Retention framework for research data and records

1. Background and scope

You may be asked to make decisions about the retention of research data and related material. This may include, but is not limited to, laboratory notebooks, biological samples, images, microscope slides, ethical approvals, questionnaires, participant information sheets and consent forms.

What the framework covers

This framework supports **risk-proportionate decision making**, providing guidance on:

- how to make risk proportionate decisions about retention and destruction
- the importance of documenting and retaining records of those decisions
- where to access further guidance and support

Decisions should not be made in isolation. They should involve appropriate stakeholders (such as Principal Investigators (PIs), sponsors, data managers, or other relevant roles).

Who is this framework for and when does MRC policy apply?

This framework is primarily for UKRI-employed staff, such as staff in MRC Institutes.

If you work in **MRC Head Office or an MRC Institute**, guidance on managing corporate information is available in the UKRI Information Management Policy and the UKRI Retention Schedule. This framework supplements these resources. You can also seek guidance from the Knowledge and Information Management (KIM) Team.

Other MRC-funded researchers, such as those working in an MRC Centre of Research Excellence, may find this framework helpful, but the MRC retention policy is advisory only. Researchers should follow the retention requirements of their employing organisation and/or sponsor(s). For example, researchers employed by a university should follow university and/or sponsor policies.

For MRC-funded units that transferred to a university or closed, MRC retention policy applies to research data and related material generated when MRC was the substantive employer or to clinical research that was sponsored by MRC-UKRI. For data and material generated after transition, university and/or sponsor policies apply.

For units that close, arrangements for data management, including any retention requirements, are likely to be set out in the unit closure or transfer agreement. If no arrangements are in place, the KIM Team can advise.

Further guidance is also available on MRC's Data Management and Sharing webpages.

Why retention matters

Retaining research data and materials is important because it:

- supports Open Research, reproducibility and transparency
- enables further research through data and sample sharing
- provides an audit trail of the research conducted within your Institute
- preserves research of historical significance or importance

However, long-term retention (including any sample storage or biobanking) has **cost and resource implications**. This framework therefore promotes a **risk-proportionate approach**, focusing retention efforts on 'high-risk' or high-value research while applying a lighter touch approach to lower risk research.

Examples of 'high-risk' or high-value research may include:

- Studies where there is a greater likelihood that data and/or samples will be shared (for example, to support further research or reproduce high-profile, high-impact, or potentially contentious findings)
- Research which has generated protected Intellectual Property (IP), or there is potential for IP to be licensed
- Research which involves an invasive medical intervention and/or vulnerable groups

2. MRC retention policy for research data and related material

MRC policy complements statutory and regulatory requirements such as the Clinical Trials Regulations and the UKRI Retention Schedule.

Minimum retention periods (see section 1 for applicability)

Basic research – **at least 10 years** after study completion[1]

Population health and clinical studies – **at least 20 years** after study completion

From 28 April 2026, the Clinical Trials Regulations require minimum retention periods of:

- **25 years** for Clinical Trials of Investigational Medicinal Products (CTIMPs)
- **30 years** for Advanced Therapy Medicinal Products (ATMPs)
- Further guidance: Archiving and retention of clinical trial records (MHRA)

Studies proposing retention periods beyond the minimum limits stated above must provide a valid justification. However, longer retention periods may be appropriate where:

- The research is likely to be reused or reproduced because it is high impact, high profile, potentially contentious, or of value to the wider research community
- Protected IP exists or may arise, meaning certain records (such as laboratory notebooks or key clinical study documentation) may need to be retained for extended periods, potentially indefinitely
- For research involving an invasive medical intervention and/or vulnerable groups (for example due to age or disability), a minimum retention period of 25 years may be more appropriate
- Clinical studies involving pregnancy or children should consider retaining records until the child reaches the age of 25, where this would exceed the retention periods specified above
- The research informs national policymaking or is of historical significance or importance. In these circumstances permanent preservation should be considered

Further guidance on reviewing what you keep, and decision making is provided in Annex 1.

[1] Concordat on Open Research Data

Active management

There should be a clear and transparent system for managing the lifecycle of research data and samples.

There should be an identified person (or persons) responsible for managing all research data and/or samples (such as the PI or Trial Steering Committee, or a librarian, Information or Data Manager, Designated Individual or equivalent)

A register, database or inventory should be in place to manage research data and/or samples. This should include at a minimum storage location, time of data collection, type of data and any decisions with respect to destruction

It should be easy for authorised users to find and access data and/or samples (via a catalogue or register including relevant metadata). This is particularly important for sample and data sharing.

Wider research policy and regulatory landscape

Research organisations will have their own data and records retention policies to meet their responsibilities under data protection[2] and research governance[3]. You must adhere to the policies relevant to your research.

Similarly, collaborators, funders, journals and others may have their own requirements. You must understand and meet the policies of **all** relevant organisations involved in your research (such as those providing regulatory oversight, collaborators, funders or journals) and consider how these expectations affect research data and/or sample retention. In some cases, these expectations may be set out in an agreement.

For example, organisations providing regulatory oversight may audit or inspect your premises (such as the Health & Safety Executive, Home Office, Human Tissue Authority, Medicines and Healthcare products Regulatory Agency). Be aware of the research data and/or samples they are likely to require and ensure that you can provide access to these.

If your research involves the NHS, they will have their own expectations and may check compliance. Be aware of these expectations, ensure that you can meet them, and consider any long-term requirements (for example, how long do they expect you to retain any research data and/or samples, and how will storage responsibilities and costs be managed).

3. Deciding what, how and where to keep

This section provides practical guidance to support decisions about what data and/or samples to retain, how they are stored, and how they are managed over time. The scenarios below illustrate how these principles can be applied in certain situations.

---

[2] In terms of primary data protection law, researchers can keep data indefinitely for research (subject to appropriate safeguards). If research data are no longer useful, or unlikely to be used in further research, then you should consider their destruction.

[3] Outlined in the UK Policy Framework for Health and Social Care Research. In addition, Archiving and retention of clinical trial records (MHRA) outline requirements for CTIMPs.

What to keep

Research projects generate and use a wide range of data and/or related material. What to retain will depend on the nature of your research, and what is needed to understand what was done when, how and why.

These decisions should be led by someone familiar with the research (such as the PI or Trial Steering Committee), and should align with organisational policy and sponsor expectations where relevant.

For some **basic research,** laboratory notebooks may provide sufficient detail to understand what was done when, how and why. For some **clinical studies**, much more data and/or records will be required to achieve the same understanding.

In some instances, for example if you've used a Secure Data Environment (SDE), you may not hold the research data directly. However, a persistent identifier, such as a Data Object Identifier (DOI) or the mechanism to recreate the dataset, should be retained. We use the term mechanism to mean the criteria or parameters used to create your research dataset, be that a subset of a larger dataset or one created from multiple sources.

Decisions about what to keep should be made on a case-by-case basis, focusing on retaining the data and/or samples needed to support open research (including data sharing), provide an audit trail, and meet regulatory or governance requirements.

You may wish to consider the uniqueness of any materials created as part of your research. Most publishers recommend that researchers should make unique research materials available to others, particularly where these can't be recreated.

Where relevant, decisions must align with participants' expectations. For example, if a participant withdraws from a study, retention of their data and/or samples should reflect what was agreed at the point of withdrawal.

A risk proportionate approach should be taken to decisions about retention and long-term archiving. There should be a proportionate assessment of risk to identify studies that may require additional management. For example, you may decide to manage high-risk studies by retaining data and/or samples for longer in their original format. That's not to say that **all** research data from high-risk studies (or indeed any study) should be kept for longer or in their original format, as the scenarios below demonstrate:

Scenario 1: research outcome

For some basic research, decisions about what to retain may be influenced by research outcome. A record of what was done will always be kept in laboratory notebooks. Where experiments fail (rather than generate negative results), the original outputs may have limited long-term value and may not need to be kept. However, the decision should be documented so the fate of the original outputs can be tracked. You may also consider retention of original outputs in the short term for protocol optimisation reasons.

Some original outputs take up a lot of physical space (such as plates and cell cultures). Plates are typically captured using a plate reader with images printed or saved electronically. The original plates are therefore no longer the sole record. Keeping a cell culture may also be less important if the initial biological starting material and any resulting cell line is preserved. As such, cell cultures and plates would usually only be kept if deemed useful and/or unlikely to deteriorate.

Similarly, not all data created as part of research needs to be kept. Intermediate data (data created during processing) may not need to be retained where it can be reliably reproduced from the retained raw data, metadata and any protocol or software used to create it. Some data processing workflows will destroy intermediate data immediately.

Scenario 2: fragile outputs

Some outputs can deteriorate over time. Gels and immuno blots are typically photographed, with the originals then discarded because they are too fragile to preserve long term. Stained microscope slides may be retained after analysis but may not be worth keeping in the long term as the staining can fade over time (and slides can take up a lot of physical space).

Scenario 3: paper records

Consider whether original paper records need to be retained longer term. If you were to recruit healthy volunteers to a volunteer database, you might ask them to complete a paper form to capture some personal details (such as name, address, contact telephone number) as well as the types of research in which they might like to take part. These details could then be entered into a digital database and the original forms stored for reference.

Is it worth keeping these original paper forms in the longer term? The information that you need is now stored in the database. Therefore, as long as there is some form of quality assurance process to verify that the data in the database is accurate and valid, the original paper forms can be destroyed after an agreed time period. The quality assurance process and the decision to destroy the forms should be approved, documented and retained for future reference, providing an audit trail of the fate of all your research data and/or samples.

It's also worth considering the risk of keeping the original paper forms in the longer term and balancing this against the risk of destruction. The forms contain 'personal data'[4] and are therefore subject to data protection law and confidentiality requirements. Storing the original paper forms in the longer term, particularly if this means storing them with an external archiving organisation, presents a potential risk to your volunteers' privacy as well as a cost burden.

Specific considerations for biological samples

Existing collections of samples can have significant value for research if they are of known quality and are linked to accompanying data. However, long-term maintenance and storage of samples require resource. Samples should therefore be assessed periodically to ensure they remain 'fit for purpose,' considering both their scientific value and any ethical issues. Samples that are no longer of value should be disposed of safely, respectfully and sensitively, in line with any legal requirements and local policy. This should include documenting the disposal decision and retaining it for future reference, so the fate of the samples can be tracked.

[4] What is personal data? | ICO

What format – physical versus electronic

The principles illustrated in Scenario 3 (paper records) can be applied more broadly, as any research data may be collected in physical form (such as biological samples, microscope slides, paper records) and/or electronically.

Where both a physical and electronic version of the same data exists, decisions can be made about whether to keep both formats long term. These decisions should be made by someone who is familiar with the data (such as the PI or Trial Steering Committee), in line with organisational policy and, where appropriate, sponsor expectations.

Converting original paper records into electronic format (for example, by data entry or scanning) is also an option. In principle, either approach can help reduce long-term paper archiving or sample maintenance costs. However, neither comes without risk nor cost as explored below.

Scenario: Scanning

An MRC researcher wished to replace paper screening and consent forms for MRI participants with scanned electronic copies. This would meet MRC policy provided there is a documented process to ensure the scanned forms are of appropriate quality, stored securely and subject to quality assurance checks. Once the process was approved and recorded, the originals could be destroyed. The destruction decision should be documented and retained for future reference to provide an audit trail of the fate of the original forms.

Decisions about format involve balancing risks and practical considerations:

- How likely is it that the data will be needed in future (for example for audit or review). Requests are more likely for high-risk or high-value research
- Would it be easier to provide access in physical or electronic format?
- What are the risks associated with long-term storage in each format?

Paper can take up more physical space than its electronic counterpart. Whilst paper and electronic records are both vulnerable to loss, accidental deletion or destruction (for example by fire, flood), electronic records can be backed up more frequently and across different storage methods (such as secure networks and external hard drives).

However, it's important to note that converting to electronic format is not without cost (for example, scanning time, server space, and environmental impact). To support environmental sustainability, consideration should be given to low-carbon storage solutions and backup arrangements that avoid unnecessary duplication.

Electronic records also need active management to ensure they remain accessible over time. This includes maintaining audit trails, tracking and ensuring security of backups, and regularly checking that files can continue to be accessed and read as software and hardware technology advances.

What are the risks of transferring paper records to electronic format?

A process of checking and validation is essential, as scanned records have limited value if the scanning process hasn't captured all of the paper records or these haven't scanned as intended and you've got poor-quality or incomplete copies.

There is also the possibility for scanned records to be amended (either deliberately or accidentally). Therefore, access to these should be controlled by an appropriate person.

Where to keep, who manages it, and how access is provided

Decisions about where data is kept are closely linked to how it will be managed and accessed over time. Initially, research data and/or samples will be collected and retained under the supervision of the PI or Trial Steering Committee (within the PI's employing organisation, Clinical Trials Unit(s) and/or other sites involved in the research).

At an appropriate point in the research lifecycle, a decision may be taken to store and curate data and/or samples centrally. Organisations should have a policy or process on when and where this should occur. The use of data repositories is becoming increasingly common to manage legal and technical requirements and support research data sharing. UKRI's commitment to open research promotes making research data openly available wherever possible.

When deciding where data will be kept, consider:

Storage and location:

- whether the research team will retain some or all of the data locally
- whether the data will be transferred to a repository or another organisation
- whether any clinical data will form part of the patient's medical record

Responsibility and management:

- will the research team maintain data and associated metadata in a useable format?
- whether agreements are required with third parties (such as a data repository or partner institution) to manage data on your behalf
- how data and metadata are maintained to support reproducibility and future research

Access and sharing:

It's important to consider how data sharing requirements will be met. This includes decisions about access, how participant privacy will be protected, how commercial sensitivities will be respected and the type of data services that will be provided. If a data repository is used, many of these considerations may be managed by the repository.

Research materials may also need to be accessed for other purposes, such as regulatory inspections, audits or investigations, and these should be considered when making decisions about retention and storage. For research involving human participants, you should also work with relevant Data Protection team(s) to decide how subject access requests will be handled.

Storage of biological samples

To maintain quality and optimise sample use, samples can be transferred to a central tissue resource or 'biobank.' In England, Wales and Northern Ireland research 'biobanks' are subject to the Human Tissue Authority's licensing scheme, overseen by a Designated Individual (or equivalent for unlicensed premises). The Designated Individual is a key contact for sample management. Samples are often only as valuable as any accompanying data; the following section is relevant where any accompanying data is identifiable.

Identifiable information and consent

In multi-site studies, participants' identifiable information will usually remain at the site where they were recruited and/or treated. However, some identifiers may be shared with a coordinating centre (or biobank) to enable future linkage or participant follow-up. Participants should be informed of, and agree to, these arrangements.

Completed consent forms contain identifiable information and are therefore typically retained at each site. Sites may incur archiving costs as a result, which should be identified early, with clear arrangements made for how they will be covered.

When sharing data and/or samples with others, sufficient information must be provided to give assurance that sharing is consistent with participants' expectations and that no restrictions prevent further sharing. This may be achieved by sharing copies of the relevant versions of template consent forms, participant information sheets and/or research ethics application. For larger collections (such as data and/or samples collected in a clinical trial or longitudinal population study), this may instead be achieved by providing records that log the consents in place or track and identify the relevant consent paperwork. In all cases, an appropriate legal agreement should be put in place.

When archiving identifiable information outside the controlled environment in which it was originally collected, the PI or the Trial Steering Committee should consider any additional risks. With guidance from the sponsor(s) and relevant Data Protection Officer(s), they should ensure that:

- only the identifiable information necessary for retention is kept (a risk-based decision)
- appropriate data security measures are in place

General requirements for storage

Wherever research data and/or samples are stored:

- Their security must be ensured
- They must be organised so they can be easily identified and located
- They must be accessible for audit, inspection, or review where required

If external storage providers are used, you must ensure that they meet these requirements.

For further guidance on approved external archiving providers, contact the KIM Team.

4. Reviewing what you keep

Retention is an ongoing process. Organisations should have a policy or procedure to review stored data and/or samples, and decide whether further retention or destruction is appropriate.

Data and samples should only be kept for as long as necessary (including to support further research, defend any scientific challenge, or meet regulatory or governance requirements). When no longer required, destruction or disposal should be considered. The decision-making flowchart in Annex 1 provides further guidance.

The process of review could be initiated by the impending departure of a member of staff from the organisation (for example a PhD student moving on, or a programme leader retiring). PIs have a key role to play in any decisions about the retention or destruction of research data and/or samples collected, generated or used within their programme of research. It is therefore particularly important that PIs are involved in any retention or destruction decisions before they (or one of their students) moves on to another research organisation, or the PI retires.

Scenario: samples and associated data

Where a PI has a large collection of biological samples and data, then it's important for them to help determine:

Whether anyone wants these samples and/or data.

Perhaps the PI accessed the samples and/or data from a biobank or a data provider, and a legal agreement is in place stating the fate of the samples and/or data; or

The PI or student is keen to take the samples and/or data with them to their new research organisation; or

Someone else within the current research team, another collaborator or a biobank wants to take responsibility (or custodianship) for the data and/or samples.

Whether the PI is confident about the quality of the samples and/or data. Have the samples been stored appropriately; is there any evidence to prove sample quality? Similarly, is the data ready for sharing? Does it have appropriate high-quality metadata? (Any recipients of samples or data are likely to want to be assured of their quality).

Whether any restrictions limit the use of the samples or data. Are there any legal agreements in place which state what can and cannot be done with the samples or data? Do the samples have specific storage requirements (for example, do they need to be kept in an ultra-low temperature freezer?) For both samples and data, does the planned secondary or future research align with donors' expectations, do you have evidence of consent? (Again, any potential recipients of the samples and/or data are likely to want assurance in terms of consent. Work with local contacts to determine what assurance is most appropriate in your circumstances).

Making transfers

Where data and/or samples will be transferred to another organisation, then the transfer arrangements should detail any conditions of access and eventual destruction (these arrangements could be outlined in an agreement). Where relevant the Data Protection Officer and/or the Designated Individual (or equivalent) in both your organisation and the recipient organisation should be notified.

All decisions (either further retention, transfer, or the destruction of samples and/or data) should be documented, signed off, and kept for future reference so that there is an enduring record to track what has happened to all research data and/or samples within your organisation. Sign off should be provided by the Institute Director, Head of Department, Designated Individual (or equivalent), Data Protection Officer or sponsor representative.

Permanent preservation

MRC recommends that research data from studies that directly inform national policymaking, or are of other historical significance or importance, should be considered for permanent preservation. In some cases, policy impact may be an explicit aim of the study; in others, this significance may only become apparent later. It may therefore be necessary to consider the impact of the study at several points in its lifecycle, particularly for studies which have long-term goals and run for many years.

Review points

Once decisions have been made about what, how and where to keep data, what is held should be reviewed on an ongoing basis. Below are some suggested timelines for review and potential decisions:

Less than 5 years

- Retain physical data in its original form unless the degradation risk is high. For example, would microscope slides fade and degrade in this time?

Between 5 to 20 years

- Review what will be retained and in what format
- If you have any remaining samples, assess fitness for purpose
- Maintain accessibility of electronic records. For example, systematically check files (including backups) upgrading any software or hardware to ensure continued access
- For basic research it will also be timely to consider destruction (as discussed below)

After 20 years
Consider destruction unless research data are subject to statutory or regulatory requirements (such as the Clinical Trial Regulations) or are of high-risk or high-value.

- In terms of risk consider whether the research was high profile, high impact, produced contentious findings; or involved an invasive medical intervention and/or vulnerable groups
- In terms of value consider whether and/or how often data have been shared; whether it's been used in any preprints or publications; generated protected IP; if there's been any policy impact or the research was historically significant or important

Annex 1: Retention decision making flowchart for research data and/or samples