

Process for completing UKMED research

Version 7 – December 2023

Table of contents

1. [Introduction](#) p 2
2. [Scope of process](#) p. 2
3. [Who can apply for a dataset from the UK Medical Education Database?](#) p. 2
4. [Application process](#) p. 3
 - [Stage 1 – prior to completing an application form](#) p. 3
 - [Stage 2 – completing an application form](#) p. 4
 - [Stage 3 – review of application form](#) p. 4
 - [Stage 4 – finalising the data specification](#) p. 7
 - [Stage 5 – Data sharing agreement/contract - GMC Data Controller](#) p. 7
 - [Stage 6 – De-identification of the data set](#) p. 8
 - [Stage 7 – Software arrangements and configuring the Safe Haven](#) p. 8
 - [Stage 8 – Review of analysis outputs against Statistical Disclosure Controls](#) p. 10
 - [Stage 9 – UKMED review of research reports](#) p. 11
 - [Stage 10 – Close of project](#) p. 12
 - [Stage 11 – Research publication](#) p. 13
5. [Change control](#) p. 13

Appendix A – [Abbreviations](#)

Appendix B – [GMC template Data Sharing Agreement](#)

Appendix C – [Data User Agreement to access the HIC Safe Haven for approved UKMED Research Projects](#)

Appendix D – [Software arrangements as per December 2023](#)

Introduction

- 1 This document sets out the process for requests for de-identified linked data from the UK Medical Education Database (UKMED) and criteria for assessing these applications.

Scope of process

- 2 This process applies if the applicant wishes to use a research dataset only available from UKMED, because the proposed dataset contains data linked from multiple sources that contribute to UKMED. For instance, the proposed data come from UCAT, GAMSAT, BMAT, the UKFPO and the GMC.
- 3 This process does not need to apply if an individual applicant only wishes to access data from one organisation that contributes to UKMED. For instance, if only GMC data are required for research the applicant should apply directly to the GMC and the GMC's own processes would apply. Equally, if the applicant only wished for access to UCAT data; they should apply directly to UCAT.
- 4 A research project using UKMED, and data not previously held in UKMED but submitted for a specific project by a research team is also within the scope of this process. Researchers wishing to augment UKMED data with their additional data are required to follow the document on [Researchers' own data](#).

Who can apply for a dataset from the UK Medical Education Database?

- 5 The application process will be open to all those who can complete an application, have funding and/or secured staff time for the project and who would be in a position to satisfy the terms of the data sharing agreement with the GMC as data controller.
- 6 All organisations contributing data will follow this process for access to linked UKMED datasets.
- 7 Previous applicants who have received data from UKMED must have demonstrated compliance with the earlier contract and delivery of the planned outputs e.g., a peer reviewed article.
- 8 The data are for research purposes only and cannot be used to make decisions about individual data subjects.
- 9 Students can apply to use UKMED data for a research project:

9.1 Students can apply directly but must have a named supervisor from an academic institution or relevant organisation.

9.2 The supervisor will co-sign the contract for accessing UKMED data for research. Accountability lies with the supervisor and their associated organisation for the supervision of the student's research activities, including compliance with the clauses set out in the contract.

10 Researchers with students on the research team

10.1 Researchers who wish to undertake a research project where students are part of the team, are recommended to make an application to UKMED prior to when the student will need to complete the work for their course.

Application process

Stage 1 – prior to completing an application form.

11 A Welcome pack for prospective applicants is available on the UKMED website.

Ethics approval for research

12 Researchers may wish to complete a self-assessment: <https://www.hra-decisiontools.org.uk/research/>. For research projects using data exclusively held by UKMED, researchers are advised to contact a local ethics committee chair to confirm whether the project meets the criteria for an exemption from the need to apply for ethics approval.

13 Researchers supplying data not held by UKMED will be required to demonstrate that the institution of the lead researcher has granted ethics approval or that their proposed study is exempt from the requirement to obtain ethics permission.

Funding for UKMED Research

14 Applicants will be expected to demonstrate that funding and/or staff time has been committed to the research or that a relevant research grant has been initiated before application is reviewed. Historically, an important proportion of projects have not been completed, therefore applications without funding/secured staff time will not be reviewed.

15 No data extracts will be prepared or shared until funding and/or other resources have been secured.

Stage 2 – completing an application form

- 16** Prospective research applicants should complete a standardised application form (illustrative example available on <https://www.ukmed.ac.uk/>) containing details about their research question, methods and researchers who will be expected to use the data requested. The online application form must be accessed and submitted by logging into the UKMED website.
- 17** The closing date/s for submitting completed applications seeking approval from a given meeting of the UKMED Advisory Board will be published in advance on the UKMED website.

Stage 3 – review of application form

- 18** The UKMED Research Subgroup will assess the completed applications against the criteria in Table 1. Members will be able to abstain from scoring a domain if they do not have the expertise pertinent to the specific case or due to conflict of interest.

Table 1 Scoring for UKMED Research Applications

Domain	Unacceptable 0	Uncertain – may be acceptable with further clarification 1	Acceptable 2
Research Question	Aims or question not clear. Question appears irrelevant to policy or practice.	Question is poorly defined or uncertain as to whether it is likely to impact on policy or practice if answered. More information or clarification may be required.	Question if answered has significant implications for policy or practice in medical education. The research question is original and is not answered in the existing literature.
Data fields requested	The data requested is not contained in the UKMED.	The data requested is within the database but does not appear capable of answering the research question identified.	The data requested is contained within the UKMED and is well linked to the research question.
Proposed Methodology	Methods not appropriate to	Methods may seem outdated, over simplistic or	Methodology takes into account the nature of and type of data available and

Domain	Unacceptable 0	Uncertain – may be acceptable with further clarification 1	Acceptable 2
	addressing research aims or question.	not well adapted to nature of data.	is suitable to address the research question.
Analysis	Analysis not appropriate to addressing research aims or question.	Analysis may seem outdated, over simplistic or not well adapted to nature of data.	Analysis takes into account the nature of and type of data available and is suitable to address the research question.
Evidence of planned output/use	No statement of intended use.	Intended uses unclear.	There is a clear statement on the intended outputs which may include publication in a peer-reviewed journal, publication on an organisation's website, reports that are evaluations of a service rather than research, or publication as a Ph.D. thesis.
Team	<p>Team is not plausible with little evidence of relevant skills or track record.</p> <p>There is no governance structure defined or there is indication that it may not be accountable.</p>	<p>Team may have obvious skills gaps or limited relevant research track record.</p> <p>Governance structure is outlined, but is not very clear.</p>	<p>Proposed team members have a good track record in related research and are likely to have the skills to employ the proposed methodology and manage data issues. FTE for each project team member provided.</p> <p>Governance structure is</p>

Domain	Unacceptable 0	Uncertain – may be acceptable with further clarification 1	Acceptable 2
			described in a clear and accountable way.

- 19** The UKMED Research Subgroup members will complete an initial assessment of each application independently. This is followed by a discussion during the meeting of the Research Subgroup.
- 20** When a request includes data pertaining to one or more medical royal college or faculty (college) exams, each college involved will be asked to review the application using the criteria in Table 1. Colleges may nominate the person best qualified to comment on their behalf. Colleges are not obliged to comment.
- 21** A representative of colleges will be invited to attend in person or remotely the relevant section of the Research Subgroup’s meeting to discuss the proposal. Attendance is not obligatory.
- 22** The Research Subgroup will provide an assessment of the proposal for the UKMED Advisory Board against all domains, summarising: the quality of the submission, the potential outcomes of the research and benefits to medical education.
- 23** The UKMED Advisory Board will consider the report from the Subgroup and provide advice to the GMC on whether or not the application should be supported and, if not why and what if any feedback should be given to the applicant(s).
- 24** UKMED Advisory Board members have committed to support research proposals that have academic merit, even if the possible outcomes may be unfavourable for data contributors or have findings that could be considered controversial.
- 25** In case of approved applications, the title, summary and contact details of the lead researcher as recorded on the application form will be published on the UKMED website.
- 26** The GMC assigns priority to production of data extracts for specific projects under direction of the UKMED Advisory Board.

Stage 4 - Finalising the data specification

- 27** Once approved, the researcher will receive an email requesting the following:

27.1 Confirmation or amendment of the proposed research dataset. This will be provided in Excel spreadsheet describing the criteria for inclusion in the extract, any further filters on specific tables and details of the table and fields that will be contained in the extract.

27.2 Details required by the GMC's Information Governance team in order to complete HIC Data User Agreements (DUA) in Appendix C such as the contact details for the person who will countersign for the user, which is typically the user's manager or head of department. Note that the DUAs can be used across UKMED projects, so only new users will be required to complete these. Only work email can be used for HIC Safe Haven accounts.

27.3 Details of how to complete the GDPR Quiz that users are required to complete before they obtain their account.

Stage 5 - Data sharing agreement/contract - GMC Data Controller

28 Once the specification is finalised the GMC as Data Controller will issue a Data Sharing Agreement (a template in Appendix B). This will contractually restrict the researcher's use of the data to that required for the completion of the research outlined in their approved proposal. It is important to note that the data cannot be used to support measures or decisions with respect to particular individuals and cannot be processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject*. The GMC require one signed Data Sharing Agreement per organisation the research team is associated with. If the researcher is employed by an organisation and they are conducting their UKMED research in the course of their employment it is not necessary to require the employer organisation to sign a data sharing agreement. If, however, the researcher's project is not part of their employment e.g., they are a PhD student, the countersignature of the supervising body is required (typically head of department).

29 The GMC will produce the extract for the study to the required specification, ensuring that the methodology of production is documented.

30 Extracts can be refreshed, by for example including additional years of data from tables already included in the Data Sharing Agreement without reapplying for access. Some researchers will find additional data becomes available during the course of their project as UKMED tables are updated annually.

Stage 6 - De-identification of the data set.

31 When providing row by row data, individual doctors will be pseudonymised. Each GMC Reference number contained within the dataset will be replaced by a unique

* See section 19 of the Data Protection Act (2018) here:
<http://www.legislation.gov.uk/ukpga/2018/12/section/19>

study ID. If the dataset contains multiple records with the same GMC number, these records will have the same unique study ID. The unique study ID will consist of a consecutive number and if multiple projects are using the same dataset a project code prefix. The GMC will hold a table that maps GMC numbers to study IDs to allow re-identification in the event of the data being queried. Study IDs will only be used for one study. This table will only be accessible to analysts working on the UKMED project. When datasets are re-issued a new study ID may be used, in which case it will not be possible to link to the earlier version of the dataset. Generally variables names are stable so scripts can be re-run, but we cannot guarantee this.

- 32** The GMC will ensure that individuals cannot be identified using a combination of demographic variables using data minimisation techniques such as by applying the concept of K-anonymity. This is satisfied if $K > 1$ for each combination of quasi-identifiers – gender, age, medical school and so forth*. To achieve this, it may be the case that some values will be recoded into broader categorisations. The GMC will minimise any reduction in utility by, where possible, only recoding variables that are not explicitly an object of the research proposal. If other techniques are used these will be outlined.
- 33** On completion of the project the GMC as Data Controller will take responsibility for the secure archiving of relevant analysis files for a period of five years, including a copy of the extract provided.

Stage 7 – Software arrangements and configuring the Safe Haven

- 34** Researchers will be completing their analysis in the University of Dundee's Health Informatics Centre (HIC) Trusted Research Environment (TRE), Safe Haven. Researchers will complete a HIC/GMC Data User Agreement – please see Appendix. C. Note that researchers will need an authorised signatory from their organisation. The GMC will countersign this agreement. Each researcher requiring access to the Safe Haven will need to complete a Data User Agreement. If the researcher has completed one for a previous UKMED project, there may be no need to complete another one.
- 35** Researchers will need to complete the *Research, GDPR and confidentiality – what you really need to know* e-learning modules and demonstrate they have passed the *Research, GDPR and Confidentiality Quiz* by saving the certificate produced and sending it to the GMC together with their signed Data User Agreement. The course and the quiz can be found at: <http://byglearning.co.uk/mrcrsc->

* 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>

lms/course/category.php?id=1. If the researcher has completed one for a previous UKMED project there is no need to complete another one.

- 36** Researchers will be remotely logging onto a secure server located within HIC to access the data and perform analysis, without being able to copy or remove the data from the secure central server. All the common tools required for the analysis (Appendix D) are provided for use within this environment and this server can be accessed securely from anywhere.
- 37** The remote-access Safe Haven utilises a cloud-based Amazon Web Services (AWS) environment. In this model data are not released externally to researchers for analysis on their own computers but placed on a server in AWS, a secure IT environment, where the researcher is given secure remote access to analyse it.
- 38** The GMC supply the data to HIC and GMC will be responsible for all queries regarding the data. Researchers will have a named point of contact at the GMC for this purpose. The GMC will transfer files to HIC via a secure file transfer. Within 48 hours HIC will transfer these files to the Safe Haven environment (except during the 2-week Christmas/New Year period when there will be no Safe Haven support available).
- 39** The GMC will supply the data to HIC as MS-SQL database backup (.BAK file) using one of the GMC's secure mechanisms. HIC will restore the database within the Safe Haven and provide access to a read only version and a read/write version of the research extract database.
- 40** The research extract database will contain:
 - 40.1** de-identified data as per the source tables described in the UKMED Data Dictionary and specified for inclusion in the data sharing agreement.
 - 40.2** A table containing linked person level data (one row per person) taken from the original tables (those in 45.1) with sets of columns for each event where the original table contains a row for each event (e.g., an attempt at a test or exam). The columns will be in chronological order, so that variables pertaining to entry to medical school will be before those containing postgraduate outcomes.
- 41** This research extract database will be created by producing views and tables as specified in the project's data sharing agreement in UKMED (known as AGORA in the GMC's Oracle database). These will use the project specific identifier from the approved proposal on the UKMED website (i.e. UKMEDP[n]) as a prefix in their name. Note that with the current versions of ORACLE and MS-SQL there is a 128 character limit on the length of the names. These tables and views will be copied into a project specific MS-SQL database in order to produce the .BAK file.
- 42** For ease of data management, when researchers require their data to be refreshed, we may only refresh particular tables in the research extract database. For example, if new postgraduate exam outcomes are available only EXAM_TOTAL_MARKS would

be refreshed. These data refreshes will be version controlled within the research extract database. These refreshes will be delivered to HIC via one of the GMC's secure routes. UKMED is refreshed twice a year: undergraduate data (UCAS, HESA and test provider data and postgraduate data (ARCP, EXAM and ORIEL). The GMC will be responsible for importing them into the research extract database and managing version control within the database by liaising with the researchers.

- 43** Researchers will be able to connect the database to the analysis packages identified in Appendix D (detailing GMC licensing arrangements) and import the dataset of their choice. This may be the entire person level table, a subset of columns from it, or another table or view constructed from the sources tables also contained in the research extract database.
- 44** Previously written customised code/syntax, libraries of reference data and so forth can be imported once approved by the GMC.
- 45** HIC is responsible (hicsupport@dundee.ac.uk) for managing access to the TRE/Safe Haven and working with the researchers to ensure the required software is available. The GMC (ukmed@gmc-uk.org) is responsible for answering any queries on the data supplied.
- 46** Arrangements for the software depend upon the software required and the nature of the license held by the researcher as detailed in Appendix D.
- 47** All software within the Safe Haven is licenced for academic research only. If the researcher is working on a commercially funded project additional fees will apply.
- 48** If a researcher wants to use a different research application to those listed in Appendix D they will need to speak with HIC to check that the Safe Haven environment can support the software. Some packages may not work correctly in the Safe Haven environment so this needs to be looked at on a case-by-case basis.
- 49** For software that is not included as standard and where HIC Safe Haven can support it, researchers must buy the necessary licence along with the software media (see Appendix D for exceptions for SAS, SPSS and STATA) and pay HIC an installation fee per install.

Stage 8 - Review of analysis outputs against Statistical Disclosure Controls

- 50** When the researcher has completed their analysis, outputs intended for the public domain, for example a table of results, will be reviewed by the GMC using the following statistical disclosure controls*:

* <https://www.hesa.ac.uk/about/regulation/data-protection/rounding-and-suppression-anonymise-statistics>

50.1 0, 1, 2 are rounded to 0

50.2 All other numbers are rounded to the nearest multiple of 5

50.3 Percentages based on fewer than 22.5 individuals are suppressed

50.4 Averages based on 7 or fewer individuals are suppressed

50.5 The above requirements relate to headcounts, Full-Person Equivalent (FPE) and Full-Time Equivalent (FTE) data. Financial data are not rounded.

51 Charts and visualisations should also apply statistical disclosure controls to reduce the risk of reidentification:

51.1 Histograms

51.1.1 Aggregating categories to distributions with long tails and small numbers;

51.1.2 Masking the scale of the X and Y axis by omitting all values and using the maximum or minimum values;

51.2 Box plots and scatter plots

51.2.1 Grouping together or averaging the minimum and maximum values as outliers may be attributable to a single observation.

51.2.2 Grouping together or averaging the original data to generate scatter plots.

52 Data output requests are typically processed by HIC, then shared with the GMC. The GMC will review the files in line with statistical disclosure controls and if approved, share the output analysis files with researchers via GMC Connect within two working days. Researchers are strongly encouraged to leave sufficient time in their plans for their output to be reviewed before being passed to them.

Stage 9 - UKMED review of research reports

53 Research output from UKMED will be reviewed by the Subgroup and reported on to the Advisory Board. The final report from a project should state on the front page the research questions from the original application. This review will assess:

53.1 Were research questions posed in the original application addressed?

53.2 Is nature of data and UKMED appropriately represented?

53.3 Did the team conduct any additional analysis of potential concern. If yes, why?

- 53.4** Are there any reasons for challenging presentation/publication. If yes, why?
- 54** In reviewing the research outputs that are not to be published in peer-reviewed journals, for recommendation to the UKMED Advisory Board, the Subgroup may have regard to the following questions:
- 54.1** Have the data sources been correctly documented and acknowledged?
- 54.2** Is the overall design appropriate and adequate to answer the research question?
- 54.3** Are the data subjects adequately described?
- 54.4** Have inclusion and exclusion criteria been described?
- 54.5** Have data quality issues such as data linkage adequacy and missing data been reported?
- 54.6** Is the sample reported representative of any population it seeks to make inference about?
- 54.7** Are the methods adequately described?
- 54.8** Is the derivation of the variables used appropriate and adequately described?
- 54.9** Are all required statistics reported?
- 54.10** Is the interpretation warranted by and sufficiently derived from/focused on the data?
- 54.11** Is the interpretation discussed in the light of previous evidence?
- 54.12** Are the study limitations noted?
- 55** The reviewers will identify any specific issues that the UKMED Advisory Board or any of the data providers may wish to consider. For instance, concerning reputation, timing of release or prior notice of sensitive findings for relevant parties.
- 56** When research output includes analysis of data pertaining to one or more college exams, each college involved will be invited to review the output using the Subgroup's criteria. Colleges may nominate any persons they feel best qualified to comment on their behalf. Colleges are not obliged to comment.
- 57** Colleges may choose to have a representative attend in person or via telephone conference the relevant section of the Research Subgroup's meeting to discuss the output. This is not obligatory.

- 58** The outcome of the review is either a recommendation to the UKMED Advisory Board to publish or request for further work before re-submission.
- 59** For any of the review questions, the Research Subgroup may recommend further work prior to submission to a journal or publication by other means.
- 60** The output of the Subgroup's peer review typically will only be shared with the submitting researchers and the UKMED Advisory Board. However, applicants should note that the GMC is subject to the Freedom of Information Act and is unable to offer a blanket restriction on disclosure.

Stage 10 - Close of project

- 61** A project will be closed when one of the following occurs:
 - 61.1** The project team have had their research paper accepted for publication.
 - 61.2** Users request the project is closed.
 - 61.3** Insufficient progress has been made, such as when the next year of data has been loaded into the database before any work has started.
- 62** Closure will trigger the archive process.
- 63** The script file should be described in a way that cross-references the script to the published outputs – for example by table number in the final report and placed in the folder Script for GitHub The GMC will then add the script files into a public shared repository..
- 64** The names of the final data files and the contact person for any queries should be emailed to UKMED@gmc-uk.org. It should be possible to generate the published analysis from the final data files and the script file.
- 65** The GMC will record these details in the register with the UKMED and HIC identifiers included. If a researcher needs it for reference purposes, they should contact UKMED@gmc-uk.org.
- 66** HIC will archive the final files in an archive project folder (with the HIC identifier) and delete all other files.
- 67** The GMC will take a copy of the archive project folder and hold it in a secure location on their system with access restricted to key GMC staff working on the UKMED project.
- 68** If the original project team or a new team wish to re-use data files or script files, an application to the UKMED Advisory Board will be required. The application should reference the files required as per the register entries.

Stage 11 - Research publication

- 69** When the research output is finalised and no later than one year after the agreed end of research project, a link to a peer-reviewed publication will be published on the UKMED website.
- 70** At least 30 days before publication a confidential copy of the submitted research output has to be provided to UKMED Advisory Board so that it may be checked for likely impact and (confidentially) shared with UKMED data contributors whose data has been utilized so they are aware and can consider any implications for their organization.
- 71** In case of a research output not intended for peer-review publication, upon being finalised and no later than one year after the project is approved by Advisory Board for publication of findings, a full output will be published on the UKMED website.
- 72** Final publications need to include the following acknowledgement for audit purposes. "Source - UK Medical Education Database ("UKMED") UKMEDP[Number] extract generated on [dd/mm/yyyy] Approved for publication on [dd/mm/yyyy]. I am 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 Records [yyyy/yy] to [yyyy/yy] 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.

Change control

- 73** Versions of this document were reviewed by the UKMED Development Group (Renamed in December 2016 UKMED Advisory Board) on 19 June, 15 October 2015 and by email in January 2019.

Version 7

Version 7 of this document contain changes:

- 74** Clarification when countersignature is required for the Data Sharing Agreement (Stage 5)

- 75 Publication of researchers' syntax file on close of project (Stage 10)
- 76 Inclusion of the text of acknowledgement required in publications (Stage 11)
- 77 HIC software updated in appendix D

Version 6

78 Version 6 of this document contains changes to:

- 78.1** Stipulate that funding and/or secured staff time are a pre-requisite of making an application for UKMED data,
- 78.2** Researchers must seek ethics approval or check whether they are exempt,
- 78.3** HIC processes migration into TRE.

Version 5

79 Version 5 of this document contains changes to:

- 79.1** Who can apply for a dataset (paragraphs 15 and 16). Information for student applicants
- 79.2** Stage 7 – Software arrangements and configuring the Safe Haven (paragraphs 44 – 48). To improve data management data extracts will now be provided to researchers as a database in the Safe Haven. GMC staff will be able to assist researchers with connecting to analysis packages to their research extract database.
- 79.3** Stage 8 - Review of analysis outputs against Statistical Disclosure Controls (paragraph 58). Statistical disclosure controls for charts and visualisations

80 Version 4 of this document reflects the following changes:

- 80.1** New datasets (BioMedical Admissions Test – [BMAT] scores and UCAS applications);
- 80.2** The option of bringing in additional datasets – paragraph 9,
- 80.3** Clarification of exemption from ethics approval – paragraph 14,
- 80.4** Archiving data extracts;
- 80.5** Note that there is a separate process for access to extracts for medical school entry profiles and workforce planning. This may be more suitable for those

working for an applicable organisation who do not wish to publish research. Please see *UKMED training pathway analysis*.*

- 81** Appendix B – GMC template Data Sharing Agreement has been updated to reflect the commencement of the General Data Protection Regulation (GDPR) on 25 May 2018.
- 82** Appendix C - Data User Agreement to access the HIC Safe Haven for approved UKMED Research Projects has been updated to reflect the commencement of the General Data Protection Regulation (GDPR) on 25 May 2018.

Version 3

- 83** Amendments have been made to **paragraphs 22, 23, 54 and 55** to ensure the involvement of colleges in requests involving exam data.

* https://www.ukmed.ac.uk/documents/UKMED_Training_pathway_analysis_extracts_reports.pdf

Appendix A – Abbreviations

Glossary

AoMRC	Academy of Medical Royal Colleges, www.aomrc.org.uk
ARCP	Annual Review of Competence Progression
ASME	The Association for the Study of Medical Education https://www.asme.org.uk/
AWS	Amazon Web Services
BMA	British Medical Association, www.bma.org.uk
BMAT	BioMedical Admissions Test, www.admissionstesting-service.org/for-test-takers/bmat/about-bmat
COPMed	Conference of Postgraduate Medical Deans (UK), www.copmed.org.uk
DHSC	Department of Health and Social Care, www.gov.uk/government/organisations/department-of-health-and-social-care
DPA	Data Protection Act, www.gov.uk/data-protection/the-data-protection-act
DSA	Data Sharing Agreement
EDI	Equality Diversity Inclusivity
EEA	European Economic Area
EPM	Educational Performance Measure
FtP	Fitness to Practise
GAMSAT	Graduate Medical School Admissions Test, www.gamsat.acer.edu.au
GDPR	General Data Protection Regulation, https://www.gov.uk/government/collections/data-protection-act-2018
GMC	General Medical Council, www.gmc-uk.org
HEIW	Health Education and Improvement Wales, www.heiw.nhs.wales
HESA	Higher Education Statistics Agency, www.hesa.ac.uk
HIC	Health Informatics Centre, https://www.dundee.ac.uk/hic
IS	Information Services (part of GMC)
LRMP	List of Registered Medical Practitioners
MAPs	Medical Associate Professions
MSC	Medical Schools Council, www.medschools.ac.uk
TRE	Trusted Research Environment
UCAS	Universities and Colleges Admissions Service, www.ucas.com
UCAT	University Clinical Aptitude Test, www.ucat.ac.uk Formerly UKCAT.
UKFPO	UK Foundation Programme Office, www.foundationprogramme.nhs.uk Data from Foundation Programme application system https://www.oriel.nhs.uk/Web/ , prior to 2016 from Foundation Programme Application System (FPAS).

Appendix B – GMC template Data Sharing Agreement

Note highlighted sections to be completed for final contract.

Dated [date of contract]

The General Medical Council and [Researcher Organisation (s)]

**Agreement relating to the provision of UKMED data for research purposes –
EXAMPLE ONLY NOT FOR USE**

Contents

DATED [DATE OF CONTRACT]

THE GENERAL MEDICAL COUNCIL AND [RESEARCHER ORGANISATION (S)]

1. DEFINITIONS AND INTERPRETATION
 2. USE OF HEALTH INFORMATICS CENTRE (HIC) DATA SAFE HAVEN
 3. HESA STANDARD ROUNDING METHODOLOGY AND ACKNOWLEDGMENT FOR USE OF DATA
 4. ETHICAL APPROVAL, LICENCES AND CONSENTS
 5. RESEARCHER AND INSTITUTION STATUS
 6. KEY PERSONNEL
 7. INTELLECTUAL PROPERTY
 9. CONFIDENTIALITY
 10. FREEDOM OF INFORMATION
 11. DATA PROTECTION
 12. WARRANTY, INDEMNITY AND INSURANCE
 13. DISPUTE RESOLUTION
 14. POLICIES AND STANDARDS
 15. ASSIGNMENT AND SUB-CONTRACTING
 16. PUBLICITY, MEDIA AND OFFICIAL ENQUIRIES
 17. AUDIT
 18. NOTICES
 19. ENTIRE AGREEMENT
 20. GOVERNING LAW AND JURISDICTION
- BACKGROUND
1. DEFINITIONS AND INTERPRETATION
 2. OBLIGATIONS OF THE GMC
 3. OBLIGATIONS OF THE RESEARCHER

This Agreement is made on [Add date]

between (the Parties):

(1) General Medical Council whose principal place of business is at General Medical Council, Regent's Place, 350 Euston Road, London, United Kingdom NW1 3JN (GMC); and

(2) [Researcher Organisation] [a company incorporated in England under number [tbc] as appropriate] whose principal place of business] is at [tbc] .

Background

A The Researcher has submitted a proposal to the UKMED Development Group for access to data which has been approved by the GMC for the purpose of answering the specified research questions.

B The Researcher requires the provision of certain data in order to support the research. The GMC has agreed to supply the data on the terms and conditions set out in this Agreement.

It is hereby agreed as follows:

1. Definitions and interpretation

1.1 In this Agreement (unless the context otherwise requires), the following words and phrases shall have the following meanings:

Agreement: these terms and conditions and each of the attached schedules as varied or amended from time pursuant to the Change Control Procedure;

Approved Research: the research referred to in recitals A and B above, set out in the proposal referred to in recital A, which has been approved by the GMC and for which the GMC is agreeing to provide the Data.

Background Intellectual Property: the Intellectual Property in works:

(i) existing before the Approved Research started; or

(ii) developed, written or prepared other than in order to provide the Approved Research;

Commencement Date: [insert];

Confidential Information: all information which is disclosed before or after the Commencement Date by one party to the other however conveyed and which would appear to a reasonable person to be confidential or which is marked confidential or is accompanied by a written statement saying that it is confidential or proprietary, which relates to the business, products, developments, trade secrets, know-how, personnel, customers and Researchers of the party disclosing it, and all information derived from the above;

The Data: the Data placed in the Safehaven by the GMC as described in clause 2.1.

Derived data: all data-fields generated by you whilst undertaking your research project using the UKMED data. This, in effect, is your final working dataset, including the analytical dataset used together with all newly-generated data-fields. This includes derived data-fields created from analysis of existing data, plus all intermediates of those derived data-fields.

Dispute Resolution Procedure: the procedure set out at clause 13;

Data Protection Legislation: the General Data Protection Regulation, the Data Protection Act 2018, and any subsequent legislation regulating the use of personal data;

Employee: any person employed or engaged by the Researcher or any of its agents, subcontractors, Researchers or associates;

Ethical Approval: any approvals or consents to be obtained from professional or other bodies that are required by the Researcher for the performance of the Approved Research or any of them;

FOI Legislation: Any or all of the Environmental Information Regulations 2004, the Environmental Information (Scotland) Regulations 2004, the Freedom of Information Act 2000 or the Freedom of Information (Scotland) Act 2002, or any subsequent legislation which amends or replaces them;

Foreground Intellectual Property: the intellectual property in works developed, written or prepared by the Researcher (or its subcontractors or consultants);

HESA Standard Rounding Methodology: is the application of the following processes to data or statistics:

- 0, 1, 2 are rounded to 0.
- All other numbers are rounded to the nearest multiple of 5.
- Percentages based on fewer than 22.5 individuals are suppressed.
- Averages based on 7 or fewer individuals are suppressed.
- The above requirements relate to headcounts, FPE and FTE data.

Or such other rounding methodology as HESA may from time to time notify to the GMC as replacing the methodology described here.

Intellectual Property: means patents, inventions, know-how, trade secrets and other confidential information, registered designs, copyrights, database rights, design rights, rights affording equivalent protection to copyright, database rights and design rights, semiconductor topography rights, trademarks, service marks, logos, domain names, business names, trade names, moral rights, and all registrations or applications to register any of the aforesaid items, rights in the nature of any of the aforesaid items in any country or jurisdiction, rights in the nature of unfair competition rights and rights to sue for passing-off;

Key Personnel: the persons specified in clause 6 of this Agreement, or otherwise notified by the Researcher to the GMC and approved by the GMC as provided for in clause 6;

Request for Information: a request for information to which this Agreement applies which falls within the FOI Legislation, or a request which appears to fall within that description;

Researcher Personnel: any employees, workers, agents, consultants and officers of the Researcher (or any subcontractor engaged by the Researcher) and any individuals contracted to the Researcher (or to any such Researcher subcontractor) who are involved in the performance of the Approved Research;

Safe Haven: the virtual research environment which is approved by the GMC for the purpose of storing and analysing the Data;

Staff: all persons employed by the Researcher to perform its obligations under this Agreement together with the Researcher's servants, agents, Researchers and subcontractors used in the performance of its obligations under this Agreement;

Syntax: the script/code used to generate your results in the software employed;

Working Day: any day from Monday to Friday (inclusive) which is not a Bank Holiday in England and Wales.

1.2 In this Agreement (unless the context otherwise requires):

1.2.1 The words **including** and **include** and words of similar effect shall not be deemed to limit the general effect of the words which precede them;

1.2.2 Reference to any agreement, contract, document or deed shall be construed as a reference to it as varied, supplemented or novated;

1.2.3 Obligations undertaken by a party which comprises more than one person shall be deemed to be made by them jointly and severally;

1.2.4 Words importing persons shall include firms, companies and bodies corporate and vice versa;

1.2.5 Words importing the singular shall include the plural and vice versa;

1.2.6 Words importing any one gender shall include either other gender;

1.2.7 References to a numbered clause, schedule or paragraph are references to the clause, schedule or paragraph of this Agreement so numbered;

1.2.8 Reference to any legislative provision shall be deemed to include any statutory instrument, by-law, regulation, rule, subordinate or delegated legislation or order and any rules and regulations which are made under it, and any subsequent re-enactment or amendment of the same; and

2. Use of Health Informatics Centre (HIC) Data Safe Haven

2.1 The GMC shall place pseudonymised data within the Safe Haven, where the Recipient can access a pseudonymised dataset to undertake the Approved Research. Such access shall be available for a specified period notified to the Recipient following signature of this Agreement.

2.2 The Recipient in accessing the Data from the Safe Haven shall:

2.2.1 Only process the Data:

2.2.1.1 To undertake the Approved Research; and

2.2.1.2 to the extent and in such a manner as necessary for the Approved Research;

2.2.2 not extract, copy, print, or in any way remove the Data from the Safe Haven without the prior approval of the GMC;

2.2.3 only extract Data subject to the HESA Standard Rounding Methodology specified in clause 3 and approved and authorised by GMC to be extracted by the Recipient from the Safe Haven as an Output under clause 2.2.6;

2.2.4 implement appropriate technical and organisational measures to protect the Data against unauthorised or unlawful processing and against accidental loss, destruction, damage, alteration or disclosure. Such measures shall be appropriate to the nature of the Data and the Purpose; and

2.2.5 notify the GMC of any accidental or deliberate unauthorised loss, destruction, damage, alteration or disclosure of the Data and take all reasonable steps to recover or ensure secure deletion of the Data; and

2.2.6 not to use the Data, or allow others to use the Data, to identify or attempt to identify individuals or to make decisions or enable any decisions to be made about an individual, or to contact or attempt to contact any individual.

2.2.7 not allow the Data to be accessed by any Researcher Personnel other than the Key Personnel.

2.2.8 not include any identifiable data in syntax used for cleaning or analysing the data extract as the syntax is exported and made public at project completion.

2.3 When the Recipient has finalised their analysis of the Data, the analysis ("the Output") will be placed in a secure lock within the Safe Haven environment to be checked for any risk that individuals are identifiable or potentially identifiable from the Output and to ensure that the HESA Standard Rounding Methodology has been applied in preparation of the Output as required by clause 2.2.3. Where no such risk is identified and the Rounding Methodology has been applied, the Output will be released to the Recipient by the GMC via a secure file transfer.

2.4 Where the Output has not been approved for release under clause 2.3, the reason for this will be explained to the Recipient who will be asked to amend the Output, including through application of the HESA Standard Rounding Methodology, to remove any risk of identifiability.

2.5 The Data will be retained by its Safe Haven provider for a period of seven years from the commencement date. On completion of the Approved Research, the Recipient may apply to the GMC to re-access the Data where this is necessary to check the accuracy of their original analysis.

3. HESA standard Rounding Methodology and acknowledgment for use of data

3.1 The Data includes information derived from data collected by the Higher Education Statistics Agency Limited ("HESA") and provided to the GMC ("HESA Data"). HESA makes no warranty as to the accuracy of the HESA Data and does not accept responsibility for any inferences or conclusions derived from its data by third parties.

3.2 Where the Recipient uses, reproduces or references in a publication HESA Data the Recipient shall:

3.2.1 acknowledge HESA as a source of the data and make clear that HESA does not accept responsibility for any inferences or conclusions derived from HESA Data by third parties e.g.

- Source:
- HESA Student Record 20XX/YYs
- Copyright Higher Education Statistics Agency Limited

The Higher Education Statistics Agency Limited cannot accept responsibility for any inferences or conclusions derived by third parties from data or other information supplied by it; and

3.2.3 ensure that the HESA Rounding Methodology as defined in paragraph 3.3 has been applied.

3.3 The "HESA Standard Rounding Methodology" is the application of the following processes to data or statistics:

- 0, 1, 2 are rounded to 0.
- All other numbers are rounded to the nearest multiple of 5.
- Percentages based on fewer than 22.5 individuals are suppressed.
- Averages based on 7 or fewer individuals are suppressed.

- The above requirements relate to headcounts, FPE and FTE data.

Or such other rounding methodology as HESA may from time to time notify to the GMC as replacing the methodology described here.

4. Ethical Approval, licences and consents

4.1 The Researcher will obtain all relevant Ethical Approvals. On written request from the GMC, the Researcher will immediately provide copies of Ethical Approvals obtained in relation to their project or evidence that Ethical Approval is not required for their project.

4.2 If any licence or consent of any government or other authority is required, the Researcher will obtain and maintain such licence or consent at its own expense and produce evidence of it to GMC immediately upon GMC's written request.

5. Researcher and Institution status

5.1 The Researcher confirms that this research project is not commercially funded and does not form part of a commercially funded project

6. Key Personnel

6.1 The following persons will require access to the Safe Haven. Subject to clause 6.2, such persons shall be Key Personnel for the purposes of this Agreement:

Name	Role/Responsibility	Email contact address
------	---------------------	-----------------------

[Insert]	[Insert]	[Insert]
----------	----------	----------

6.2 The Researcher may appoint additional or replacement Key Personnel to those individuals specified in clause 6.1, provided that such appointments shall be subject to the prior written approval of the GMC (such approval not to be unreasonably withheld or delayed).

6.3 The Researcher shall ensure that all Key Personnel are subject to explicit contractual obligations to comply with the conditions on access to the Data which are set out in this Agreement and generally to keep the Data confidential and secure.

7. Intellectual Property

Background IPR

7.1 GMC's Background Intellectual Property will remain vested in GMC.

7.2 Background Intellectual Property of HESA or any other third party shall remain vested in HESA or another third party (as the case may be).

7.3 The Researcher's Background Intellectual Property will remain vested in the Researcher (or third party proprietor as the case may be).

7.4 Where, as part of its Background Intellectual Property, GMC provides third party Intellectual Property to the Researcher, it shall use reasonable endeavours to ensure that the Researcher is authorised to use such third party Intellectual Property for the purposes of the research project to which this Agreement relates.

7.5 All Foreground Intellectual Property will vest in the Researcher.

7.6 All rights in the name and logo of each party shall remain the absolute property of that party and the other party shall not by virtue of this Agreement acquire any rights or interest therein. Neither party shall use or exploit the other party's name or logo in any way without the prior written consent of the other party.

8. Review by UKMED Advisory Board prior to publication

8.1 The GMC acknowledges that the Researcher has a responsibility to make academic publications to demonstrate carrying out of its primary purposes; that is the advancement of education through research. The GMC is content to support this subject to the conditions below:

8.1.1 The Researcher will provide a copy of all proposed publications, derived data and syntax to the GMC for presentation to the UKMED Advisory Board not less than thirty (30) days in advance of the submission for publication.

8.1.2 The GMC retains the right to veto (and in doing so, withdraw consent for the Researcher to use any material to which the GMC owns the intellectual property rights in academic publications) any proposed publication that is factually inaccurate with regard to the role, structure and function of the GMC or other contributing organisation.

8.1.3 The GMC shall notify the Researcher of any veto within thirty (30) days after the receipt of the material, failing which the Researcher shall be free to assume that there is no issue and they are free to publish.

8.1.4 The researcher will acknowledge UKMED in the abstract and acknowledgements of their publication.

8.2 Where the Researcher wishes to use the information for any other purpose that is not related to the original purposes of the research, it shall request the GMC's specific prior written approval.

General

8.3 Where either party becomes aware of any infringement or allegation of infringement of Intellectual Property, that party shall:

8.3.1 Promptly notify the other party in writing of such infringement or allegation of infringement;

8.3.2 Where the other party is the owner of the infringing or allegedly infringing Intellectual Property:

(a) Allow the other party to conduct all negotiations and proceedings and give the other party all reasonable assistance; and

(b) Make no admission relating to the infringement or alleged infringement.

8.4 Notwithstanding any other provision of this Agreement, the licences granted under clause 7 will survive any termination of this Agreement.

9. Confidentiality

9.1 All Confidential Information given by one party to the other, or otherwise obtained or developed by one party relating to the other, shall be kept secret and confidential by the receiving or developing party (as applicable) throughout the term of this Agreement and following its termination or expiry for any reason and the receiving or developing party shall not use such Confidential Information other than for the purposes of the proper performance of this Agreement without the prior written consent of the other party.

9.2 The obligations of confidentiality in this clause shall not extend to any information which the receiving party can show:

9.2.1 Was independently disclosed to it by a third party entitled to disclose the same; or

9.2.2 Is required to be disclosed under any applicable law, or by order of a court or governmental body or authority of competent jurisdiction.

9.3 The obligations contained in this clause shall not apply to:

9.3.1. The use or publication of any material in accordance with clause 10 of this Agreement; or

9.3.2. Any disclosure made pursuant to the FOI Legislation.

9.4. If confidential data is provided to the Researcher in an anonymised/pseudonymised form, the Researcher should not, and should not support or allow third parties to:

9.4.1. Attempt to deduce the identity of individuals by a process of triangulation, which involves using other datasets to bypass anonymisation/pseudonymisation to identify individuals.

9.4.2. Use information, accidentally or otherwise, gleaned from identifying individuals by triangulation to make any decisions in relation to an individual.

9.5 In relation to section 171 of the Data Protection Act 2018, the Researcher acknowledges that they do not have the consent of the GMC to re-identify data subjects in the data extract.

10. Freedom of information

10.1 The Researcher acknowledges that GMC is subject to the requirements of the Freedom of Information Act 2000 (FOIA) and shall assist and cooperate with GMC to enable GMC to comply with its information disclosure obligations.

10.2 The Researcher shall and shall procure that its sub-contractors shall:

10.3 Transfer to GMC all Requests for Information that it receives as soon as practicable and in any event within two (2) Working Days of receiving a Request for Information;

10.4 Provide GMC with a copy of all information in its possession or power in the form that GMC requires within five (5) Working Days of GMC's request; and

10.5 Provide all necessary assistance as reasonably requested by GMC to enable GMC to respond to the Request for Information within the time for compliance set out in section 10 of FOIA or regulation 5 of the Environmental Information Regulations 2004.

10.6 GMC shall be responsible for determining in its absolute discretion and notwithstanding any other provision in this Agreement or any other agreement whether any information is to be disclosed or published, or is exempt from disclosure in accordance with the provisions of the FOIA.

10.7 In no event shall the Researcher respond directly to a Request for Information unless expressly authorised to do so by GMC.

10.8 The parties acknowledge that, except for any information which is exempt from disclosure in accordance with the provisions of the FOIA, the content of this Agreement is not confidential.

or

10.9 Each party acknowledges that the other party is subject to the requirements of the FOI Legislation and shall assist and cooperate with the other party to enable the other party to comply with its information disclosure obligations.

10.10 Where either party receives a Request for Information which relates to information held pursuant to this Agreement, the other party shall respond within five (5) Working Days to any request by the receiving party for:

- 10.11 Assistance in determining how to respond to a Request for Information; and/or
- 10.12 The provision of information held by the other party pursuant to this Agreement.
- 10.13 Where either party receives a Request for Information which relates to information held pursuant to this Agreement, prior to responding to such request, that party will consult with the other party in relation its response and will allow the other party to make representations with respect to:
- 10.14 Any information which the other party regards to be confidential; and/or
- 10.15 Any similar or identical request received by the other party and (where applicable) the other party's response to that request.
- 10.16 Notwithstanding the provisions of clauses 10.9 to 10.13 (inclusive) each party acknowledges that the decision on whether any exemption applies to a request for disclosure of recorded information is a decision solely for the party receiving a Request for Information.
- 10.17 The parties acknowledge that, except for any information which is exempt from disclosure in accordance with the provisions of the FOI Legislation, the content of this Agreement is not confidential.

11. Data protection

- 11.1 The parties shall comply with the Data Protection obligations which are set out in Schedule 2 to this Agreement.
- 11.2 Notwithstanding the general obligation in clause 11.1, the Researcher shall:
- 11.2.1 Provide the GMC with such information as the GMC may reasonably request to satisfy itself that the Researcher is complying with its obligations under the Data Protection Legislation;
- 11.2.2 Promptly notify the GMC of any breach of the security measures to be put in place pursuant to this clause; and
- 11.2.3 Ensure that it does not knowingly or negligently do or omit to do anything which places itself or the GMC in breach of obligations under the Data Protection Legislation.

12. Warranty, indemnity and insurance

- 12.1 The Researcher warrants, represents and undertakes that:

12.1.1 It has full capacity and authority to enter into and to perform this Agreement and that this Agreement is executed by a duly authorised representative of the Researcher;

12.1.2 It will obtain full Ethical Approval prior to the date at which such Ethical Approval is required in order to commence or continue to provide the Approved Research or any part thereof;

12.1.3 It has obtained and will maintain all necessary rights, consents, authorisations to perform its obligations in accordance with this Agreement;

12.1.4 It will undertake the Research with all the skill, care, diligence and foresight that would be expected from a Researcher, to a high standard including (without limiting the generality of this clause) in accordance with the Researcher's own established internal procedures;

12.1.5 All personnel assigned will have the requisite training, experience, qualifications and knowledge necessary to carry out the tasks assigned to them and in so doing will adopt reasonable and proper standards of behaviour.

12.1.6 The Researcher will, for the entire duration of the Agreement (including any extension period) and for one year following the end of the Agreement (including any extension period), insure and keep itself insured with a reputable insurance company against all insurable liabilities under this Agreement and in respect of the Approved Research including, without limitation against all the Researcher's liabilities under this clause 10.

13. Dispute resolution

13.1 If any dispute arises out of this Agreement the dispute shall be referred to the [insert Researcher Senior Representative] and the Assistant Director, Research of GMC who will attempt to settle it by negotiation. If the parties are unable to settle any dispute by negotiation within ten (10) Working Days the parties may elect to refer the dispute to mediation or an alternative form of dispute resolution. However, nothing in this clause shall prevent the parties commencing or continuing court proceedings at any time.

14. Policies and Standards

14.1 When performing the Approved Research, the Researcher shall comply with the policies and standards set out in Schedule 3 ("the Policies and Standards") as amended from time to time in the event that there is a conflict between the Policies and Standards and this Agreement the terms of this Agreement shall take precedence.

15. Assignment and sub-contracting

15.1 The Researcher shall not assign, transfer or otherwise deal with any of its rights or obligations under this Agreement, or sub-contract the performance of any of its obligations under this Agreement without the prior written consent of GMC.

15.2 Any sub-contract entered into by the Researcher shall be on no less onerous terms than those set out in this Agreement, provided that no sub-contracting by the Researcher shall in any way relieve the Researcher of any of its responsibilities under this Agreement.

16. Publicity, media and official enquiries

16.1 Without prejudice to the GMC's obligations under the Freedom of Information Act 2000, the Researcher shall not make any press announcement or publicise this Agreement or any part thereof in any way, except with the prior written consent of the GMC and other contributing organisation, as appropriate.

16.2 The Researcher shall take reasonable steps to ensure that their servants, employees, agents, sub-contractors, Researchers, professional advisors and consultants comply with clause 12.1.

17. Audit

17.1 The Researcher shall keep and maintain until six years after the end of the Agreement, or as long a period as may be agreed between the Parties, full and accurate records of the Agreement.

18. Notices

18.1 Any notice given by one party to the other under this Agreement must be in writing and may be delivered via email, personally or by pre-paid first class post and in the case of post will be deemed to have been given two Working Days after the date of posting. Notices shall be delivered or sent to the addresses of the parties on the first page of this Agreement or to any other address notified in writing by one party to the other for the purpose of receiving notices after the date of this Agreement. Each party may specify by notice to the other a particular individual or office holder to whom any notices served on it are to be addressed, in which case a notice shall not be validly given unless so addressed.

19. Entire agreement

19.1 This Agreement sets out the entire agreement and understanding between the parties in respect of its subject matter and supersedes all former warranties, statements, representations, understanding, undertakings and agreements (in each case whether written or oral) made by or between the parties relating to such subject matter and including without limitation the terms set out in any document produced by the

Researcher, notwithstanding the terms of any previous agreement or arrangement expressed to survive termination.

20. Governing law and jurisdiction

20.1 This Agreement and any matter arising from or in connection with it shall be governed by and construed in accordance with English law.

20.2 Each party irrevocably agrees to submit to the non-exclusive jurisdiction of the English courts over any claim or matter arising from or in connection with this Agreement.

In witness whereof the parties have executed this Agreement the day and year first above written.

Schedule 1

Part A – Data Specification

Schedule 2

Data Protection

Background

In order to perform the Approved Research on GMC's behalf, the Researcher may require access to certain Data (as defined below) to be made available to it by GMC. The main part of this Agreement sets out the basis on which such access will be provided the arrangements for such access and some of the conditions on the grant of such access. The provisions below further regulate the provision and use of the Data, with the aim of ensuring that access to and use of Data under this Agreement complies fully with the Data Protection Legislation.

1 Definitions and interpretation

1.1 In this Schedule 2 (unless the context otherwise requires), the following words and phrases shall have the following meanings:

Legislation: the General Data Protection Regulation and the Data Protection Act 2018.

Data Controller: the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.

Data Processor: natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

Personal Data: any and all personal data (as such term is defined in the Legislation) provided or made available (whether directly or indirectly, including the Data provided through the Safe Haven, and in whatever form) by the GMC to the Researcher pursuant to this Agreement. This will include data provided in a pseudonymous format and any further information relating to individuals which one might reasonably expect to regard as confidential.

Data Subject: an identified or identifiable natural person.

European Economic Area: The EEA countries are currently the EU countries plus Iceland, Liechtenstein and Norway and for the purposes of this Agreement will include any country on the EU Commission's list of countries or territories providing adequate protection for the rights and freedoms of data subjects in connection with the processing of their personal data.

2 Obligations of the GMC

The GMC shall provide access to Personal Data to the Researcher through the Safe Haven together with such other information as the Researcher may reasonably require in order for the Researcher to perform the Research.

3 Obligations of the Researcher

3.1 The Researcher will be a Data Controller and they are entirely responsible for full compliance with their obligations under the Legislation.

3.2 The Researcher shall:

3.2.1 Implement appropriate technical and organisational measures to protect Personal Data against unauthorised or unlawful processing and against accidental loss, destruction, damage, alteration or disclosure. Such measures shall be appropriate to the harm that might result from unauthorised or unlawful processing or accidental loss, destruction or damage to Personal Data and to the nature of Personal Data to be protected;

3.2.2 comply with all applicable laws, regulations, regulatory requirements and codes of practice, in connection with its obligations under this Agreement, including without limitation by complying with all the provisions of the Legislation and not do anything which may result in a breach by the GMC of the same; and

3.2.3 promptly notify the GMC if it receives a request from a Data Subject to have access to Personal Data within the data accessed under this Agreement or any complaint or request relating to the GMC's obligations under the Legislation and provide full cooperation and assistance to the GMC in relation to any such complaint or request (including, without limitation, by allowing Data Subjects to have access to their Personal Data if agreed by the GMC).

3.3 The Researcher shall not process Personal Data outside the European Economic Area or transfer Personal Data outside the European Economic Area without the advance agreement of the GMC's Information Policy Section that it will be adequately protected.

3.4 At the reasonable request and expense of the GMC, the Researcher shall assist the GMC in complying with its obligations in its activities as a Data Controller.

3.5 In this Schedule 2 (unless the context otherwise requires) an obligation in this Agreement on the Researcher to do, or refrain from doing, any act or thing shall include an obligation to procure that its employees, agents and sub-contractors (if any permitted) also do, or refrain from doing, such act or thing. This includes ensuring that any subcontractor or agent shall comply with obligations equivalent to those of the Researcher.

Schedule 3

Policies and Standards

Data Protection and Information Security from the researcher's host organisation.

Execution

Signed by)

for and on behalf of)

General Medical Council)

Signed by)

for and on behalf of)

[the Researcher])

Appendix C - Data User Agreement to access the HIC Safe Haven for approved UKMED Research Projects



Data User Agreement to access the HIC Safe Haven for approved UKMED Research Projects

Data User Responsibilities

The Approved Data User is reminded of the sensitive nature of the data being accessed and is required to maintain the security and confidentiality of the Project Dataset in accordance with this HIC Data User Agreement and the GDPR Data Protection Principles (attached in Appendix (i) forming a part of this Agreement). HIC requires Approved Data Users to report inadvertent events that are in breach of the terms of the HIC Data User Agreement to enable improvements to be made. Contact the HIC Governance Manager in the first instance to report the incident, who will initiate a Significant Event Report and report any outcome to the GMC.

The Approved Data User:

- 1) Will not reuse the Project Dataset for purposes outside the scope of the Project; attempt to link it to other datasets; or to de-anonymise it.
- 2) Agrees to ensure that individual-level data is not transferred outside the Safe Haven via any means including, for example (not exhaustive); photographic, recording, screen grabbing and note taking. If you are in doubt please contact HIC for further advice and assistance.
- 3) Must never share or disclose their Safe Haven login details.
- 4) Must never allow people who do not have approved access to a Project to see individual level data from that Project within the Safe Haven.
- 5) Must make every effort to stop people who do not have personal access to a Project viewing data about that Project on the Safe Haven screen (e.g. refraining from using via a laptop in public spaces).

When the Project is complete the data and analysis syntax used will be archived by HIC.

Approved Data Users may utilise existing University of Dundee agreements for the use of a set of standard statistical software packages within the Safe Haven for non-commercial academic studies, where the Approved Data User works for an academic institution. Due to the different licensing rules of statistical software providers different terms and conditions will apply. Statistical software access for each study and the terms and conditions of each package can be confirmed with the Manager of HIC's Data Linkage Service.

The Approved Data User is *(tick as appropriate):*

- An academic
- A non-academic

The Approved Data User works for:

- An academic institution
- A commercial institution
- Other (please specify) _____

Glossary

Project: means a unique study with a PI, specified cohort, aims and methods that are logged into the HIC Project Management System and all required governance approvals have been provided to GMC.

Project Dataset: means the data that has been anonymised uniquely and is specifically for use within a Project. The dataset must relate to the cohort and purpose defined for the Project and is placed in the Safe Haven Project folder to be accessed by an Approved Data User. A signed Data User Agreement will permit an Approved Data User to access Project Datasets, as authorised by a Project PI and recorded on the HIC Project Management System.

Signatures

a) Declaration by Applicant for Approved Data User Status

By signing and dating below (but not as a party to this HIC Data User Agreement) you hereby declare and confirm that you understand and agree to the terms of this HIC Data User Agreement and the Data Protection Principles which will govern the Project Dataset(s) being provided to the Institution from GMC which you are applying to access.

Name: _____ ("the Applicant")

Position: _____

Institution: _____

Signature _____ Date signed: _____

b) Applicant's Organisation

Approved Data Users must have this section signed by an authorised signatory from their organisation, the Institution named below.

An authorised signatory for the project collaborator's Institution is to sign the following declaration: The Institution named below hereby agrees that the Applicant named in section a) above is a bona fide employee or student of this Institution engaged in a reputable Project for which all relevant required permissions have been granted, and that the Project Dataset requested can be entrusted to this person in the knowledge that they will conscientiously discharge their obligations in regard to the confidentiality of the Project Dataset. This Institution agrees to abide by the terms of this HIC Data User Agreement and shall take responsibility for ensuring that the proposed Approved Data User complies with the terms of this HIC Data User Agreement, Appendix A and all applicable statutory and regulatory permissions and Data Protection requirements. The Institution agrees to provide a secure working environment and suitable technical resources to meet this obligation.

The Institution agrees that a breach of this HIC Data User Agreement may lead to the withdrawal of access to the HIC Safe Haven for this Institution, its staff and students, and that GMC and HIC have a duty to report serious legal or regulatory breaches to the appropriate authorities (such as the Data Protection Commissioner and professional regulatory bodies).

Name: _____

Position: _____

Signature _____ Date signed: _____

For and On behalf of: _____
(Name of Institution)

c) GMC Data User Authorisation

GMC authorised signatory confirming the above Applicant is approved to access GMC data within the HIC Data Safe Haven. Specific Project access approval is provided separately, per Project.

Name: _____

Position: _____

Signature _____ Date signed: _____

Appendix(I)The GDPR Data Protection Principles

Article 5 of the GDPR sets out seven key principles which lie at the heart of the general data protection regime.

Article 5(1) requires that personal data shall be:

“(a) Processed lawfully, fairly and in a transparent manner in relation to individuals (‘lawfulness, fairness and transparency’);

(b) Collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall not be considered to be incompatible with the initial purposes (‘purpose limitation’);

(c) Adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (‘data minimisation’);

(d) Accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay (‘accuracy’);

(e) Kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes subject to implementation of the appropriate technical and organisational measures required by the GDPR in order to safeguard the rights and freedoms of individuals (‘storage limitation’);

(f) Processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures (‘integrity and confidentiality’).”

Appendix D Software arrangements as per December 2023

Software vendor	Product	Version	License arrangements
IBM	SPSS Statistics Premium	22	Please note we only have licenses for 2 * Professional and 2 * Standard versions of SPSS to run concurrently.
Open Revolution R	Revolution R Open 3.5.1 (Community Edition)	3.5.1	Included as standard
RStudio	R Studio Desktop (Open source edition)	1.1.463	Included as standard (for academics only). Users must be academics working for a recognised academic institution. Commercial projects, even when resourced by academic staff, are NOT covered by HIC's academic licencing agreements. Users would need to buy their own commercial licence.
SAS	SAS	9.4 TS Level 1M2 X64_7PRO	Included as standard (for academics only). Users must be academics working for a recognised academic institution. Commercial projects, even when resourced by academic staff, are NOT covered by HIC's academic

Software vendor	Product	Version	License arrangements
			<p>licencing agreements. Users would need to buy their own commercial licence.</p> <p>HIC are providing this software by agreement with the software provider based upon HIC's current organisational software licences. However, these agreements are subject to change outside of HIC's control and therefore the benefit of this licence could be revoked and any time.</p>
<p>STATA</p> <p>Plugins must be requested individually for installation</p>	<p>STATA MP14 (2 concurrent users only). Once the max number of users is reached the application will no longer be available. If this becomes an issue then a user could buy their own licence.</p> <p>STATA 15</p>	<p>64 bit 3-Aug-2015</p>	<p>Included as standard (for academics only).</p> <p>Users must be academics working for a recognised academic institution.</p> <p>Commercial projects, even when resourced by academic staff, are NOT covered by HIC's academic licencing agreements. Users would need to buy their own commercial licence.</p> <p>HIC are providing this software by agreement with the software provider based upon HIC's current organisational software licences. However, these agreements are subject to change out of HIC's control and therefore the benefit of this licence could be revoked at any time.</p>
<p>Notepad-plus-plus.org</p>	<p>Notepad ++ (Open Source)</p>	<p>Build Time 01/10/15</p>	<p>Included as standard</p>

Software vendor	Product	Version	License arrangements
Apache OpenOffice	OpenOffice	4.1.1	Included as standard
Microsoft	SQL Server Management Studio		Included as standard.
Microsoft	Windows	7	Included as standard
Libra Office			Included as standard
University of Bristol	MLWin	2.35	Included if requested (for academics only)
Variable	Specific application software, not on this list	Variable	Subject to HIC approval, HIC will require 5 working days to install on receipt of the software and license key. A fee of £250 is payable for this. All software must be Windows based. HIC are unable to support UNIX or Linux versions of applications.

Figure 1 Process for completing UKMED Research

