Schedule of Retention Periods for Research Data

Contents

1. PURPOSE 2
2. EXCLUSIONS 2
3. DEFINITIONS 2
4. RESPONSIBILITIES 2
4.1 DATA RETENTION SERVICES 3
4.1.1 KCL Infrastructure 3
4.1.2 External storage options 3
4.2 When researchers leave 3
4.3 Intellectual Property Ownership 4
5. DATA RETENTION 4
5.1 Judgment necessary 4
5.2 Open Access/ Open Research 4
5.3 Clinical Trial Data 5
6. PERSONAL DATA AND UK DATA PROTECTION LAW IN RESEARCH 5
6.1 UK Data Protection Law principles: 5
7. RETENTION OF PERSONAL (IDENTIFIABLE) RESEARCH DATA: 6
8. RETENTION OF CLINICAL HEALTH RESEARCH DATA: 7
8.1 Medical records 7
8.2 Clinical Trials with Marketing authorisation 7
9. TABLE OF MINIMUM RETENTION PERIODS FOR RESEARCH DATA 8
10. RESEARCH ABANDONED BEFORE PROJECT COMPLETION 14
11. RETAINING DATA BEYOND THE MINIMUM PERIOD 14
11.1 Highly significant research data 14
11.2 When results are challenged 14

1. PURPOSE

The King's Schedule of Retention Periods for Research Data is intended to guide researchers, both students and staff, in deciding on retention periods for research data.

2. EXCLUSIONS

Where research is funded, the funder may have a policy relating to the retention of research data collected as part of the given project. This schedule does not usually supersede these policies. Where a funder does not provide specific retention guidelines, this policy may be followed.

3. DEFINITIONS

'Researcher' or 'researchers' refers to the lead researcher on a given research project, or the primary research supervisor in the case of undergraduate and post-graduate research students.

'Research data', or 'data' refers to any representation or other objects that are created or gathered for the purposes of producing research or scholarship, and which can be used to validate or reproduce original research findings. This definition aims to be technology neutral, omitting any reference to the format in which the research data are collected, created, stored etc.

Research data can cover a diversity of form and content, including (but not limited to) numbers, text, images, audio, simulations, models, interview recordings, questionnaires, laboratory notebooks, videos, algorithms, codebooks, test results, specimens, chemical samples, biological materials, human tissue, historical documents in archives, databases, or any combination of these. Many factors will affect whether it is feasible or legal to retain physical objects used in research. Factors include the nature and durability of the materials, ownership, agreements with sponsors and funders, practices in the discipline, and legislation on matters such as hazardous materials, animal protection, use of human tissue, health, and safety. It is not possible to list all types of material that might be used in research and prescribe retention periods.

4. RESPONSIBILITIES

University

KCL is responsible for the development and delivery of in-house IT infrastructure, digital hosting and storage arrangements, including the provision of maintenance and support for these systems and components used by King's researchers in support of all relevant policies and procedures, including but not limited to this retention schedule, the KCL Research Data Management Procedure and funder policies.

Researchers

Researchers are responsible for the forward planning of long-term data retention requirements of the research project before the project begins, as well as determining the

most suitable data retention service (whether in-house or external) for the research data, usually through the creation of a Data Management Plan.
The responsibility of determining suitable data retention options falls to the Researcher. Further guidance on external data storage options can be found within the KCL Research Data Management Procedure document.

4.1 DATA RETENTION SERVICES

4.1.1 KCL Infrastructure

Where research data are retained within the provided KCL infrastructure, the University is responsible for meeting the retention requirements laid out in this schedule, including any other relevant conditions.

4.1.2 External storage options

Where research data are not retained within the provided KCL infrastructure, the chosen storage location must meet the conditions that comply with security, safety, privacy, and confidentiality requirements of any ethical clearance the project was granted, funding agreements, data licence agreements, and any further provisions set out under KCL policies. Where data are stored in an identifiable format, consent must be obtained from participants to store the data offsite.

In such cases, as per the KCL Research Data Management Procedure, a metadata record should be made publicly available non-exclusively via the King's Research Data Management Service, irrespective of where the actual data are hosted, except in cases in which the research project and metadata describing it is itself confidential.

4.2 When researchers leave

When researchers leave, the University will no longer be responsible for the storage of their research data and primary materials, unless:

- the research data has been deposited within KCL provided IT infrastructure or digital hosting services; or
- an agreement to this effect has been made between the researcher and the University; or
- a funding or other agreement to this effect has been made between the University and a third party; or
- the University is the legal owner of the research data or the intellectual property in the data.

In cases where the University is no longer responsible for the storage of said research data, the Researcher will be responsible for organising the ongoing custodial responsibilities or continuing storage, management, and disposal of the retained data for the minimum period as indicated by this schedule, under conditions that comply with security, safety, privacy, and confidentiality requirements of any ethical clearance the project was granted, any funding agreement, and any data licence agreement. The subsequent responsibilities may fall to supervisors, heads of department or research units, depending on the type of research in question.

4.3 Intellectual Property Ownership

University policy on the ownership of intellectual property in research data and other outputs of research is found in the Code of Practice for Intellectual Property, Commercial Exploitation and Financial Benefits. The use of University facilities to store research data does not alter the ownership arrangements. For example, if copyright subsists in the research data, and the copyright is owned by a staff member, student, or external organisation, storing the data on University facilities will not of itself cause a transfer of copyright to the University.

5. DATA RETENTION

The University recognises that:
research data are an essential component of any research project involving KCL staff and students regardless of whether the project is supported by external or internal funds.
the availability of accurate and retrievable data is necessary to verify and validate when required, the process and outcomes of the research.
the value and benefits of data re-use, as recommended by the KCL Research Data Management and funder policies.

The Table of Minimum Retention Periods for Research Data at the end of this Schedule prescribes the minimum periods for which research data should be retained for the purposes of the intended research project.

5.1 Judgment necessary

While this policy prescribes minimum retention periods for research data, it is important to evaluate the research data on a case-by-case basis to determine the value and risk of retaining such research data.

In cases where it is feasible and legal to retain research data, the guiding principle is to retain only sufficient data to justify the published or reported outcomes of the research and to enable the researcher to defend the outcomes if they are challenged. However, there may be additional considerations, such as the requirements of funder policies, open access/ open research principles, the cost-effectiveness and availability of data storage options, the need to satisfy legal obligations or reduce litigation risk, as well as the nature of the data itself (for example, data that by their nature cannot be re-measured) (see also section 11: Retaining data beyond the minimum period).

It may be less important to retain data from research projects that do not result in dissemination of findings, because their findings, if any, are unlikely to be challenged or arouse any interest. However, as above, this should be assessed on a case-by-case basis.

5.2 Open Access/ Open Research

As research data sets are increasingly being published, both to support and validate published research findings in journal articles, and as recognised research outputs in their own right to be re-used and cited in the move towards ‘open research’, there is the need for them to be retained, preserved and made accessible for the long term. A common retention period for published data sets is 10 years, but with an aspiration to go beyond this, even to maintain perpetuity when data sets are deemed to have long term value to the research

community. In light of this, Open Access/ Open Research should be taken into consideration when determining the appropriate retention period for research data, alongside any recommendations present within the Research Data Management policy and procedure and those of any funder policy, as outlined in the section 2. Exclusions.

5.3 Clinical Trial Data

Whether the findings are disseminated or not, data from clinical trials are a special category. Clinical trial data are preserved for longer periods than is usually necessary for other kinds of research, in case evidence emerges of unexpected late effects.

6. PERSONAL DATA AND UK DATA PROTECTION LAW IN RESEARCH

UK Data Protection Law includes the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018 and applies to identifiable 'personal data'. Personal data are any information where a living person can be identified or who is identifiable, directly from the information in question or, where a person can be indirectly identified from that information in combination with other information. In practice this means that an individual might be identifiable just from the data collected as part of a research project (direct) or, the data collected could be combined with other information, such as newspaper articles or social media content (even if this is not collected as part of the research project), leading to an individual being identified (indirect).

For further guidance on personal data and the UK Data Protection Law for research, please refer to the KCL Research Governance Office's webpages.

6.1 UK Data Protection Law principles:

UK Data Protection Law highlights the following 7 key principles which apply to the use and retention of personal data in research. Data must be:

1. processed lawfully, fairly and in a transparent manner in relation to individuals;
2. 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;
3. adequate, relevant and limited to what is necessary for the purposes for which they are processed;
4. 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;
5. 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 the implementation of the appropriate technical and organisational measures required by UK Data Protection Law to safeguard the rights and freedoms of individuals;
6. processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and accidental loss, destruction or damage, using appropriate technical or organisational measures"; and
7. the controller shall be responsible for, and be able to demonstrate, compliance with the principles.

7. RETENTION OF PERSONAL (IDENTIFIABLE) RESEARCH DATA:

Participants must be notified of the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period (UK GDPR Article 15(1d)). Researchers should base the criteria on whether it is necessary to store personal research data to verify and defend, when required, the process and outcomes of the research.
To abide by the fifth Data Protection principle, it may be necessary for identifiable (personal) research data to be anonymised as soon as there is no requirement for the data/information to be identifiable. For example, there may only be a requirement to collect identifiable data from a participant to link it to other data collected from that participant. In this case, data should be anonymised once the data has been linked together (for example, by deleting the identification link between the code used to keep the data confidential and the identification of the participant).

If there is only a requirement for personal data collected from student research to be retained for data authentication purposes up until the degree has been awarded, then the research project data should be deleted or anonymised as far as possible thereafter whilst abiding by the retention schedule for research data.

Advice regarding specific types of personal research project data:
Video and audio recordings are personal data if individuals can be identified from the recordings (e.g. the individual has not been disguised on a video and the voice has not been distorted). Interviews are often audio-recorded and once the recording has been transcribed it might not be necessary to keep identifiable audio recordings. There might be exceptions, such as where interview recordings should be kept for data authentication purposes.

Human tissue samples must be stored as research project data for a minimum period after the project has finished, following publication or degree award. Identifiable human tissue samples must be stored confidentiality so that they can still be traced to the corresponding signed consent form. The human tissue samples should not be anonymised as this would make the samples untraceable, which would breach the Human tissue Act (2004).

Signed consent forms should be kept for as long as the research project data are retained (by the researcher or an archive). As consent forms contain personal data, they should not be archived alongside research data, but should be kept safely under the responsibility of the Researcher. A blank consent form and information sheet should be archived alongside research data to provide background information on the research data. (UK data archive).
Where participants withdraw their consent, a record of the withdrawal including what was agreed regarding withdrawal (e.g. the researcher may retain data already collected, but the participant will not be contacted further etc.) should be retained in the same way as signed consent forms.

Contact details must be stored securely and separate from the research data. Contact details should also only be stored for as long as they are required to administrate the research (e.g. to collect data, reimburse participants, notify of prize draws etc.). With the participant's consent, contact details can be kept for longer periods for specific reasons – for example, so that the participants can be contacted about future research. At the end of the retention

period, identifiable personal research data should be reviewed and deleted, unless there is a justifiable reason for keeping it.

8. RETENTION OF CLINICAL HEALTH RESEARCH DATA:

It is not legitimate to archive clinical health records for historical purposes. If such records are no longer required to provide care, then they should not be retained in identifiable form or used for other purposes, without patient consent or support under Section 60 of the Health and Social Care Act 2001.

Retention periods for clinical trials/ health studies are outlined in the Table of Minimum Retention Periods for Research Data.

8.1 Medical records

Subject's medical records should be retained following applicable legislation and per the maximum period permitted by the hospital, institution or private practice. The documents can be retained for a longer period, however, if required by the applicable regulatory requirements or by agreement with the sponsor. It is the responsibility of the sponsor to inform the hospital, institution or practice as to when these documents no longer need to be retained.

8.2 Clinical Trials with Marketing authorisation

Research data pertaining to Clinical Trials with Marketing authorisation must be retained as per the Commission Directive 2003/63/EC.

9. TABLE OF MINIMUM RETENTION PERIODS FOR RESEARCH DATA

The Table of Minimum Retention Periods for Research Data prescribes the minimum periods for which research data should be retained for the purposes of the intended research project. The table categorises the data by type of research, researcher type, ethical clearance status, and dissemination status as factors relevant to determining the minimum period.

Type of research study/ trial	Researcher Type	Ethical clearance	Dissemination status of research findings	Minimum period of Retention	Responsibility for retention	Based on
Advanced medical therapy research (ATIMP)	Staff	External Ethical Clearance obtained	Disseminated	Advanced medical therapy research Trial Master Files – 30 years minimum following completion of the trial, plus 2 years after last marketing authorisation (this includes paediatric studies). Source data* & Investigator Site File – Completion of trial + 5 years, destroyed only with Sponsor permission.	Researcher	European Medicines Agency guidelines for investigational advanced therapy medicinal products in clinical trials.
NHS clinical trial (CTIMP) 'Clinical trials' includes any studies that involve the investigation of an investigational medicinal product (IMP), including	Staff	External Ethical Clearance obtained	Disseminated	Requirements of specific Trusts should also be considered, as well as those of the Sponsor. For marketing authorisation holders, please refer to Section 6 above. Data should be retained for	Researcher	Medicines for Human Use (Clinical Trials) Regulations Trust requirements

combined IMP/ device studies etc. Refer to Clinical Trials Regulation No 536/2014 for full definitions.				whichever duration is deemed the longest. Where KCL is the Sponsor: • Essential clinical trial documentation will be retained, in accordance with the Medicines for Human Use (Clinical Trials) Regulations, for a minimum period of 25 years following completion of the trial, plus 2 years after last marketing authorisation (this includes paediatric studies), unless a longer retention period is stated in the essential trial documents. • Source data* & Investigator Site File – Completion of trial + 5 years, destroyed only with Sponsor permission.		Sponsor requirements
Other clinical or social care studies not falling under the above categories, where KCL is the sponsor:	Staff (or student projects with a supervisory CI)	External Ethical Clearance obtained	Disseminated	Requirements of specific Trusts should also be considered, as well as those of the Sponsor. For marketing authorisation holders, please refer to Section 6 above. Data should be retained for	Researcher	Medicines for Human Use (Clinical Trials) Regulations Trust requirements

				whichever duration is deemed the longest. Where KCL is the Sponsor: • Research involving children – completion of trial + 5 years or 3 years after the youngest subject reaches 18 years of age (whichever is longest) • Research involving those who lack capacity to consent – 25 years • Other research study data - Completion of trial + 5 years, unless a longer period has been stipulated by the funder and has been approved by the regulatory bodies.		Sponsor requirements
Commercial clinical trial or health research	Staff	External Ethical Clearance obtained	Disseminated	Refer to Commercial Sponsor requirements.	Researcher	Commercial Sponsor requirements.
Other population health studies	Staff	External Ethical Clearance obtained	Disseminated	20 years unless specified otherwise by the funder of the research.	Researcher	As recommended by MRC principles and guidelines for

						good research practice
Healthy volunteer studies, non-clinical or overseas clinical studies that do not require HRA approval	Staff Student – PhD/ MPhil	KCL CREC, all ethical risk levels, including projects where ethical clearance is not required.	Disseminated**	10 years following publication which fundamentally relies on the data, unless specified otherwise by the funder of the research.	Researcher	UKRI Concordat on Open Research Data
Healthy volunteer studies, non-clinical or overseas clinical studies that do not require HRA approval	Student – Taught Postgraduate/ Undergraduate without intention to publish	KCL CREC, all ethical risk levels, including projects where ethical clearance is not required.	N/a	Destroy within 1 month following receipt of award.	Researcher	KCL
Healthy volunteer studies, non-clinical or overseas clinical studies that do not require HRA approval	Student – Taught Postgraduate/ Undergraduate *with* intention to publish	KCL CREC, all ethical risk levels, including projects where ethical clearance is not required.	Disseminated**	**Pre-publication:** Destroy within 1 year following receipt of award. If there is a requirement to retain data for a longer period this must be discussed and agreed with the Research Governance Office rgo@kcl.ac.uk **Post-publication:** At least 5 years following publication which fundamentally relies on the data, unless	Researcher	KCL

				specified otherwise by the funder of the research, or by the journal.		
Animal research: • Project licences • Personal licences • ASPA inspectors' reports • Information mentioned in 2010/63/EC art 30.1 • Records of regulated procedures • Health records of animals • Records of the source, use and disposal of protected animals used in procedures • Lab books	All levels	All	All	5 years, as per legal requirements	Researcher/ BSU	1986 c 14 ss 4(5), 5E(1), 18(2E), Sch 2C para 8(b) Home Office guidance
Animal research: • Records of AWERB advice provided under 2010/63/EC art 27.1 • Records specified in 1986 c 14 s 5G(2)	All levels	All	All	3 years, as per legal requirements	Researcher/ BSU	1986 c 14 Sch 2C paras 6(4), 9(1)(e) 1986 c 14 s 5G(1)

Individual history files of dogs, cats or NHPs which have died or have been set free or rehomed						
Other anonymous research data not covered elsewhere in this schedule	N/a	N/a	Disseminated or non-disseminated	To be reviewed after 10 years following completion of research to determine whether continued storage would be beneficial.	Researcher	As recommended by MRC principles and guidelines for good research practice

*Source Data = All information in original records and certified copies of original records of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies). https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-good-clinical-practice-e6r2-step-5_en.pdf

**Where research data is not disseminated following collection or analysis, and dissemination is not intended, sufficient justification must be made for the continued storage of the research data.

10. RESEARCH ABANDONED BEFORE PROJECT COMPLETION

Where a research project is abandoned before completion (this does not include postponement) research data should be destroyed upon abandonment, unless it is a Clinical Trial (See section 8), or a justification for retention is provided.

11. RETAINING DATA BEYOND THE MINIMUM PERIOD

A guiding principle is that research data should be retained for sufficient time to allow reference to them by other researchers and interested parties; this may be for as long as interest and discussion persist following dissemination.

However, in deciding on extended retention, researchers should consider the potential value for future research. This is especially the case where the research would be difficult or impossible to repeat or where repeating the research would place a significant burden on human participants or animals. For research involving humans, guidance on ethical issues in the re-use of data and its retention beyond the minimum period should be discussed with the approving Research Ethics Committee.

Additionally, in accordance with the UK GDPR, identifiable personal data must not be stored for longer than is necessary for the intended purpose. Therefore, if identifiable personal data are to be stored for longer than the minimum period, this must be justified appropriately in compliance with UK Data Protection Laws. Please also refer to section 6.

11.1 Highly significant research data

If the data have community or heritage value or directly informs national policymaking, consideration should be given to permanent retention, preferably within a national collection or archive. Each research project should be assessed individually to determine whether this data should be kept indefinitely, and it is recommended that researchers discuss this with KCL Archives.

11.2 When results are challenged

If results from research are challenged, all relevant data must be retained until the matter is resolved.