MRC/CSO SPHSU DATA SHARING POLICY

Version 1 30th May 2012

We are committed to maximizing the use of MRC SPHSU data to advance knowledge and welcome proposals for collaborative projects and data sharing which meet the MRC's mission of improving human health. Our policy reflects the MRC's guidance on data sharing.[1] We have developed it with the aim of making data as widely and freely available as possible to bona fide researchers[2] while safeguarding the privacy of participants, protecting confidential data, and maintaining the reputation of the study.

Unit datasets covered by the policy

1. The MRC's policy states that all 'significant' MRC population and patient studies with 'ongoing data collection or analyses', and/ or 'legacy data sets' should have a data sharing policy.
2. At the present time we are including the Unit's large quantitative datasets, with either longitudinal or repeat cross-sectional data collection, in this policy. While we recognise that scientific value can also be gained by sharing qualitative data, the confidentiality and other issues are rather different, and need to be separately considered.
3. What constitutes an MRC study is not always clear, as many studies often have multiple funders; our definition is that fieldwork has been funded as part of MRC's QQR budget or by a MRC grant.
4. In situations where a project is jointly funded, for example, an external grant paid for direct fieldwork costs but significant Unit staff time was also involved unfunded, the member of staff leading project within the Unit needs to establish with other grantholders and funder what the appropriate data sharing policy should be.
5. Table 1 (attached) outlines which studies within the Unit fall under the MRC's data sharing policy.
6. We are taking two broad approaches to data sharing for existing studies in the Unit, although both within MRC framework. First most studies are following a Unit-wide approach to data sharing and a single Unit Data Access Committee (DAC) will oversee this process. Secondly, a few large studies which have their own steering committees are developing their own data sharing policies, which will be overseen by their current governance arrangements. In general these studies have more extensive data sharing activities and a more intensive approach to data sharing.
7. In relation to all new data collection undertaken by Unit staff (but not students), these should include consent and ethical approval for data sharing.
8. Where **funders** allow and the scale of data collection/topicality of subject seems to warrant data sharing, in addition to the main research activities, specific resources should be requested as part of **all** future grant applications and QQR bids for the one off costs of documenting new data

[1] http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/Policy/PHSPolicy/index.htm

[2] http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/Policy/PHSPolicy/inpractice/index.htm#P60_5454

to DDI3 standards, publicising such documentation in a relevant way, and the ongoing costs of data-sharing.

9. At the start of all new studies, a **Data Management Plan(DMP)**[3] should be prepared and agreed with the Study's steering group (or the Unit's Data Access Committee if there is not a study specific committee). A DMP template is available from MRC guideline document but briefly it should include: a short description of the data to be collected in the study, how they are to be stored, what data sharing and governance arrangements will be developed and what privileged use period, if any, will be specified. Where projects are external funded, many research councils now require this plan as part of the funding application. Otherwise, information should be documented as part of Unit's RPRG system.

Approach to data sharing

10. The Unit's data sharing policy requires: a study-specific website; publicly available criteria and partnership agreements for data sharing; a process and application form for considering data access requests, including an internal contact with expertise in each dataset; study documentation and variable metadata and a dataset that can be shared; and an internal governance and monitoring process. Different studies within the Unit are currently at different stages with these components. Table 2 (attached) briefly summarises each study's current situation in relation to these components. All studies plan to be fully compliant with this policy by 31st December 2012 if not before.
11. **A study website** should include:
    general information about the study
    criteria and partnership agreements for data sharing
    application form and process for data access
    a contact for enquiries
    variable level information on the data available

    In addition to holding this information on the Unit website, it is also a requirement that study level metadata, at a minimum are held on the MRC's Data Gateway.

Criteria and partnership agreements for data sharing

12. The following criteria are employed to consider applications for data sharing.
    - The research question(s) proposed should aim to develop understanding of how to improve human health and be of good scientific value, ethically and legally sound.
    - The analytical techniques proposed should be appropriate and robust and the data should be able to address the question(s).
    - The skills and experience of the applicants should be appropriate to undertake the project.
    - There should be an intention to publish the research findings for wider scientific scrutiny and benefit in open access journals.
    - The Unit resources required to create the necessary dataset for the project should be reasonable.
13. The partnership agreement should specify:

[3] http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/DMPs/index.htm

- Whether the project should be undertaken collaboratively, and if so if there is an expectation of joint authorship, and what support external collaborators can expect from the relevant Unit staff.
- The need for the approval of the Unit Director for any publications, presentations or media contact resulting from analyses of Unit datasets. The process for seeking such approval should be specified if no Unit author.
- The requirements of external users for data security – these need to match current Unit policies in managing personal data, and include:
    - Not to give data to third parties nor to attempt to identify specific respondents within the study.
    - Data and related syntax files should be stored on password protected network drives (to ensure regular back up and security). PCs should have automatic password lock features to ensure their security.
    - All data transferred between SPHSU and external researchers, and vice versa, can be sent on a CD or data pen or by email. However, whatever mode of transfer adopted, all data transferred must be encrypted during transfer, and if posted should be sent by registered mail, couriered or hand delivered.
    - If MRC data are to be stored on a laptop while held by external researchers, the device must be encrypted and password protected.

    Depending on the nature of the Study, the Unit custodian may add further security measures, for example employing random id numbers, not allowing external users to have detailed data which are potentially sensitive, and further checks may need to be carried out to ensure no possibility of deductive disclosure.
- What is required of the external user in terms of reporting back to the study on progress, changes in project plans, and what should happen to the data at the end of project ie whether the data should be returned or destroyed, and if documentation on derived variables, cleaning etc is required.

Process for applying for data

14. In general all active studies will have an identified person on their website and other documentation who can be contacted about access to the dataset. This needs to be a recognised part of their current role. Unit leads for each current study are identified in Table 2. Given we have no Unit-wide data management staff, we currently have no capacity for providing data sharing support for studies where there are no active research leads employed in the Unit.
15. Once external users make contact, informal discussions should take place about whether the dataset is appropriate for the research question under consideration, and who might be involved in the project. If appropriate, a formal application should be made. An outline application can be found in Appendix 1. After initial consideration by study-specific staff, the application should be passed to the Unit Data Access Committee for formal approval. The committee should consider whether the project meets the criteria for data access; whether the data security

arrangements are suitable and what Unit resources the project will require. If approved a formal agreement should be established between the Unit and the external users (letter in Appendix 2), and the data released. If there are queries these should be discussed between study lead and external users and hopefully resolved. If external users are unhappy with the process or decision, which can not be resolved satisfactorily by the data access committee, then the matter should be raised with the Unit Director and if still unresolved with the MRC Programme Manager.

16. Individual study leads need to monitor the number of enquiries they receive and their outcome. This needs to be reported to the Unit Data Access Committee annually and included in eVal etc.
17. Individual studies may wish to vary this process and application form etc to suit their particular needs. This should be agreed with the Data Access Committee. The final access process should be described on the study website with timescales for each stage.

High quality data and metadata

18. MRC policy requires that all active studies should ensure their metadata meet a reasonable standard. Current MRC-wide discussion suggests this should be a minimum of DDI 2 standards. http://www.ddialliance.org/ which includes study level summary info; variable level data: name, variable and value labels, missing values, question text, whether derived or not, categories (broad topics), questionnaire name, page number, question number. The data themselves should be cleaned with appropriate documentation. Both study level and variable level metadata, along with broad background information on the study and publications, should be provided on the study website and the MRC Gateway, once this is set up to accommodate this.

Governance process

19. We will establish a Unit data access committee (DAC), which will have five basic functions:
    - To review and approve DMPs for all new primary data collection studies (via RPRG)
    - To review and approve external applications to use data for studies without their own steering group, after the applications has been processed by internal lead(s) for that study.
    - To review annual information on data sharing activities and outputs for all Unit studies.
    - To hold a unit wide database of studies (to be extracted where possible from RPRG) - info on study, consent, location of files, etc, etc.
    - To report to the SRS annually on data sharing and data access issues.

1. The committee proposed is Michaela Benzeval as Chair, the Unit lead of each study covered by it, Geoff Der as Unit statistician and Mary-Kate Hannah as the only (study specific) Unit Data Scientist. We very hope to create a Head of Data services post who can take the lead on all our data sharing activities in the near future.

Internal processes to monitor compliance towards the MRC Data sharing requirements

For new studies as they are set up monitoring by RPRG to ensure they have all basic building blocks up to level appropriate for scale of study. These include:

- Consent and ethics approval for data sharing (main RPRG form)
- A DMP should be attached to the RPRG form and reviewed by DAC.
- Once the privileged use period is over studies need to confirm they have established the necessary components of a data access process: a study specific website, criteria for access and principles of partnership agreements, an application form and process, a metadata catalogue and that information on the Study is included in the MRC's Data Sharing Gateway. (RPRG initial monitoring form, reviewed by DAC)
- Once a study is actively sharing data, this process needs to be monitored by a study specific steering group or the Unit Data Access Committee.

Michaela Benzeval, version 1 30th May 2012

Appendix 1 Generic application form

Proposal for access to XXXXX Study data

Section 1: The proposal

1. Title of the project

2. Anticipated start and end dates of project

Start | | End | |

3. Justification for using Study data for this project.

In addition please attach a 1-2 page outline of your proposal which includes the hypotheses to be tested, methods of analysis planned, and justification for the specific variables requested

Section 2: Research Governance

4. Key External Researcher (designated PI)

Name Institution Address   Telephone Fax E-mail

5. Other external researchers involved in proposed project (and institutions)

6. (Proposed) funder of research

Is funding for this proposal confirmed or pending a decision?

Confirmed ☐ Application Pending ☐ Application to be submitted ☐

If pending when is funder's decision expected? If still to be submitted when is deadline for application?

7. Has this project been (or will it be) peer reviewed?

YES ☐ NO ☐

If so by whom?

Section 3: Data requirements

<will need amending to individual study requirements>

8. General description of variables required

9. Data confidentiality and security requirements

a) How and where will data be stored at your site?

b) If data are to be stored on a laptop at all, please note this and identify the encryption software to be used to protect the data.

c) What measures will be undertaken to ensure respondents' confidentiality is preserved in data storage, analysis and reporting?

d) Who will be able to access data stored at your site?

e) By whom will data be analysed?

f) Should it be necessary to transfer data between team members how will this be done?

Section: 4 Outputs

10. Please specific planned outputs from this project with anticipated timing (publications, presentations, media contact, derived variables, other added value to Study)

Each member of the proposed team who will have access to data should sign a separate copy of this agreement and return it to the proposed link person. The Head of Department of the PI, and the link person,. need to sign one copy.

AGREEMENT FORM

The data in the XXXXXX Study are highly confidential and have been given by respondents on the understanding they will be treated with the utmost confidentiality and respect. All users must ensure that respondents' confidentiality and the reputation of the study are safeguarded at all times.

Failure to uphold this agreement may result in all further access to MRC/CSO SPHSU data by you and your team being denied.

Title of Project ....

....

External Researcher ....

Unit lead on project..... .....

Declaration

I have read the pro-forma on the use of XXXXXX data by external researchers and agree to the conditions therein.	☐
I have also read and agree to abide by the requirements of the SPHSU's policies on data protection and confidentiality and on data management, by the MRC's guide to the handling of personal information in medical research and by the MRC's guides to good research practice.	☐
I will not share the data with any third party other than those who have also signed this agreement. Nor will I attempt to match the dataset covered by this agreement with any other XXXX data I currently hold. I will make no attempt to identify any individual within the study.	☐
I will not submit any papers for publication or presentation or have any media contact about results from XXXXXX data without the prior consent of the SPHSU Director.	☐

At the end date of the project I will return all data, including documentation for any variables I have derived, and destroy all copies of the dataset I hold. ☐

Signature(s)

External Researcher ..... date…….

Unit Lead ..... date…….

Sponsorship of research

I confirm that this institution is willing to act as joint sponsor of this research project with MRC SPHSU and will ensure the confidentiality, protection and appropriate ethical use and management of XXXXXX data at all times.

Head of Department ..... date…….

Project Number ……

Date agreed by Data Access Committee …….

Chair ..... date…….

Collaboration approved

Unit Director…….…….…… date…….

GUIDANCE ON COMPLETING THE APPLICATION FORM

Section 1: The proposal

Please complete the data access application form and attach a 1-2 page proposal of your research plans, which specifies the research questions addressed, the analytical strategy and the variables or set of variables to be used.

Section 2: Research Governance

The MRC is working towards compliance with the Research Governance Framework. To this end, all projects drawing on SPHSU data are required to identify the organisations and individuals who will hold key roles in the project's governance – the principal investigator, sponsor, employer and funder – and who will undertake the associated responsibilities as set out in the Framework. All projects employing SPHSU data will be jointly sponsored by the MRC SPHSU to employ the appropriate and ethical use of the data.

Section 3: Data requirements

Your proposal should specify the subset of variables required for the project. This should be documented as part of the application, and if further variables are required in due course this may require a further application. The data should not be used for any purpose other than stated in the proposal.

Data provided for the use of external researchers at all times remains the property of the MRC. The external researcher(s) will be responsible for maintaining secure storage of data or files provided, for ensuring the confidentiality of any personal information on these files and for ensuring that no data are passed on to any third party at any time. No attempt should be made to identify specific individuals form the dataset.

The external researcher(s) will also be obliged to abide by the requirements of the SPHSU's policies on data protection and confidentiality and on data management (attached), by the MRC's guide to the handling of personal information in medical research (http://www.mrc.ac.uk/pdf-pimr.pdf) and by the MRC's guides to good research practice (http://www.mrc.ac.uk/pdf-good_research_practice.pdf).

All data provided to external researchers are anonymised, and should not include information that allows the possibility of deductive disclosure, if this is possibility analyses must be conducted inhouse. Data for each individual project will be supplied with a random id number system, so that data cannot be combined across different projects to create possible deductive disclosure possibilities.

External users of the data must adopt the same level of data security as is upheld within the Unit.

- Data and related syntax files should be stored on password protected network drives (to ensure regular back up and security). PCs should have automatic password lock features to ensure their security.
- All data transferred between SPHSU and external researchers, and vice versa, can be sent on a CD or data pen or by email. However, whatever mode of transfer adopted, all data transferred must be encrypted during

transfer, and if posted should be sent by registered mail, couriered or hand delivered.

- If MRC data are to be stored on a laptop while held by external researchers, the devise must be encrypted and password protected.
- At the end of the project the dataset should be destroyed.

Section4: Outputs

a) Publications

We are keen to see the work undertaken by or for external researchers leading to publications. Unless otherwise indicated, it will be assumed that this means joint publication with one or more members SPHSU. Normally, it will be the case that the external researcher takes the lead in writing up (literature review, etc.) and therefore is entitled to be first author(s). However, authorship should be discussed on a paper by paper basis and based on the Unit's policy on authorship, attached. Where ever possible papers should be published in Open Access Journals in line with the MRC's policy.

In exceptional circumstances, the Unit staff involved reserve the right to disagree with the conclusions of the external researchers and withdraw their names from any publication. This would not necessarily preclude external researchers from publishing in their own right. Although we reserve the right to veto publication in certain circumstances, for example, to prevent deductive disclosure or if it is felt publication might jeopardise future waves of interviews, such consent will not be unreasonably withheld.

All presentations and publications based on SPHSU data, and any associated media contact, must have the **prior** approval of the Director of the Unit.

All publications based on analyses of the XXXX study must include an acknowledgement of the Study.

> *The XXXXX Study is funded by the UK Medical Research Council and the data were originally collected by the MRC Social and Public Health Sciences Unit (MC_A540_XXXXX). Information on how to apply for access to the data can be found at:* studywebsite *We are grateful to all of the participants in the Study, and to the survey staff and research nurses who carried it out. The data are employed here with the permission of the SPHSU Data Access Committee (Project No. XXX). XXXX are also funded by the MRC (MC_A540_XXXX).*

B) Derived variables

Derived variables created by the external collaborator which may be of wider value to users, should be returned to the study together with appropriate documentation (e.g. coding frames or syntax used).

Agreement

The external researcher(s), their head of department and the Unit study lead should sign the attached Agreement form, indicating they have understood and will meet these obligations.

Failure to uphold this agreement may result in all further access to SPHSU data by you and your team being denied.

Appendix 2 Letter of agreements – should be tailored to individual studies

Address

Date

Dear XX

Application number & title

Your proposal to access XXXXXX data for the above project has been approved. I am enclosing a copy of your signed agreement confirming this, which describes the terms of the agreement under which you can hold the data.

Please liaise with XX to provide you with the data that you require for your project and to keep us informed of your progress. Should your analysis plans change significantly from those specified, you may need to submit a further application to the Unit. I'd be happy to advise you about this.

All publications based on analyses of the XXXX study must include an acknowledgement of the Study.

> *The XXXXX Study is funded by the UK Medical Research Council and the data were originally collected by the MRC Social and Public Health Sciences Unit (MC_A540_XXXXX). Information on how to apply for access to the data can be found at:* studywebsite *We are grateful to all of the participants in the Study, and to the survey staff and research nurses who carried it out. The data are employed here with the permission of the SPHSU Data Access Committee (Project No. XXX). XXXX are also funded by the MRC (MC_A540_XXXX).*

The Unit Director, must give prior approval for all papers – oral and written - based on XXXXX data, and any media coverage of the Study.

Good luck with your project; please do let me know if you have any further queries.

With best wishes,

Yours sincerely

Study lead