



Knight, Gareth (2015) LSHTM - EPSRC Research Data Management Compliance Report. London School of Hygiene & Tropical Medicine. (Unpublished)

Downloaded from: <http://researchonline.lshtm.ac.uk/3716463/>

DOI: [10.17037/PUBS.03716463](https://doi.org/10.17037/PUBS.03716463)

Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: <http://creativecommons.org/licenses/by/2.5/>

EPSRC Research Data Management Compliance Report

Contents

Introduction	2
Approval Process.....	2
Review Schedule	2
Acknowledgement	2
EPSRC Expectations.....	3
1. Awareness of EPSRC principles and expectations	3
2. Data Access Statement in published papers.....	4
3. Institutional policies and processes for maintaining awareness of research data	5
4. Management and sharing of non-digital data	6
5. Catalogue of public-funded research data holdings.....	7
6. Barriers to data sharing.....	8
7. Data retention.....	10
8. Curation throughout the data lifecycle.....	11
9. Funding for curation activities	12

Report Version Control

Version	Date	Author	Change Description
1.0	18 January 2013	Gareth Knight	First version of RDM Roadmap produced, outlining how LSHTM intend to comply with each of the EPSRC Expectations
1.1	30 March 2015	Gareth Knight	Majority of text re-written to reflect current status of RDM support in preparation for EPSRC May 2015 deadline
1.2	30 April 2015	Gareth Knight	Minor revision based on feedback from RDM Steering Group and other staff consulted.

Introduction

This report describes the approach taken by the London School of Hygiene & Tropical Medicine to enhance data management practice within the institution and fulfil expectations published by the Engineering and Physical Sciences Research Council (EPSRC). Specifically, it addresses requirements outlined in two documents

1. EPSRC Policy Framework on Research Data: Expectations
< <http://www.epsrc.ac.uk/about/standards/researchdata/expectations/> >
2. Clarification of EPSRC Expectations on Research Data Management
< <http://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement/> >

The Engineering and Physical Sciences Research Council (EPSRC) 'Policy Framework on Research Datasets'¹ identifies research data as a 'public good produced in the public interest' that should be managed and shared in order to protect and maximise the investment made in its creation. To ensure that adequate resources are allocated to this task, the EPSRC sets out nine 'expectations'² that the host institution must provide to researchers. These requirements are accompanied by a penalty – if an institution is found to be seriously failing to comply with the expectations, or deliberately obstructing their application by the 1st May 2015 deadline, a decision may be taken to withdraw existing funding, and declare the institution ineligible for future funding until the issues are addressed.

Although the London School of Hygiene & Tropical Medicine currently receives little funding from the EPSRC, it is considered important that it fulfils their stated requirements for two reasons:

1. To ensure that LSHTM researchers are able to seek EPSRC funded research grants in the future
2. To ensure that research data management services operated by the School comply with good practice and remain competitive with similar work performed in the broader academic sector.

Approval Process

This document was reviewed by the LSHTM RDM Steering Group in April 2015.

Review Schedule

This document will be reviewed annually by the RDM Support Service and further details added on new initiatives that address obligations and respond to emergent expectations as required.

Acknowledgement

We would like to acknowledge the influence of several other EPSRC Roadmaps³ upon the development of LSHTM's compliance report, most notably the University of Bath's EPSRC Roadmap⁴.

¹ <http://www.epsrc.ac.uk/about/standards/researchdata/>

² <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/expectations.aspx>

³ <http://www.dcc.ac.uk/resources/policy-and-legal/epsrc-institutional-roadmaps>

⁴ <http://www.bath.ac.uk/rdso/University-of-Bath-Roadmap-for-EPSRC.pdf>

EPSRC Expectations

1. Awareness of EPSRC principles and expectations

“Research organisations will promote internal awareness of these principles and expectations and ensure that their researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data” (EPSRC Expectation 1)

No.	Description	Action Taken	Responsibility	Status
1.1	Promote awareness of EPSRC principles & expectations			
	<i>The EPSRC expect researchers and students to be aware their requirements</i>	The importance data management is promoted in: <ul style="list-style-type: none"> Guidelines on Good Research Practice Research Degrees Handbook RDM Policy and RDM introduction information 	RDM Service	Green
		LSHTM policy suite covers topics such as Information Management and Security, Data Protection, and Freedom of Information.	Information Services Directorate and other units	Green
		Individual units, such as Clinical Trials, maintain Standard Operating Procedures (SOPs) that cover the design, management, conduct, oversight, regulatory requirements and completion of research projects within specific environmental contexts.		
		LSHTM RDM Policy incorporates relevant EPSRC principles	RDM Service	Green
		EPSRC summary guide published on RDM website	RDM Service	Green
		Tailored guidance on EPSRC requirements provided through RDM modules for staff and students	RDM Service	Green
1.2.	Ensure researchers & students are aware of regulatory environment			
	<i>The EPSRC expect researchers and students to be aware of regulatory and funding environment in which they work</i>	Relevant LSHTM Policies listed in (among other locations): <ul style="list-style-type: none"> Funder Data Management Plan (DMP) examples Clinical Trial Data Management SOP LSHTM RDM webpages 	Information Services Directorate and other units	Green
		Domain and country specific regulations outlined in the academic teaching programme, seminar series & staff/student training modules, Many guides are tailored to specific research areas, e.g. clinical trials.	Information Services Directorate and other units	Green

No.	Description	Action Taken	Responsibility	Status
1.3.	Awareness of reasons to withhold research data			
	<i>The EPSRC do not expect researchers to breach regulatory, contractual, or other obligations and expect institutions to guide researchers on the topic.</i>	Guidance published on LSHTM RDM website recommends a measured approach to data sharing that addresses all relevant, but potentially overlapping obligations	RDM Service	Green
		RDM training modules for staff and PhD students offer decision tree for determining ability to share data	RDM Service	Green
		'RDM helpdesk' offers one-to-one support for staff and students seeking to meet journal and funder sharing obligations while addressing other conflicting obligations	RDM Service	Green

2. Data Access Statement in published papers

"Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed." (EPSRC Principle 2)

No.	Objective	Action Taken	Responsibility	Status
2.1.	Improve awareness of data as a citable object			
	<i>Researchers should be made aware of the function of data as a distinct data output that may be used for further research</i>	Guidance on reasons to cite data and practice to be applied offered via various resources: <ul style="list-style-type: none"> • RDM training modules for staff & students • LSHTM RDM website 	RDM Service	Green
		Case studies of new models of publication provided through: <ul style="list-style-type: none"> • Lunchtime RDM seminars • Regular RDM workshops 	RDM Service	Green
2.2.	Embed understanding of data citation practice			
	<i>Researchers should be aware of good practice for data citation</i>	The LSHTM RDM website provides: <ul style="list-style-type: none"> • Reasons to cite data • Recommendations on best practice (e.g. as recommended by ESRC and others) • Practical examples of citing public, restricted access and unpublished data 	RDM Service	Green
		LSHTM training modules for staff and students outline <ul style="list-style-type: none"> • Reasons to cite data 	Information Services Directorate	Green

No.	Objective	Action Taken	Responsibility	Status
		<ul style="list-style-type: none"> Recommendations on best practice (as advocated by ESRC and others) Practical examples of citing public, restricted access and unpublished data 		
		'RDM Helpdesk' offers one-to-one support for staff/students wishing to address publisher/funder obligations for data sharing		Green
2.3.	Provide tools for citing research data			
	<i>Researchers should be aware of tools and services that will help them to cite data</i>	<p>The LSHTM Research Data repository provides:</p> <ul style="list-style-type: none"> A pre-generated Harvard/Vancouver citation for each metadata record A plug-in which allows citations to be exported to Endnote and other tools Data collections that meet eligibility criteria are assigned a Digital Object Identifier (DOI) 	RDM Service	Green

3. Institutional policies and processes for maintaining awareness of research data

"Each research organisation will have specific policies and associated processes to maintain effective internal awareness of their publicly-funded research data holdings and of requests by third parties to access such data; all of their researchers or research students funded by EPSRC will be required to comply with research organisation policies in this area or, in exceptional circumstances, to provide justification of why this is not possible." (EPSRC Expectation 3)

No.	Objective	Action Taken	Responsibility	Status
3.1	Policy framework establishes institutions commitment to maintaining research data			
	<i>Document should exist that establish institution's commitment to maintaining research data</i>	LSHTM Research Data Management Policy approved by Senior Leadership Team in June 2014. Review schedule established	RDM Service;	Green
3.2.	Institutional processes for describing public-funded research data are in place			
	<i>Ensure that data created and/or managed by the School is identified and catalogued</i>	Importance of recording research data established in RDM Policy, principle 4	RDM Service; Academic Affairs	Green
		Processes for liaising with funded projects towards the end of the grant to support data cataloguing and sharing activities are being trialled and will be developed further following data repository's launch.	RDM Service; Research Operations	Green
		Self-deposit procedures for creating description records of key outputs available through LSHTM Research Data Repository and promoted in School	RDM Support Service;	Green

No.	Objective	Action Taken	Responsibility	Status
3.3.	Record details of 3rd party requests to access data			
	<i>Ensure consistent processes are in place to record access requests to research data</i>	EPSRC-funded projects made aware of need to notify RDM Service of access requests sent directly to project team via: <ul style="list-style-type: none"> • EPSRC summary guide on LSHTM RDM website • Contact email sent to projects in first months of project 	RDM Service	Green
		RD students funded by EPSRC grant made aware of requirements through: <ul style="list-style-type: none"> • EPSRC summary guide on LSHTM RDM website • RDM training modules in TSP programme 	RDM Service	Green
		LSHTM Research Data repository configured to notify Project's 'Data Contact' and RDM Service of access requests	RDM Service	Green
3.4.	Conditions for restricting access to content			
	<i>Ensure conditions for restricting access to content are understood and applied appropriately</i>	The ability to withhold information on data and review access requests is established in the RDM Policy's 4 th principle, which allows projects to request a waiver.	RDM Service	Green
		Deposit guidelines for the LSHTM Research Data repository provide examples of data that should be recorded	RDM Service	Green
		One-to-one advice on research data that should be shared is provided through the RDM Service 'helpdesk'	RDM Service	Green

4. Management and sharing of non-digital data

Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner) (EPSRC Expectation 4)

No.	Objective	Action Taken	Responsibility	Status
4.1	Ensure paper-based research are managed and shared appropriately			
	<i>Non-digital research data should be managed in compliance with relevant legal, regulatory and contractual obligations</i>	Processes for deposit and storage of physical records to LSHTM Archives are published on the LSHTM intranet. A staff role is identified.	Archives & Records Management Service	Green
		Well documented procedures exist for cataloguing physical records using CALM. Catalogue records are incorporated into aggregation services, such as AIM25.	ARM Service	Green

No.	Objective	Action Taken	Responsibility	Status
		Facility for on-site access available by booking an appointment to visit the LSHTM Archives	ARM Service	Green
		Facility for scanning a small amount of paper material available through LSHTM Archives	ARM Service	Green
4.2.	Ensure physical samples are managed and shared appropriately			
	<i>Physical samples produced by LSHTM labs should be managed in compliance with relevant legal, regulatory and contractual obligations</i>	Processes for storing physical samples on-site and off-site (e.g. UK BioBank) are well documented.	Quality & Governance Manager	Green
		Register of human tissue samples maintained in compliance with Human Tissue Act.	Quality & Governance Manager	Green
4.3.	Maintain linkages between research outputs held in different forms			
	<i>Ensure link-up between metadata records for digital and non-digital outputs</i>	The LSHTM research data repository allows depositors to: <ul style="list-style-type: none"> • Perform a title lookup of papers held in LSHTM ResearchOnline • Reference other resources by providing a title and URL 	RDM Service;	Green
		The CALM system support the ability to link to resources held elsewhere	ARM Service	Green

5. Catalogue of public-funded research data holdings

“Research organisations will ensure that appropriately structured metadata describing the research data they hold is published (normally within 12 months of the data being generated) and made freely accessible on the internet; in each case the metadata must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it. Where the research data referred to in the metadata is a digital object it is expected that the metadata will include use of a robust digital object identifier (For example as available through the DataCite organization)” (EPSRC Expectation 5)

No.	Objective	Action Taken	Responsibility	Status
5.1.	Ensure key data outputs from public funded research are identified and catalogued			
	<i>Key research outputs should be documents as evidence of research and institutional activity</i>	Processes exist for acquiring and cataloguing paper records using the archives’ CALM system.	ARM Service	Green
		Processes for liaising with funded projects towards the end of the grant to support data cataloguing and sharing activities are being trialled and will be developed further following the repository’s launch.	RDM Service; Research Operations	Green
		Self-deposit procedures for creating description records of key outputs available through LSHTM Research Data Repository website following launch.	RDM Service	Green
		Research projects that have a funder or journal requirement to catalogue and/or publish data within 1 year of generation are provided with:	LSHTM RDM Service	Green

No.	Objective	Action Taken	Responsibility	Status
		<ul style="list-style-type: none"> Website guidance on choosing an appropriate repository Tutorials on use of the LSHTM research data repository Tailored advice from the RDM Service on request 		
5.2.	Provide a robust digital object identifier for digital objects			
	<i>A persistent identifier should be assigned to key data collections</i>	LSHTM maintain a contract with the British Library, the UK representative of DataCite, that allows it to assign Digital Object Identifiers (DOIs) to Data Collections submitted to the research data repository that meet eligibility criteria.	Library and Archives Service; RDM Service	Green
		Other identifier schemes, such as grant code identifiers and ORCIDS, are captured in the research data repository as optional fields	RDM Service	Green

6. Barriers to data sharing

“Where access to the data is restricted the published metadata should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example ‘commercially confidential’ data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.” (EPSRC Expectation 6)

No.	Description	Action Taken	Responsibility	Status
6.1.	Embed understanding of good practice for responsible data sharing			
	<i>Researchers should be made aware of the benefits of data sharing, obligations to be met, and reasons that some data can be shared</i>	Guidance on data sharing is provided through: <ul style="list-style-type: none"> The LSHTM RDM website⁵ RDM training modules for staff and students The RDM ‘helpdesk’ maintained by the RDM Service 	RDM Service	Green
6.2.	Personal & sensitive information should be handled appropriately			
	<i>The EPSRC encourage data sharing for purpose of validation and reuse, but do not expect personal & sensitive data to be put into the public domain.</i>	Guidance on data sharing will provide: <ul style="list-style-type: none"> A decision tree outlining steps to be taken to archive data Advice on ‘de-identifying’ data and reference to relevant resources Pointers to country, domain and resource specific guidance 	Research ethics RDM Service;	Green
		The LSHTM Data Repository Collection Policy and deposit guidelines		Green

⁵ <http://www.lshtm.ac.uk/research/researchdataman/share/index.html>

		explicitly state that it is intended for non-personal/non-sensitive data only.		
6.3.	Enhance awareness of embargo periods			
	<i>The EPSRC recognise that delays in publication of research findings and data may be appropriate, e.g. in order to protect Intellectual Property.</i>	Guidance on embargo periods are available on the RDM website through: <ul style="list-style-type: none"> Funder RDM profiles Journal specific guidance, such as that published for PLOS 	RDM Service; Open Access team	Green
		Tailored advice on legitimate and illegitimate reasons for withholding data are available through <ul style="list-style-type: none"> Data sharing training for staff and students One-to-one support available through the RDM Service 		Green
6.4.	Published metadata should state the reason for data restrictions & how access may be granted			
	The EPSRC state that published metadata that describes restricted data should outline the restriction reason and conditions under which access may be granted.	Each file published through the data repository is accompanied by: <ul style="list-style-type: none"> A statement on who can access the data A request form for use by non-authorised users 	RDM Service	Green
		LSHTM data repository deposit guidance states that Data Transfer Agreements (or other legal documents that an end user must sign before gaining data access) must be made public to provide an indication of access conditions to be met.	RDM Service	Green
6.5.	3rd party collaborators should be made aware of the need to validate published findings			
	Public funded research involving 3 rd parties must be planned and executed in a way that enables published findings to be scrutinised and validated by others	The PI is responsible for ensuring team members and collaborators are aware of obligations to be met	Project PI	Green
		LSHTM's default position is to publish RDM guidance and data repository policies on the public website, unless there is a reason that it cannot be made more widely available	RDM Service	Green

7. Data retention

“Research organisations will ensure that EPSRC-funded research data is securely preserved for a minimum of 10-years from the date that any researcher ‘privileged access’ period expires or, if others have accessed the data, from last date on which access to the data was requested by a third party; all reasonable steps will be to ensure that publicly-funded data is not held in any jurisdiction where the available legal safeguards provide lower levels of protection than are available in the UK”

No.	Description	Action Taken	Responsibility	Status
7.1.	Ensure awareness of retention obligations			
	<i>Data must be kept for a minimum of 10 years after a ‘privileged access period’ or last access, whichever is longest</i>	<p>Generic guidance is provided on the following topics:</p> <ul style="list-style-type: none"> <i>Selecting data to be preserved:</i> Described in retention guidance on the RDM website, as well as in the LSHTM Retention and Disposal Schedule. <i>Selecting data to be shared:</i> Described in Data Sharing guidelines on the RDM website. <p>Retention obligations for Clinical Trials are outlined in the LSHTM Data Management SOP (held on the LSHTM intranet)</p>	RDM Service; Archives & Records Management; Quality and Governance Manager	Green
		<p>Guidance on EPSRC retention requirements are provided through:</p> <ul style="list-style-type: none"> The EPSRC funder summary, retention guidelines and data sharing guidelines on the LSHTM RDM website RDM training offered to LSHTM staff and students 	RDM Service; Archives & Records Management	Green
		<p>Specific advice on identifying data of value to the project is available through several sources:</p> <ul style="list-style-type: none"> <i>Non-digital:</i> Archives & Records Management Service <i>Digital data:</i> RDM Service <i>Clinical Trials:</i> LSHTM Quality and Governance Manager 	RDM Service; ARM Service; Quality and Governance Manager	Green
7.2.	Ensure data is held for full retention period			
	<i>To ensure data is kept for the required period, a record must be maintained on notable dates in its lifecycle</i>	For each data collection held on LSHTM’s secure server, metadata is captured on key dates for account creation, file creation (on drive) and appraisal.	Archives & Records Management Service	Green
		Processes for appraising data and deciding next action (e.g. keep until next appraisal date/delete) are documented and applied consistently		
		For each file uploaded to the LSHTM data repository, the depositor must specify the minimum retention period (10 or more years after upload).	RDM Service	Green
		Metrics will be used to monitor demand for EPSRC (and other) data and determine retention requirements. Access and impact metrics are captured	RDM Service	Green

		using the following: <ul style="list-style-type: none"> • <i>Open data</i>: The IRStats plug-in records downloads of publicly available files • <i>Restricted data</i>: Records of controlled access requests submitted to the data repository are captured. In addition, the EPrints alt metrics plugin is used to monitor wider interest in the data.		
7.3.	Ensure an Exit strategy is in place			
	The EPSRC expect institutions to make arrangements for the preservation and continued accessibility of data, in the event of mergers, acquisitions and closures.	The RDM Service maintains an Exit Plan that will be enacted in the event that institutional funding ceases. Arrangements for management of LSHTM Data Collections is one component of this plan.	RDM Service	Green
7.4.	Management & preservation of 3rd party data			
	<i>Institutions are not expected to preserve & manage 3rd party data, but may wish to do so if their research results depend upon the continued availability of the data</i>	Decisions on the need to curate, preserve and share 3 rd party research data will be made on a case-by-case basis, in conjunction with individual researchers or project teams. If the 3 rd party data relates to a specific project, the Principal Investigator may be specified as the designated data contact.	RDM Service	Green

8. Curation throughout the data lifecycle

“Research organisations will ensure that effective data curation is provided throughout the full data lifecycle, with ‘data curation’ and ‘data lifecycle’ being as defined by the Digital Curation Centre. The full range of responsibilities associated with data curation over the data lifecycle will be clearly allocated within the research organisation, and where research data is subject to restricted access the research organisation will implement and manage appropriate security controls; research organisations will particularly ensure that the quality assurance of their data curation processes is a specifically assigned responsibility”;

No.	Description	Action Taken	Responsibility	Status
8.1.	Ensure awareness of relevant research lifecycle models			
	<i>Research often involves performance of several common activities, irrespective of the objectives methodology and outcomes. To effectively plan data management activities, it is essential that they apply a holistic approach.</i>	LSHTM’s RDM website is structured to reflect key stages of the research lifecycle and the activities to be performed	RDM Service	Green
		A Data Management Plan is mandated for all LSHTM research projects and recommended for student projects from January 2015, as part of the RDM Policy	RDM Service	Green
		The research and data lifecycle are used as a starting point to frame data management discussion in Data Management Planning modules offered to staff and students	RDM Service	Green

8.2.	Ensure research data is curated and preserved following project completion			
	<i>Research data serves as evidence of research activity and institutional investment. There is a need to ensure data management processes address all stages of its lifecycle, from its creation to (potential) destruction</i>	A research data repository will be available to LSHTM staff from May 2015 onwards.	RDM Service	Green
		Policies and processes for ingesting, curating, preserving and removing data submitted to the data repository are in place	RDM Service	Green
		Secure storage suitable for holding personal and sensitive data is available in-house and procedures for obtaining an account are available on the LSHTM intranet.	IT Services/ Archives	Green

9. Funding for curation activities

“Research organisations will ensure adequate resources are provided to support the curation of publicly-funded research data; these resources will be allocated from within their existing public funding streams, whether received from Research Councils as direct or indirect support for specific projects or from higher education Funding Councils as block grants.”

No.	Description	Action Taken	Responsibility	Status
9.1.	Build awareness of data management costing			
	<i>There is a need to build awareness of the EPSRC's (and other funders) position on data management costs and rules that should be followed.</i>	The LSHTM RDM website offers guidance on: <ul style="list-style-type: none"> Funder positions on data management costs Rules and models to be followed 	RDM Service	Green
		Training on Data Management Plans cover funder positions and RDM costing rules. Uptake of training is monitored on an ongoing basis	RDM Service	Green
		Research projects are provided with support at the bid stage as part of the 'Helpdesk' function to ensure they understand costs of short and long-term data management.	RDM Service	Green
		All LSHTM-led funded projects are required to write a Data Management Plan on confirmation of funding from January 2015, even if one is not required by their funder. Oversight of this activity is provided by the RDM Service.	RDM Service	Green
		MPhil/PhD/DrPH degree students encouraged to write a Data Management Plan and include it in a Upgrade/review report for review by their supervisor.	Student supervisor	Green
9.2.	Ensure institutional Data Management services are sustainable			
	<i>The curation and preservation of data collections is dependent upon the allocation of</i>	An itemised list of operation costs for the RDM Service have been identified and submitted to the relevant committee for approval each year.	RDM Service	Green

	<i>appropriate funds</i>	Cost recovery mechanisms for use of IT Services are being explored within the School.	IT Services	Green
9.3.	Embed data management infrastructure into institutional processes			
	<i>There is a need to encourage uptake of RDM Services across the institution, in order to ensure the longevity of the service</i>	The functions provided by the RDM Service are promoted at department meetings, marketplace events and on the RDM website, among other places.	RDM Service	Green
		Tailored advice on use of LSHTM's data management facilities is available on request. Further resources will be developed on an ongoing basis.	IT Services RDM Service	Green