissions cycles). HESA data are linked to test provider data (UKCAT and GAMSAT) using the UCAS person id.¹ HESA cases are matched to the GMC register using the medical school code and the medical school's internal identifying number for each of their students,² which the GMC also receives as part of the provisional registration process. The use of database identifiers to link gives an efficient and reliable mechanism.
- 4** The GMC number provides links to postgraduate data, including data from applications to foundation training, Annual Review of Competency Progression (ARCP) outcomes and fitness to practise data; a full list is in the data dictionary³.
- 5** The project is supported by the UKMED Development Group, an advisory group, including current and future potential data contributors as well as data subject (medical students and doctors in training) representation. Members are listed in Annex A.

Progress Update

- 6** The UKMED website, hosted by the Medical Schools Council was launched in October 2015 and applications for research datasets were invited over the winter.
- 7** The applications were reviewed by a research sub-group (the members are listed at Annex A), against a set of published criteria to determine academic rigour and feasibility⁴. The criteria were developed collaboratively by the research sub-group and reviewed against three test research applications submitted by data contributors before being endorsed by the UKMED stakeholder group and piloted.
- 8** Data sets have been provided to support the three test research applications and work continues on these projects.
 - a** Predicting Fitness to Practise issues from admission profiles in UK medical school entrants.
 - b** What has been the impact of accelerated graduate-entry medicine courses in terms of educational and sociodemographic profile, success at medical school, completion of Foundation training, and specialty entry?
 - c** Do the Educational Performance Measure decile score and SJT predict successful completion of the foundation programme?
- 9** We received ten research applications through the on-line application process and all of these were reviewed by the research sub-group members. Four are scheduled to be supported over the summer:
 - a** The role of academic attainment in understanding sex differences in specialty choice and fitness to practise.
 - b** The allocation of doctors to specialties and general practice training posts by demographic and socio-economic characteristics.
 - c** "Getting on" in medicine: a programme of study of careers trajectories and decisions of doctors.
 - d** What demographic and educational factors predict doctors' decisions to apply for training programmes in particular medical specialties?
- 10** A summary of these seven projects is provided on the research page of the UKMED web site - <http://www.ukmed.ac.uk/research>
- 11** We ensure compliance with the Data Protection Act by de-identifying the data: cases are assigned their own unique Study – Id in place of a GMC or HESA number. We ensure that combinations of quasi-identifiers do not uniquely identify one case in the extract⁵. The extract is shared via a safe haven to minimise the risk of re-

identification; it allows the researchers to run analysis on the extract using the statistical package of their choice, whilst preventing the export/import of data thereby preventing re-identification through linkage. Researchers are under contract to only use the data for the purposes of the approved research proposal. Outputs are reviewed by the development group prior to publication.

- 12** Note that UKMED is a mechanism for providing protected access to the data; the research projects are not funded by UKMED.

Initial scoping of datasets identified for consideration for Phase 2

- 13** We are preparing an options paper for UKMED phase 2 for consideration at the 15th June 2016 UKMED Development Group.
- 14** Phase 1 of the project was designed as a prototype to enable correlation of selection tests at entry to medical school with performance in the first few years of practice. Currently the database contains only two cohorts so one option for Phase 2 is to simply include future cohorts and more recent versions of medical school selection tests and continue to focus on entry to medicine.
- 15** However there may be capacity to include further datasets and broaden the benefits of UKMED. In this case, the Development Group have given a steer that the building of the Database should be led by clear research questions that contribute to improving our understanding of medical education and training pathways.
- 16** The table in Annex B describes the datasets UKMED Group members have previously suggested for consideration; both in terms of expanding existing data sets beyond the 2007 and 2008 cohorts currently in UKMED and the possibility of including new data sets. Each of the datasets identified in the table includes commentary on expected benefits. Where the dataset would require further work before it could be considered for Phase 2, the status and additional work is described.
 - a.** For Table 1 (existing data sets), of existing data sets, are there any other issues for further exploration?
 - b.** For Table 2 (potential new data sets)
 - i** What are the key benefits that should be sought seeking from the next phase?
 - ii** Are there data sets that should be included, amended or disregarded?
 - iii** Are there additional considerations we will need to explore for any of the data sets identified?

c. Of the possible data sets and related benefits, which should we prioritise?

17 The working assumption is that only data sets which already have appropriate information governance in place to enable sharing into UKMED will be considered for Phase 2.

Next steps

18 Please share feedback via email (quality@gmc-uk.org) with UKMED in the subject line of the email.

19 To join the UKMED mailing list and for general queries please go to <http://www.ukmed.ac.uk/contacts/new>

Annex A

UKMED Development Group Members

Prof Steve Thornton (Chair)	Vice Principal (Health) and Executive Dean of the Barts and the London School of Medicine and Dentistry
Professor Jane Dacre	Academy of Medical Royal Colleges
Dr JP Lomas	Academy of Medical Royal Colleges Trainees
Charlie Bell	British Medical Association Medical Students Committee
Harrison Carter	British Medical Association Medical Students Committee
Mark Shannon	BioMedical Admissions Test
Prof Bill Reid	Conference of Postgraduate Medical Deans
Alan Robson	Department of Health
Marita MacMahon Ball	Graduate Australian Medical Schools Admissions Test
Veronica Vele	Graduate Australian Medical Schools Admissions Test
Kirsty White	General Medical Council
Martin Hart	General Medical Council
Paul Buckley	General Medical Council
Jonathan Howes	Health Education England
Prof Derek Gallen	The Wales Deanery
Dr Katie Petty-Saphon	Medical Schools Council
Duncan Henderson	NHS Education for Scotland
Prof Keith Gardiner	The Northern Ireland Medical and Dental Training Agency (NIMDTA)

Prof Jon Dowell	Research Subgroup Chair
Prof Chris McManus	Professor of Psychology and Medical Education, University College London.
Rachel Greatrix	UK Clinical Aptitude Test
Charlotte Dainter	UK Foundation Programme Office

UKMED Research Sub-group members

Prof Jon Dowell (Chair)	Head of Division of Undergraduate Medical Education Dundee University
Prof Steve Thornton	Chair of Development Group
Daniel Smith	General Medical Council
Dr Paul Garrud	Graduate Australian Medical Schools Admissions Test
Dr Sarah McElwee	BioMedical Admissions Test
Dr Paul Tiffin	Centre for Medical Education Research
Prof Peter Johnston	NHS Education for Scotland
Prof Chris McManus	Professor of Psychology and Medical Education, University College London
Dr Kim Walker	UK Foundation Programme Office
Sandra Nicholson	UK Clinical Aptitude Test
Olga Sierocinska King	Medical Schools Council
Siobhan Fitzpatrick	Medical Schools Council
Kirsty White	General Medical Council

Annex B

Table 1: Existing datasets

Dataset	Current Status	Further considerations
Higher Education Statistics Agency (HESA)	HESA data for the 2007 and 2008 intakes are included in UKMED.	<p>GMC holds HESA data for 2002/03 - 2014/15 and the data sharing agreement enables these data to be included however we would need to explore purchase of the annual data set beyond 2014/15.</p> <p>In the future HESA data could contain marks for individual modules. We have asked HESA to supply a list of modules that they currently collect data on whether students successfully complete. If these modules look to be useful, HESA might consider supplying the mark instead. We would also need to determine what supplementary information would be required to interpret these module marks – for example the pass mark of each sitting.</p>
Annual Reviews (ARCP)	ARCPs related to the 2007 and 2008 cohorts are included in UKMED.	GMC has collected ARCP data for all foundation and speciality programmes by GMC number annually and hold records from 05/08/2009 onwards. Note that foundation records were only collected from 01/08/2012 when the foundation programme moved to an ARCP process. The GMC's annual collection will continue.

Dataset	Current Status	Further considerations
Foundation Programme Applications (UKFPO)	UK graduate data is included in UKMED for those applying to foundation programme in 2012, 2013, 2014 and 2015. Application data from non-UK graduates is not included.	<p>The data sharing agreement would need to be revised to include non-UK graduate applicants.</p> <p>We would need to clarify the impact of impending system changes. We understand that for 2016 the data will come from the existing system, but from 2017 onwards it will come from ORIEL.</p>
Recruitment data (HEE central and IT suppliers)	We currently collect application/offer/acceptance data from ORIEL. We hold first round recruitment outcomes from 2012 onwards. 2015 is the first year these data have come from ORIEL.	<p>May additionally include:</p> <ol style="list-style-type: none"> 1. For ST1/CT1 entry applicants who are not in a current foundation programme, or have not completed a foundation programme within the last three years are required to complete an "Alternative Competency Form". This has been in practice for a number of years although revised year on year. The forms and guidance can be found via the below link https://www.oriel.nhs.uk/Web/ResourceBank/Edit/MTAyMw%3d%3d 2. Documentation on the derivation of interview scores and shortlisting scores for each specialty and each year? 3. ORIEL 2016 data will include more scores for individual sections of the recruitment process – tests, stations and so forth. A set of columns in ORIEL will contain these data, what a given column means will depend on the specialty and is defined in a specification template. We will need to vary the current Data Sharing Agreement (DSA) to show that we intend to include these fields in the 2016 extract. We will need to obtain four Nation sign-off on the DSA revisions.

Dataset	Current Status	Further considerations
UK Clinical Aptitude Test (UKCAT)	Data for the 2007 and 2008 cohorts.	<p>Data for applicants to medical school from 2009 to 2014, i.e. for all those cases which can be identified as starting medical school using the HESA data already available to the GMC.</p> <p>We can consider adding information on who applies for a bursary to take the test.</p> <p>The UKCAT has developed over time e.g. in later years it included non-cognitive tests and a situational judgement test, so some additional fields may be available for certain years.</p>
Graduate Admissions Test (GAMSAT)	Data for the 2007 and 2008 cohorts.	Data for applicants to medical school from 2009 to 2014, i.e. for all those cases which can be identified as starting medical school using the HESA data already available to the GMC.

Table 2: Potential new data sets

Dataset	Current status	Further work required	Benefits	Contacts
Common fields across test provider registration forms	To be discussed with all test providers in the review of their data sharing agreements.		To allow consistent data capture of key demographic variables of interest that are not available in the HESA extract.	BMAT, GAMSAT and UKCAT
Biomedical Admissions Test (BMAT)	Privacy notices are being reviewed to enable future cohorts to be included in UKMED	We continue to explore with BMAT the potential for historical cohorts to be included. This may be possible if we contact the data subjects to notify them that their BMAT data will be included within UKMED.	Enable medical schools using BMAT to evaluate their selection methods. Enable full-cohort analysis. BMAT includes a test of scientific knowledge and applications. ⁶ UKCAT does not include this.	BMAT
Prescribing Skill Assessment (PSA) data	A data sharing agreement has been signed. The data fields have been defined. MSC plan to submit the data to the GMC in the autumn of	The data quality and coverage is unknown. The PSA has been used differently across schools e.g. in some it is a summative assessment, in others a formative assessment.	Ability to assess the predictive validity of the PSA. The four health departments have agreed to inclusion of the PSA within the UK Foundation Programme.	MSC Assessment

Dataset	Current status	Further work required	Benefits	Contacts
	2016.			
Royal College exam data	<p>The GMC have collected identifiable data since 2014 by GMC number for all exams sat since 1 August 2013⁷ and these are reported on publically⁸.</p> <p>They were considered for inclusion within Phase 1 at the UKMED meeting in October 2013.⁹ However, at that time the quality and complexity of the data was unknown.</p>	<p>Review of privacy notices and revision of existing data sharing agreements.</p> <p>The data collected varies by exam and has been determined by data that were available in the college systems when the project started.</p> <p>A paper will go to the AoMRC Assessment Committee on 13th October outlining how UKMED could include these data and the potential benefits to the colleges concerned.</p>	<p>Researchers will be able to obtain college exam data linked to measures of attainment on entry to medical school (A-levels and test provider data).</p> <p>The Royal College exam data will give more reliable and valid criterion scores to use in predictive validity studies than currently available from ARCP and recruitment events. Furthermore the measures are continuous (e.g. score relative to pass mark is available) giving more possibilities in the statistical analysis.</p>	Academy of Medical Royal Colleges, the individual medical royal colleges and faculties submitting data to the GMC.
Individual medical school selection data – Multiple Mini Interviews (MMI)	<p>Approximately, 18 medical schools use MMI in selection.</p> <p>The MMI data are held by the</p>	<p>Review of privacy notices.</p> <p>Clarification of:</p>	Ability to assess the predictive validity of MMIs, one of the most widely used selection tools.	One contact per medical school is required.

Dataset	Current status	Further work required	Benefits	Contacts
	individual medical schools.	<ul style="list-style-type: none"> ■ The schools using MMI ■ The years the MMIs have been used. ■ Format (one row per person per interview) ■ Identifiers (e.g. UCAS Personal ID – the ten digit number) ■ Data fields for interviews: descriptions, length, content, scoring methodology. ■ Data fields for interviewers: (gender, position and so forth) 		
Individual medical school selection data – statements and references	Individual medical schools may hold data used in their selection processes, for example scoring of personal statements, work experience forms and/or references.	To clarify the range of selection tools used by different schools and the availability of data.	<p>Ability to assess the predictive validity of other measures used in selection and for schools to demonstrate the validity of the tools they use.</p> <p>Ability to understand contextual</p>	One contact per medical school is required.

Dataset	Current status	Further work required	Benefits	Contacts
	Individual medical schools may flag applicants as being eligible for a contextual offer and may collect other data relevant to the widening participation agenda.		admissions.	
Student Fitness to Practice (FtP)	Since 2015 medical schools have provided final year student fitness to practice data to the GMC for provisional registration applicants.	These data have been collected for the specific purpose of registration and we would need to review the information governance arrangements and engage with students and medical schools.	<p>These data contain information such as the outcome of the student FtP proceedings that are not contained in the character declaration data. This could improve the analysis of selection methods and whether they can predict medical school fitness to practice processes.</p> <p>One of the first UKMED research studies is investigating whether declarations of fitness to practise can be predicted by measures obtained on entry to medical school, including some non-cognitive tests (purporting to measure personal attributes) piloted by UKCAT in 2007 and 2008.</p>	Medical schools and GMC registration.

Dataset	Current status	Further work required	Benefits	Contacts
Additional protected characteristics	Since 18 January 2016 doctors have had the option to share new information through their GMC Online account about their religion and belief, sexual orientation and disability, in the same way as they do with other organisations.	Information governance review	Ability to extend differential attainment research to consider the experience and outcomes for other protected characteristics.	GMC
E-portfolio data	Foundation trainees and trainees in each specialty use e-portfolios to record workplace based assessments. The system supplier and the available data vary by specialty and may vary by year.	To ascertain how useful e-portfolio data might be and the work involved, it might be best to select one or two e-portfolios for inclusion on a pilot basis. The foundation e-portfolio will be used by all UKMED cases and is an obvious candidate for any pilot.	Ability to assess the predictive validity of workplace based assessments used in national training programmes.	Deaneries for foundation e-portfolio. Individual colleges for specialties.
Clinical outcomes for individual consultants/GPs	Ludka-Stempień ¹⁰ (2015) notes three major groups of data that could be used as criterion measures for	Significant work to assess the availability, range and quality of the individually identifiable data sets.	Potential to link doctors' training outcomes to clinical practice, see for example Norcini et al (2014) ¹⁴	The Healthcare Quality Improvement Partnership

Dataset	Current status	Further work required	Benefits	Contacts
	<p>predictive validity studies:</p> <ol style="list-style-type: none"> 1. Measures of malpractice e.g. complaints against medical professionals 2. Specific measures of clinical performance e.g. morbidity rate and mortality rate. The Society for Cardiothoracic Surgery (SCTS) publishes outcome data by individual operator, so a precedent exists.¹¹ However the data source for this is the National Adult Cardiac Surgery Audit (NACSA) not Hospital Episode Statistics. 3. Subjective ratings of clinical performance e.g. information from patient satisfaction questionnaires, 	<p>We would need to explore whether there are particular procedures where it would be reasonable to attribute the event to the responsible consultant (instead of a team/service).</p> <p>HSIC note that activity linked to the GMC number of the lead consultant responsible for the care of the patient will not be directly attributable to that consultant and can only be attributed to the 'consultant team' as it incorporates (although does not currently distinguish between) the work of the whole team including junior doctors, anaesthetists, nurse practitioners and other members of staff.</p> <p>It may be more appropriate to use outcome measures from national clinical audits, which will have addressed</p>		

Dataset	Current status	Further work required	Benefits	Contacts
	<p>perceived quality of medical services, and peer skills assessment</p> <p>The <i>HSIC Hospital Episode Statistics (HES) Analysis Guide</i> (March 2015)¹² suggests that data are held at a clinician identified level, but we would need to engage with HSIC to establish whether data could be placed into UKMED on an identifiable basis.</p>	<p>some of the data quality issues in the design of the audit.</p> <p>The Healthcare Quality Improvement Partnership¹³ lists a number of national audits which may include some containing patient outcomes linked to a GMC number. These audits may also have addressed issues such as case mix.</p>		
Full placement history	A full history of each trainee's training placements, as opposed to the annual NTS snapshot.	This may be contingent on developments to the LETB and deanery systems to allow transfer of this volume of data to the GMC.	An understanding of whether the posts a trainee rotates through is associated with performance on particular elements of an exam. An understanding of whether exposure to a specialty is associated with specialty choice.	LETBs and Deaneries databases.
Electronic Staff Record data from each nation for primary and	These data are held by the GMC for mapping doctors to responsible officers for revalidation purposes	Ensure the data sharing agreements would allow inclusion in UKMED. Ascertain how much	The ability to look at employment post CCT, for example who goes on to become a consultant. To improve workforce planning, which the Public Accounts	Departments of Health

Dataset	Current status	Further work required	Benefits	Contacts
secondary care.		preparation of these data would be required to make them useful for UKMED purposes.	Committee has recently indicated requires improvement. ¹⁵	
Revalidation ¹⁶	GMC hold revalidation data, the following statuses are available for each doctor: recommendation to revalidate, approved to defer (insufficient evidence to support a recommendation to revalidate), recommendations to defer (participating in an ongoing process) or non-engagement.		Understanding of which factors predict revalidation status.	GMC Registration

References

¹ https://www.hesa.ac.uk/component/studrec/show_file/10051/a%5E%5EUCASPERID.html

² https://www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&mnl=09051&href=^ ^a^ ^OWNSTU.html

³ http://www.ukmed.ac.uk/documents/UKMED_data_dictionary.pdf

⁴ http://www.ukmed.ac.uk/documents/UKMED_research_process.pdf

⁵ See L. Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), 2002; 571-588. <http://dataprivacylab.org/dataprivacy/projects/kanonymity/kanonymity2.html>

⁶ <http://www.admissionstestingservice.org/for-test-takers/bmat/about-bmat/>

⁷ <http://www.gmc-uk.org/education/14171.asp>

⁸ <http://www.gmc-uk.org/education/25495.asp>

⁹ Dowell, J. (September 2013). *Paper i – Introduction - A UK Medical Education Database (UKMED) Data describing the selection and progression of doctors from medical school application to specialisation and beyond*

¹⁰ Ludka-Stempień Katarzyna (September 2014). *Predictive validity of the examination for the Membership of the Royal Colleges of Physicians of the United Kingdom* (Thesis submitted for the degree of Doctor of Philosophy (PhD) - Division of Medical Education, Medical School, University College London

¹¹ <http://www.scts.org/patients/>

¹² Health & Social Care Information Centre *HSCIC Hospital Episode Statistics (HES) Analysis Guide* (March 2015). Available at: http://www.hscic.gov.uk/media/1592/HES-analysis-guide/pdf/HES_Analysis_Guide_Jan_2014.pdf

¹³ <http://www.hqip.org.uk/national-programmes/a-z-of-nca/>

¹⁴ Norcini, J. J., Boulet, J. R., Opalek, A. & Dauphinee, W. D. (2014) 'The Relationship Between Licensing Examination Performance and the Outcomes of Care by International Medical School Graduates'. *Academic Medicine*, 89 (8). pp 1157-1162/10.1097/ACM.0000000000000310.

¹⁵ <http://www.parliament.uk/business/committees/committees-a-z/commons-select/public-accounts-committee/news-parliament-2015/nhs-staff-numbers-report-published-15-16/>

¹⁶ <http://www.gmc-uk.org/doctors/revalidation/revalreports.asp>

