LONDON
SCHOOL *of*
HYGIENE
&TROPICAL
MEDICINE

Program: electronic library and information systems

Building a research data management service for the London school of hygiene & tropical medicine
Gareth Knight,

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

Building a research data management service for the London school of hygiene & tropical medicine

Gareth Knight

London School of Hygiene and Tropical Medicine, London, UK

Abstract

Purpose – The purpose of this paper is to present a case study of work performed at the London School of Hygiene and Tropical Medicine to set-up a Research Data Management Service and tailor it to the needs of health researchers.

Design/methodology/approach – The paper describes the motivations for establishing the RDM Service and outlines the three objectives that were set to improve data management practice within the institution. Each of the objectives are explored in turn, stating how they were addressed.

Findings – A university with limited resources can operate a RDM Service that pro-actively supports researchers wishing to manage research data by monitoring evolving support needs, identifying common trends and developing resources that will reduce the time investment needed. The institution-wide survey identified a need for guidance on developing data documentation and archiving research data following project completion. Analysis of ongoing support requests identifies a need for guidance on data management plans and complying with journal sharing requirements.

Research limitations/implications – The paper provides a case study of a single institution. The results may not be generally applicable to universities that support other disciplines.

Practical implications – The case study may be helpful in helping other universities to establish an RDM Service using limited resources.

Originality/value – The paper outlines how the evolving data management needs of public health researchers can be identified and a strategy that can be adopted by an RDM Service to efficiently address these requirements.

Keywords Digital curation, Public health, Archives, Research data management, Digital repository, RDM service

Paper type Case study

1. Background
1.1 Introduction
The Research Data Management Service is an increasingly familiar unit within a university, providing a range of services to support researchers who are creating, managing and sharing their research data. As a result, discussions associated with these services have moved from a need to justify their set-up, to consideration of how these services can be

sustained. This is a particular challenge for small universities, which are under constant pressure to demonstrate that infrastructure investment is allocated to the correct services.

This paper presents a case study of work performed at the London School of Hygiene & Tropical Medicine (LSHTM) to establish a Research Data Management Service and tailor it to the needs of health researchers. It describes the objectives that the service sought to address and the challenges that were encountered when obtaining institutional approval for the RDM policy and mandate the creation of data management plans for all research projects. It goes on to describe the requirements gathering activities that were performed in the school through an institution-wide survey and ongoing monitoring of support requests.

1.2 Institutional environment

The LSHTM is a university specialising in research and postgraduate education in public health and tropical medicine. It is a constituent member of the University of London and The Bloomsbury Colleges Group, located across two sites in central London. As an institution, it is relatively small, however, its research has significant influence in its field, having the largest volume of world-leading research in the areas of public health, health services and primary care submitted for REF2014 (Jump, 2014). It employs over 1,000 academic staff, many of whom work overseas.

The institutional environment has a significant influence upon the approach taken to provide Research Data Management Services within the institution. Public health researchers generate a large amount of data which must be stored and managed in a secure manner. However, the regulated domain and diverse countries in which they perform work presents challenges – a factor raised during requirements gathering activities. In many cases, it is not necessary to encourage researchers to apply data management practices, but to provide guidance on addressing conflicting data management obligations.

1.3 RDM at LSHTM

The London School has recognised the role that data plays in its research for many years, but has not actively provided central support until recently. The need for an institutional data management infrastructure was first noted in 2002 as part of a JISC-funded study to develop an institutional retention schedule for paper and digital records. This study recognised that paper records were being deposited with the Archives & Records Management (A&RM) Service, but no equivalent process took place for data:

> Even if a future researcher discovers a dataset through the research grants department or information in the annual report, will the data still exist in the School, will staff know where it is and will it be accessible in a readable format? (Cranna, 2003, p. 7).

Although there was a recognition that a central service should exist to help researchers manage their data and ensure it was curated and preserved over time, the A&RM Service did not possess sufficient resources or expertise to take a proactive role at the time. Instead, computing officers located within research faculties were encouraged to take on a support role.

The need for central support for data management was re-visited in 2009. Driven by demand from senior academics to enhance and standardise data management practices, as well as reduce duplication of effort across the school, LSHTM's senior management team set-up a Research Data Working Group (RDWG), chaired by Professor David Leon, to review researchers' data management practices and national/international initiatives in the area. The report, submitted to the SMT in

July 2011, made a series of recommendations to improve data management practice. This included the creation of relevant institutional policies, an RDM web site, an institutional portal for data discovery and an investigation on "how far Archives Service should be strengthened to provide relevant central support and guidance on these issues, and what the resource implications of this may be". (LSHTM Research Data Working Group, 2011, p. 3). The placement of RDM within the Archives, as opposed to its more common location in the library's remit in other universities, was considered to be essential to ensure that institutional assets which must be held in the long-term could be curated and preserved in a consistent manner, irrespective of whether they are held in digital or physical form.

1.4 An institutional Research Data Management Service

The LSHTM Research Data Management Service was established in 2012 as a unit within the A&RM Service. Using funds provided by the Wellcome Trust Institutional Strategic Support Fund – a funding scheme through which UK universities can bid for funding to build or re-develop strategically important institutional infrastructure[1] – the LSHTM were able to recruit two full-time RDM posts, a Project Manager and Software Developer[2] to take forward the RDM work. Wellcome Trust funding is provided until July 2015, after which the RDM Service will be supported through institution funding. The RDM Service is supported by a Steering Group, chaired by the School's Deputy Director and comprised of selected academic and professional support staff from each faculty, which champion the service within upper management and provides direction on a management by exception basis.

From the outset, it was recognised that the Project Manager could not support the large number of LSHTM staff and students working in London and overseas, and that a three-stage triage approach was needed to prioritise support. First, a decision was taken to prioritise support staff working on funded research. Unfunded researchers and students wishing to manage their data would be provided support and advice, but only on a reactive basis through training modules and support requests. A second step utilise existing expertise within the school where possible. Questions related to clinical data management, for example, would be directed to LSHTM's Quality & Governance Manager. Third, a decision was taken to record details on each RDM support query submitted to the RDM Service and use it as a basis to plan development work. For example, through provision of written guidance and organised workshops. This would reduce the likelihood that the same question was submitted on multiple occasions, or as a minimum, reduce the time required to process these request.

2. RDM service objectives

For its initial three-year funding period, the RDM Service was given a broad remit to enhance data management practice within the institution. This was broken down into three objectives:

(1) strengthen the institution's policy framework to ensure research data management needs are addressed;

(2) enhance infrastructure to support data management activities; and

(3) improve data management practice among researchers within the institution.

2.1 Strengthen the institution's policy framework
The first objective was to ensure research data management was addressed in the institution's policy framework. Primarily, this work focused upon the creation of a Research Data Management Policy (Knight, 2014a) that addressed the RCUK's Common Principles on Data[3] and matched similar documents being developed at other institutions[4]. Work was also performed to update other documents, such as the school's research handbook.

It was recognised from the outset that the policy could serve as a method for improving researchers' data management practice and embedding new services. Therefore a decision was made to focus upon the principles that researchers should follow to ensure their data were managed and shared in accordance with current best practice (Knight, 2014a). Consideration of the institutional services needed to support these activities would be addressed elsewhere, such as through the RDM web site and service level agreements.

A first version of the policy was distributed in October 2012 and a 12 month consultation process undertaken, during which time its implications were considered by various committees, departments and research groups within the school. The feasibility of complying with the nine principles was a key concern during the consultation period – would public health researchers be able to comply with a mandatory requirement? If not, should the activity be made conditional or only recommended? As a result, a small number of principles were changed to allow certain groups to be exempt. Projects with industry collaborators or gathering data in certain countries, for example, are often subject to collaborative licensing agreements that prevent them complying with principle 2. Similarly, it was difficult for many projects to comply with principle 6's requirement that they offer data to a repository or enclave, due to the large number of external obligations with which they need to comply. In practice, the only new activities that could be mandated were those comprised of internal procedures that did not affect the project's external operation (principles 3 and 4) or are increasingly accepted within the research community (principle 9). Activities such as applying a non-exclusive licence, transferring data to an appropriate repository/enclave, making data available for access and claiming data management costs are encouraged, but can only be performed if there are no external or internal barriers.

Following the consultation, the policy was submitted for approval to the relevant committee. There was an initial reluctance to introduce a new policy which may present research barriers. Health research is a highly regulated environment that is already subject to a number of similar requirements. However, the submission of additional evidence on how the policy would help the institution to meet its research, legal and contractual obligations, as well as strong upper management support – a recognised need for gaining policy approval in other institutions as well (Fitt *et al.*, 2015; Jones *et al.*, 2013) – ensured that it was approved.

2.2 Enhance infrastructure to support data management
The second objective was to scope and develop the data management infrastructure within the institution, in conjunction with IT services. Building upon the findings of the requirements gathering exercise (see Figure 1), which identified a need for guidance on processes for "archiving" data following their completion, ongoing work has focused upon two aspects:

(1) RDM support development: training and guidance has been developed for LSHTM staff to help them locate internal or external systems suitable for their data. Guidance focuses upon four criteria at present: subject domain, sensitivity, content type and collection size.
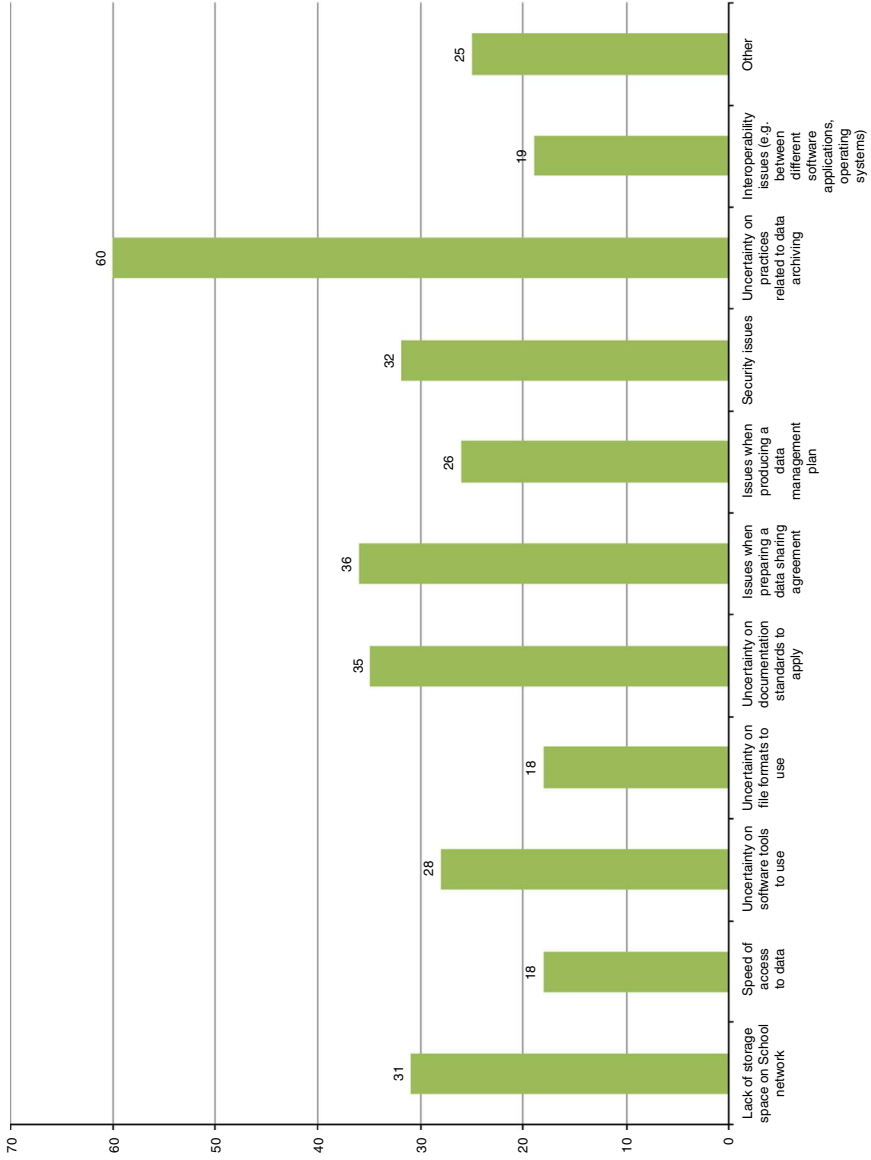
Figure 1.
Data management
challenges and
issues encountered
during the
research process

(2) Repository development: an institutional repository that could be used to catalogue and share research data assets was scoped, in order to address a gap in current support.

To ensure the data management infrastructure was sustainable, a decision was made to only store research data in-house in the LSHTM repository if it cannot be hosted elsewhere (e.g. deposited to the UK data service). This strategy differs significantly from the approach taken by LSHTM's A&RM Service, which accepts responsibility for managing the institution's paper records. However, it was considered essential to enable a limited-resource RDM Service to fulfil its responsibility to curate and preserve LSHTM research data.

The LSHTM Research Data Repository is currently in-development and will be launched in mid-2015. Following the creation of a functional specification and evaluation of several common repository software tools (Alfresco, CKAN, DSpace, EPrints and Fedora), a decision was made to adopt the University of Southampton's EPrints platform [5] and tailor it to our needs through several third party plugins (such as recollect[6] and collections[7]) and further in-house developments. This will be hosted with the University of London Computing Centre, to enable the school to take advantage of their extensive experience of EPrints development.

2.3 Improve researcher data management practice

The third and final objective was a broad recommendation that data management practice should be enhanced within the institution. This objective was divided into three tasks:

(1) determine research data management needs within the institution;

(2) ensure data management is considered by researchers from the project outset; and

(3) support the evolving data management needs of researchers.

To ensure that resources were allocated appropriate, an analysis was performed of stakeholders to be supported and most effective method of communication.

2.3.1 Determine research data management needs within the institution. Much of the research performed at the London School focuses upon public health and tropical medicine. However, this does not simplify the process of determining data management needs within the institution. The requirements of social scientists using mobile devices to perform surveys in the field differ significantly from researchers working in a lab environment. In the early stages of the project, a needs analysis was performed to identify areas where researchers needed help and determine the third party obligations that affected their approach to data management. This investigation was conducted using a combination of methods:

• Academic researchers: a web survey was conducted with academic staff to gain a better understanding of their research, data and perceived needs.

• Upper management: desk research was performed to review policies and procedures related to research operations, IT services and academic departments, as well as requirements of the RDWG.

• Research funders: a list of research funders that frequently support LSHTM projects was produced and desk research performed to determine institution and project-specific RDM requirements (Knight, 2012a).

- Regulatory obligations: desk research was performed of national/international obligations that affect research data, particularly that which involves human participants, such as the Data Protection Act.

Initially, it was planned that the Data Asset Framework (Jones *et al.*, 2009) would be applied to produce three-four case studies on researchers' projects, similar to those produced by the University of Edinburgh (Ekmekcioglu and Rice, 2009), Oxford (Martinez-Uribe, 2008) and Southampton (Gibbs, 2009). However, the Project Manager's unfamiliarity with the institution at the time meant it was difficult to locate researchers willing to participate, while a central services re-organisation taking place meant that many professional support staff were unavailable.

The RDM Service had greater success with the web survey, conducted using the Bristol Online Survey tool[8] over a five-week-period (4 October-18 November 2012). This produced 117 responses, representing 16.25 per cent of academic staff. The RDM survey consisted of 15 questions (Knight, 2012b) drawn from the data asset framework and DRAMBORA (McHugh, 2007) toolkits. Rather than apply the project-by-project approach encouraged by these methodologies, questions were re-worded to examine common data management practices applied by the researcher across all of their current projects. This approach provided a high-level overview of a large number of research projects taking place at the time, but did occasionally produce anomalous responses. For example, a small number of respondents indicated they held data in eight different locations, without indicating if it was the same data being replicated across these locations, or data from different projects. The survey also provided insight into the environment in which health research is performed and the challenges encountered. It was recognised that data management practices are influenced by many stakeholders, including standards bodies (e.g. ICH Good Clinical Practice), research funders, journal publishers, federal agencies (e.g. US Food and Drug Administration) and governments (depending upon the location of data collection and partner institutions). Meeting these requirements is resource-intensive, particularly when conflicting obligations require the researcher to re-negotiate an agreement.

Figure 1 outlines key challenges that researchers encounter when creating, managing and sharing data within the school (Knight, 2013).

The research issues encountered by LSHTM researchers are comparable to those identified in RDM surveys performed at other institutions, such as the University of Nottingham (Parsons *et al.*, 2013, p. 22), Royal Veterinary College (Harrison, 2013, p. 12) and the University of Northampton (Alexogiannopoulos *et al.*, 2010, p. 28). Although differences in survey methodology make it difficult to make an exact comparison, topics such as writing data management plans, use of institutional storage systems, security processes and documentation standards are common across institutions. Similarly, the recognition that many LSHTM staff were unaware of institutional services available or could not access them when working in the field matched findings expressed by Alexogiannopoulos *et al.* (2010) that Northampton's institutional storage was underutilized due to capacity concerns and difficulty in obtaining external access. It is only when researchers' needs are examined in detail, as explored in Section 2.3.3, that differences in working practice begin to emerge.

The survey results identified several areas where data management practice should be improved, most notably a lack of procedures for managing data following project completion. In many cases, researchers would store the data in the personal or

department area of the school's network and move onto the next project, except in rare cases where a funder required them to deposit the data with a third party data service. To address this gap, staff and students training sessions were organised to address ethical and technical issues associated with post-project management. Meetings were also held with IT services and research ethics to discuss issues raised and consider how they could be addressed within the institution.

2.3.2 Ensure data management is considered from project outset. Data management is increasingly perceived as a key component of good research practice that should be considered early to ensure opportunities are identified, risks are mitigated and appropriate resource are allocated before work commences. In many cases, researchers already describe their data management processes, either within a funder-mandated data management plan or a domain-specific research protocol document[9]. However, there remains a small number of projects that are not covered by either requirement and, potentially do not consider data management until they undertake the research process itself. To address this gap, the RDM service sought to introduce a requirement in its RDM policy (principle 3 in Table I) that all research projects must create a DMP and submit it to the RDM Service for review. This approach is increasingly becoming standard practice across the academic sector, implemented by a growing number of research-intensive universities (Horton and DCC, 2014).

It was recognised from the outset that the introduction of an institution-wide DMP would be resource-intensive, both for research projects producing the plan and the RDM staff that would review it. Therefore, three requirements were set to minimise the work involved:

(1) Prioritise key projects: the institution is responsible for managing data produced by projects where it is the named grant holder, data creator or data

No.	Principle	Status
1	Research data must be created, maintained and shared in accordance with contractual, legislative, regulatory, ethical and other relevant requirements	Mandate for all
2	Rights assigned to research data should not unnecessarily restrict its management, sharing or use	Recommend
3	A data management plan should be produced for all research projects that are creating or capturing data	Mandate for some
4	Research data created or captured by researchers must be registered with the school, irrespective of whether it is hosted at the school or elsewhere	Mandate for all
5	Research data must be held in a managed storage environment throughout the period of retention	Mandate for all
6	Research data created or captured by LSHTM researchers should be offered to an appropriate data repository or enclave designated by the school or funder, except in circumstances that would breach IPR, ethical, confidentiality or other obligations	Recommend
7	Research data that substantiate research findings should be made available for access and use in a timely manner, within the boundaries of conditions established by contractual, legislative, ethical or other requirements	Recommend
8	Research data produced and/or used during research must be cited in research outputs	Mandate for all
9	Management and sharing of research data should be supported through the allocation of research funding, where permitted	Recommend

Table I.
LSHTM RDM
policy principles

manager. These projects should be identified and steps taken to ensure their data is managed appropriately.

(2) Avoid unnecessary duplication: projects that have produced a funder DMP or research protocol document should not be required to complete an institutional DMP.

432

(3) Offer guidance, not judgment: the DMP should be easy to complete and offer suggestions on approaches that may be taken in the local context. Supporting information, example responses and multiple choice options should be provided wherever possible.

To comply with the three rules, a principal investigator would need only create a LSHTM data management plan if it was LSHTM-led, working with new data for which it was responsible and funded by an external body that does not have existing DMP requirements. Projects that had written a funder DMP or research protocol could fulfil requirements by forwarding an existing document. Projects that are analysing secondary data only, not working with data (e.g. those funded for the purpose of organising meetings and workshops), unfunded or performed as part of a consultancy are exempt.

A 15-question institutional DMP template was created[10], building upon the Digital Curation Centre's DMPOnline generic template[11] and similar work performed by the University of Bath (Cope, 2013), which underwent testing during 2014. Initial feedback has been generally positive, indicating that the questions were easy to answer, requiring 15-20 minutes on average. However, several participants commented that the DMP forms are difficult to complete for large-scale, multi-site projects, proposing that a separate DMP form be tailored for these projects. There was also a general concern that the DMP would be critically evaluated and the Principal Investigator labelled "a bad data manager" if they were not able to respond.

To determine if the approach could be sustained with only one full-time RDM staff member, we sought to establish the number of LSHTM projects that might be eligible for the new DMP requirement in a given year. We began by analysing a list of proposals handled by the research operations team in 2013 – during which 627 applications were submitted (the largest number of bids submitted in recent years) – and classified them according to eligibility (LSHTM-led, submitted to funders with no DMP requirement, not a consultancy). It was not possible to determine the project objective, which may have indicated if it was undertaken to perform new research or co-ordinate workshop/meetings. However, the information was sufficient to establish that a maximum of 247 of the 627 applications (39 per cent) would be expected to complete a LSHTM DMP form at the submission stage if the requirement had been in place in 2013. Although it was considered feasible to review and comment upon this number of applications, it would require a significant amount of time, which would prevent other activities being performed. In addition, the dependency on a single person to review DMPs would create a bottleneck, preventing research bids being submitted if that person was on leave or unwell.

The recognition that a pre-award DMP would be unsustainable prompted investigation of a second approach – request eligible projects complete a DMP following funding confirmation. It was recognised that this may be too late to provide input into projects that had not allocated sufficient resources to data management. However, it would enable the

RDM Service to focus only on the smaller number of proposals that are guaranteed funding – an approach also applied by the University of Hertfordshire[12]. By re-analysing the research operations data set, 197 projects were identified with a start date in 2013, of which only 97 met the eligibility criteria. This figure was considered to be easier to manage with the limited resources available to the RDM Service.

The proposal to mandate the creation of a data management plan as a post-award requirement proved to be essential for gaining senior management approval. Although the value of data in health research was well recognised, there was a reluctance to increase the Principal Investigator's workload when they are already subject to many external requirements. By introducing the requirement following confirmation of funding, they would have more time to spend on discussing their data management approach with team members.

2.3.3 Supporting evolving data management needs. Finally, there was a need to provide tailored advice in response to researcher requests. This is a core activity for many RDM Services, but represents an unknown activity for those wishing to plan resource allocation. A fledgling RDM Service is unlikely have accurate figures on which to estimate demand and there is currently little information available on the number and type of support requests handled by comparable institutions.

In the early stages of service planning, we attempted to estimate potential demand in different scenarios, following the recommendations of the Keeping Research Data Safe project (Beagrie *et al.*, 2010) and LIFE project (Ayris *et al.*, 2008). For example, if an average of 300 research bids submitted each year have a DMP component, of which 20 per cent require three hours of support, it may be estimated that 180 hours will need to be spent reviewing research bids. Following service launch, we began to record details on each RDM support request submitted. Initially, the information was captured to inform the RDM Steering Group of progress. However, over time it offered insight into the faculties that were using the RDM Service (and those that were not) and areas where advice was most needed.

The number of support requests submitted to LSHTM's RDM Service has grown steadily over the past two years. Following the RDM Service's launch in November 2012, 13 queries were received and processed in November-December 2012, increasing to 88 queries for 2013, and 120 during 2014. This number can be easily processed by one person, in conjunction with other work. The number of support requests broadly match those reported by the University of Southampton, which indicate that 90 requests were submitted to their RDM Service e-mail account in the first year following their "soft launch" (White and Coles, 2014, slide 26). However, they are much smaller than the 158 RDM support requests processed over a three month period by the University of Manchester (Beard, 2014, slide 8).

Following recommendations by Jones *et al.* (2013), we encourage staff and students to submit questions via the RDM Service e-mail account, to provide continuity if and when staff leave. However, in most cases support requests are sent to the RDM Project Manager directly via e-mail. When asked the reason for this, researchers indicated they prefer to contact a named person, on the basis that they are likely to receive a quicker response. A small number of queries of support requests are communicated via telephone, scheduled drop-in sessions or simply visiting the RDM office for advice.

Figure 2 provides a breakdown of LSHTM support requests from each of LSHTM's three research faculties (Epidemiology and Population Health (EPH), Infectious and
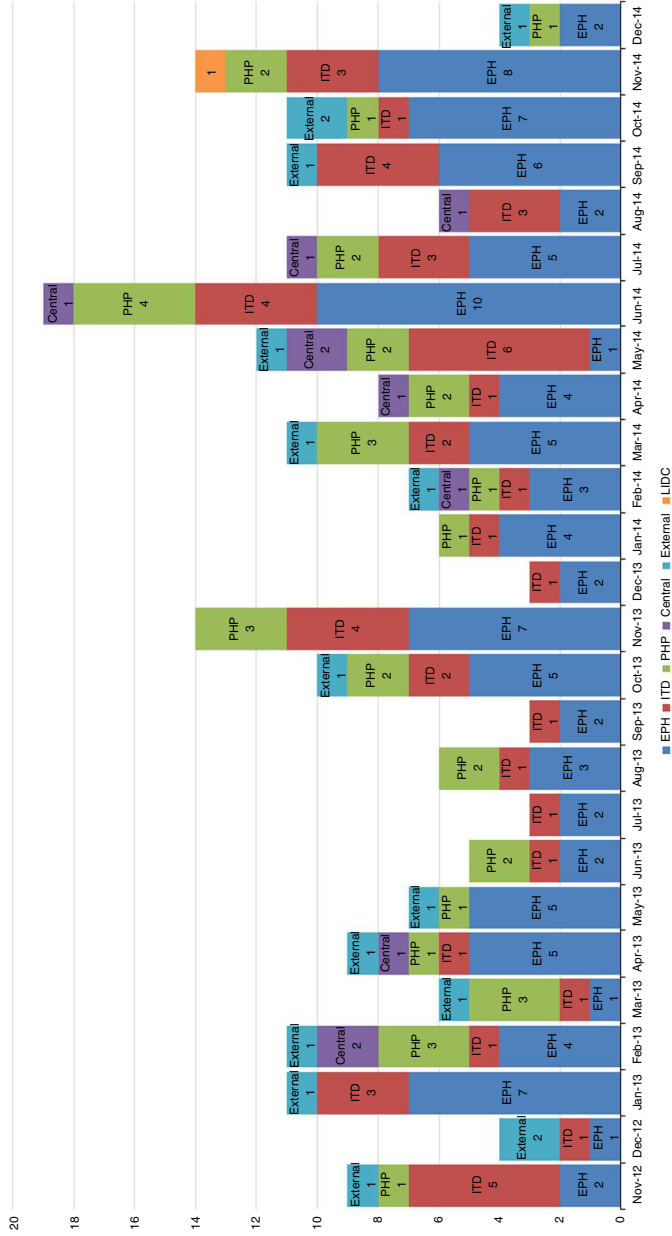
Figure 2.
Number of support
requests per month
and their source

Tropical Diseases and Public Health and Policy), as well central services (student registry, library, etc.), London International Development Centre and external researchers (researchers and students in other institutions).

An analysis of the number and source of support requests submitted each month can be used to provide insight into which departments are using an RDM Service and which are not. Staff and students in the EPH faculty show greater awareness of the RDM Service as a place for advice, having submitted 105 queries over a 26-month-period. However, less than half of contacts were from ITD (52 queries) or PHP (37 queries). On investigation, it was found that many researchers in these faculties consulted domain experts for advice, such as the LSHTM's Quality and Governance Manager, or visited specific web sites, such as the UK Data Service. Although it is reassuring that researchers utilise existing resources where possible, it has led to a recognition that much of the guidance available through the RDM Service focuses upon broad needs and that further work could be done to produce resources tailored to specific subject domains.

There does not appear to be a correlation in the number of support requests submitted during particular months, beyond a recognition that fewer people get in touch at the end of term (July and December) when many staff are on leave. However, there is a noticeable increase in support requests during months where dissemination activities are organised (lunchtime seminars in January, February, April and October 2013, a drop-in session during June 2014, and half-workshop in November 2013-2014). External factors also influence the number of requests received during specific months – the majority of queries during March 2014 were prompted by growing awareness of the PLOS Data Policy[13], whereas many of the May 2014 requests focused upon a MRC funding call.

To identify broad themes or trends in the support requests submitted, the RDM Service began to assign a label to each query at the start of 2014. This proved to be difficult due to the wide-ranging and overlapping topics covered in many enquiries, but was helpful in identifying common needs.

Although topics covered in Figure 3 are likely to be encountered by many RDM/IT services across the academic sector, the underlying questions often reflected the specialised environment in which health research is performed. The majority of funder data management plans produced/reviewed by the RDM Service were prepared for submission to health-related funders, such as the Medical Research Council, Wellcome Trust, Cancer Research UK and Economic and Social Research Council; correspondence on the PLOS data sharing requirements are a result of its impact in the medical and health science field; and the large number of requests on data sharing agreements reflect the international environment in which health research is performed. Other requests, such as those related to data storage, security, encryption and sharing also often possessed a health-related element, such as consideration of the practicalities of working with personally identifiable information.

By identifying where support is most needed, an RDM Service can plan work to reduce the number of basic requests that will be submitted on a topic and the time needed to process them. For example, we produced a PLOS Data Policy summary guide (Knight, 2014b) and a worked example of a Wellcome Trust Data Management Plan (Knight, 2014c). In addition, the planned theme of a half-day workshop in November 2014 was changed to focus upon the challenges of sharing health data in compliance
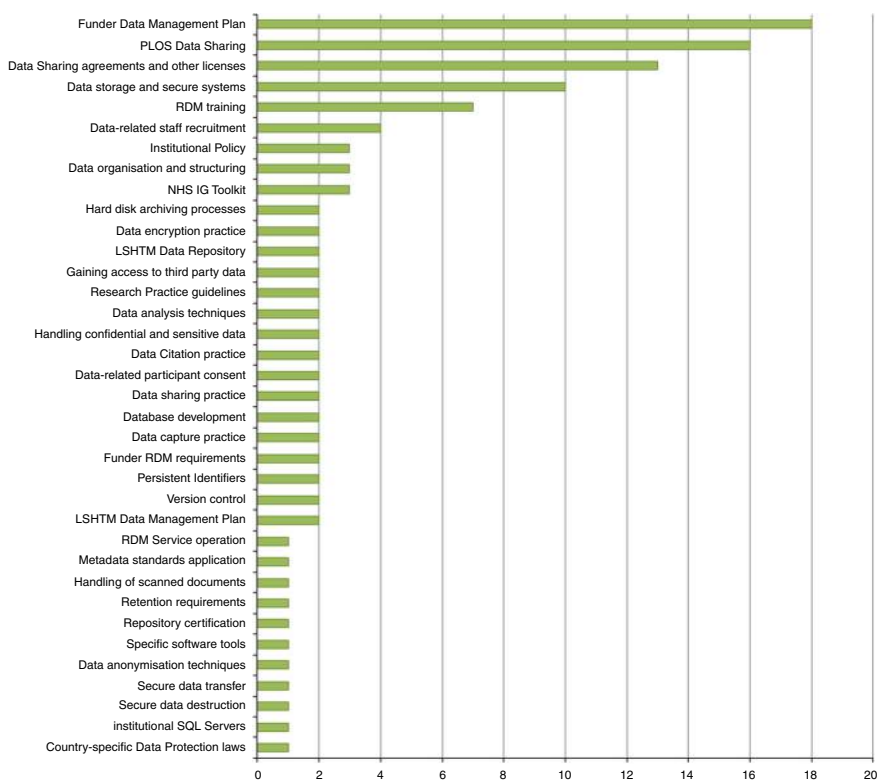
Figure 3.
Themes covered in
RDM support
requests during 2014

with journal and funder expectations. In the event that similar questions are raised at a later date, researchers can be directed to existing resources, reducing the processing time required to handle these queries.

3. Conclusion

This paper has described work performed at the LSHTM to establish a small-scale Research Data Management Service and develop resources suitable for the needs of health researchers. Key to the set-up of a central service was the provision of Wellcome Trust funding, which enabled the recruitment of full-time staff with the necessary expertise to work with researchers and develop RDM resources. It is likely that progress would have been slower if existing library and archives staff had been asked to perform the work as a portion of their existing role. It also benefited from strong management support from the Deputy Director and researchers, who recognised the importance of research data to their work and the need for careful management.

The primary challenge that a small-scale RDM Service must address is how it will support the needs of a large body of academic researchers, while also introducing improvements in practice, in a sustainable and resource-efficient manner. As described, this was addressed at LSHTM by identifying the obligations that needed to be met and resources available from the outset, and planning activities that would allow it to fulfil its objectives in a resource effective manner. At the institution level, the RCUK's expectation that universities implement an RDM policy was used as a basis to

introduce new research practices that ensured funded researchers consider data management from the outset of their project and take appropriate steps to preserve data following its completion. This was supported by the performance of activities necessary to monitor support their needs, which were subsequently used as a basis to inform development plans. The resultant outputs – written documents and training events – enabled efficiencies to be made in the support process by reducing the amount of time that RDM staff must spend on handling repeat queries and improve the quality of information available to researchers over time.

Notes

1. LSHTM's funding bid requested funding for several components, including re-development of lab space. Research data management represented a small component.

2. The Project Manager joined LSHTM in June 2012. However, the recruitment of the software developer was delayed until September 2014.

3. www.rcuk.ac.uk/research/datapolicy/

4. www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies

5. www.eprints.org/software/

6. http://bazaar.eprints.org/226/

7. http://wiki.eprints.org/w/ReCollect

8. www.survey.bris.ac.uk/

9. A document commonly produced in clinical trials and other health-related areas that describes the low-level processes and procedures to be performed in a project, including those associated with the capture, validation, analysis, storage and archiving of data.

10. www.lshtm.ac.uk/research/researchdataman/plan/lshtm_dmplan.html

11. https://dmponline.dcc.ac.uk/

12. www.herts.ac.uk/rdm/planning/data-management-plans

13. www.plos.org/policies/#sharing

References

Alexogiannopoulos, E., McKenney, S. and Pickton, M. (2010), *Research Data Management Project: A DAF Investigation of Research Data Management Practices at The University of Northampton*, The University of Northampton, Northampton, available at: http://nectar. northampton.ac.uk/2736/ (accessed 31 March 2015).

Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H. and Wheatley, P. (2008), "The LIFE2 final project report", available at: http://discovery.ucl.ac.uk/11758/ (accessed 30 January 2015).

Beagrie, N., Lavoie, B. and Woollard, M. (2010), "Keeping research data safe 2", available at: www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf (accessed 30 January 2015).

Beard, L. (2014), "Research data management at the University of Manchester", (Powerpoint presentation). Directions for Research Data Management in UK Universities, 7 November, Moller Centre, Cambridge, available at: www.slideshare.net/JISC/lorraine-beard-rdm-at-the-university-of-manchester (accessed 31 March 2015).

Cope, J. (2013), "University of bath data management plan template", available at: http://opus. bath.ac.uk/36360/ (accessed 30 January 2015).

Cranna, V. (2003), "Development of a retention schedule for research data at the London School of Hygiene & Tropical Medicine", JISC final report, London School of Hygiene and Tropical Medicine, London, available at: www.lshtm.ac.uk/library/archives/records/finalreport_pdf. pdf (accessed 30 January 2015).

Ekmekcioglu, C. and Rice, R. (2009), "Edinburgh data audit implementation project", final report, available at: http://repository.jisc.ac.uk/283/ (accessed 30 January 2015).

Fitt, A., Rouse, R. and Taylor, S. (2015), "RDM: an approach from a modern university with a growing research portfolio", *International Journal of Digital Curation*, Vol. 10 No. 1, pp. 154-162, available at: www.ijdc.net/index.php/ijdc/article/view/10.1.154/388 (accessed 31 March 2015).

Gibbs, H. (2009), "Southampton data survey: our experience and lessons learned", available at: www.disc-uk.org/docs/SouthamptonDAF.pdf (accessed 30 January 2015).

Harrison, S. (2013), "Research data management: analysis of research data management survey", *Royal Veterinary College*, pp. 1-19, available at: www.rvc.ac.uk/Media/Default/Research% 20Data%20Management/RDM%20Survey%20Analysis%20v1.0-1.pdf (accessed 31 March 2015).

Horton, L. and DCC (2014), "Overview of UK institution RDM policies", Digital Curation Centre, available at: www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies (accessed 31 March 2015).

Jones, S., Pryor, G. and Whyte, A. (2013), "How to develop RDM services – a guide for HEI", *Digital Curation Centre*, pp. 1-24, available at: www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services (accessed 30 January 2015).

Jones, S., Ross, S. and Ruusalepp, R. (2009), "Data audit framework methodology", V1.8, available at: www.data-audit.eu/DAF_Methodology.pdf (accessed 30 January 2015).

Jump, P. (2014), "REF 2014 results: table of excellence", available at: www.timeshighereducation. co.uk/news/ref-2014-results-table-of-excellence/2017590.article (accessed 30 January 2015).

Knight, G. (2012a), *Funder Requirements for Data Management and Sharing*, London School of Hygiene and Tropical Medicine, London, available at: http://researchonline.lshtm.ac.uk/ 208596/ (accessed 30 January 2015).

Knight, G. (2012b), "Survey on research data management and sharing: working paper", London School of Hygiene and Tropical Medicine, London, available at: http://blogs.lshtm.ac.uk/ rdmss/files/2012/10/WP4-WebSurveyV1.pdf (accessed 30 January 2015).

Knight, G. (2013), "Research data management at LSHTM: web survey report", London School of Hygiene and Tropical Medicine, London, 8 April, available at: http://blogs.lshtm.ac.uk/ rdmss/files/2013/04/LSHTM-RDM-Web-Survey-Report.pdf (accessed 30 January 2015).

Knight, G. (2014a), *LSHTM Research Data Management Policy*, London School of Hygiene and Tropical Medicine, London, available at: http://researchonline.lshtm.ac.uk/612422/ (accessed 30 January 2015).

Knight, G. (2014b), "The PLOS data policy: an overview", available at: http://blogs.lshtm.ac.uk/ rdmss/plos-data-policy-overview/ (accessed 30 January 2015).

Knight, G. (2014c), "Writing a wellcome Trust Data Management & Sharing Plan", available at: www.lshtm.ac.uk/research/researchdataman/plan/wellcometrust_dmp.pdf (accessed 30 January 2015).

LSHTM Research Data Working Group (2011), "Raising standards – lowering barriers: documentation of, access to and preservation of research data at the London School of Hygiene & Tropical Medicine", available at: http://blogs.lshtm.ac.uk/rdmss/files/2012/07/ LSHTM-RDWG-LoweringBarriersReport.pdf (accessed 30 January 2015).

McHugh, A. (2007), "Digital repository audit method based on risk assessment. version 1.0 (draft)", available at: www.repositoryaudit.eu/download/ (accessed 30 January 2015).

Martinez-Uribe, L. (2008), "Using the data audit framework: an Oxford case study", available at: www.disc-uk.org/docs/DAF-Oxford.pdf (accessed 30 January 2015).

Parsons, T., Grimshaw, S. and Williamson, L. (2013), "ADMIRe research data management survey", University of Nottingham, Nottingham, available at: http://admire.jiscinvolve.org/wp/files/2013/02/ADMIRe-Survey-Results-and-Analysis-2013.pdf (accessed 31 March 2015).

White, W. and Coles, S. (2014), "Relating the research workflow and the publication workflow", (Powerpoint presentation), RDMF11: workflows and lifecycle models for data management, 20 June, University of Westminster, Westminster, available at: www.dcc.ac.uk/events/research-data-management-forum-rdmf/rdmf11 (accessed 31 March 2015).

About the author
Gareth Knight is Project Manager of the Research Data Management Service at the London School of Hygiene and Tropical Medicine. He has a background in digital archives, working with researchers across a range of academic disciplines and departments to ensure digital data are curated and preserved in the long-term. Gareth Knight can be contacted at: gareth.knight@lshtm.ac.uk

This article has been cited by:

1. KoltayTibor, Tibor Koltay. 2017. Research 2.0 and Research Data Services in academic and research libraries: priority issues. *Library Management* **38**:6/7, 345-353. [Abstract] [Full Text] [PDF]