

TEDS Data Access Policy

29 May 2019, latest update 11 December 2020

Contents

1. Introduction
2. Principles of access to TEDS data
3. Governance of access
4. Time frames for processing data requests
5. Data sharing agreements
6. Pre-registration of analysis
7. Guidelines for use of the data
8. Guidelines for publication of results
9. Costs
10. Analysis of genetic data
11. Linked data
12. Exclusive use of data by the TEDS research team
13. The TEDS data dictionary
14. Summary: how to request data
15. Resources

1. Introduction

The 2010 MRC Review of Mental Health Research notes that 'TEDS is used by many researchers and has been a platform for a wide range of published studies.' More than 400 collaborators have carried out research using the TEDS resource, which we expect to continue with the extensive data that we have collected at age 21 and whose value is exponentially increased by the previous 20 years of data on this unique sample. These collaborations provide the best possible peer review of data at the practical level of using the data for analyses and publications. We have many emails from collaborators extolling the user-friendliness, easy access and high quality of the TEDS dataset. The major source of discovery of the TEDS resource has been through the many TEDS publications, especially a series of papers that outline TEDS and its measures ([Trouton et al \(2002\)](#), [Oliver et al \(2007\)](#), [Howarth et al \(2013\)](#), [Rimfeld et al \(2019\)](#)).

We encourage and facilitate the sharing of TEDS data in collaborative academic research projects, subject to the terms described in the policy below. We aim to comply with the data sharing policies of our funding body, the Medical Research Council (MRC).

There are two aspects to the TEDS study which make it a particularly valuable source of data. The first is that the participants are all twins, enabling the estimation of genetic and environmental influences on traits and their associations. The second is that the data have been collected over a period of 25 years. We would expect most users of TEDS data to incorporate twin analyses into their work, but we also recognise that our prospective design results in a dataset that is invaluable for exploring longitudinal hypotheses, thus we are also open to working with researchers for whom the key question is phenotypic (e.g. associations over time) rather than about establishing aetiology.

2. Principles of access to TEDS data

Requests for TEDS data will be considered under the following general conditions. More specific details are given in the sections below.

- The data are requested by a bona fide researcher affiliated to a recognised academic research institution.
- If the researcher is not a KCL employee, then TEDS data will only be shared subject to a formal data sharing agreement (see below).
- The purpose of the project is for scientific academic research, compatible with the wider aims of TEDS.

- The project, if successful, will have a tangible and measurable research outcome. Usually, this will be the publication of a paper in a peer-reviewed academic journal.
- The data are requested for a specific analytical project with clearly defined aims and using statistical methods that are achievable with the available data.
- The request is submitted with the agreement and involvement of a member of the core TEDS team (as listed on the [data request form](#)) who will be a co-author in any resulting publication.
- The aims and objectives of the project do not overlap with other TEDS research projects already underway (refer to our list of [current data requests](#)).
- The applicant does not have any other current and unfinished TEDS data request. Exceptions to this rule are made only for internal TEDS researchers, who typically work concurrently on more than one research project. Core TEDS collaborators may be named on multiple data requests.
- Each data request should produce only one publication (if publication is the principal objective). Applicants aiming to publish more than one paper using TEDS data must complete one data request before starting another.
- Applicants wishing to use TEDS data for an unpublished dissertation must submit a data request in a similar way as for a publication.
- If analysis is to be carried out by a research student (e.g. a PhD or BSc student) then the applicant should be the student's supervisor.
- Applicants wishing to use TEDS data for a grant application must submit a data request before writing the grant application, as a method of gaining prior permission to use of the data in the event that the grant is funded (see below).
- The submission of the data request and the use of the data comply with the detailed guidelines on this page.
- The applicant has not previously been in breach of our guidelines in relation to an earlier data request.

A data request for a grant application need not necessarily involve sharing a dataset for analysis before the grant is funded; if no dataset is required for the application, then of course no data sharing agreement will be required at this stage. However, the data request form should clearly state details of the proposed research, including the research objectives such as expected publication(s); the measures that would be needed for the research, without necessarily providing a full list of named variables; the types of analysis expected to be used; and the names and locations of collaborators who are expected to be involved in analysis. The end date for the initial data request will be the submission date for the grant application. If and when the grant is funded, the applicant must submit further, more detailed, data requests in the usual way for projects aimed at publication.

Before proceeding with a data request, an applicant should make full use of the resources on this web site (see also section 15 below) for planning their research project as follows:

- Ensure that the research will make appropriate use of the TEDS twin dataset, for example using twin model-fitting analyses and/or analysing longitudinal associations. Please note that in the case of two closely related applications, one that incorporates use of the twin aspect of the data will be granted in preference to one that does not.
- Confirm that appropriate measures or variables are available for the planned research, by searching the [TEDS data dictionary](#). There are some limitations due to exclusive use of new data by the TEDS team (described in section 12 below).
- Check that there is no obvious overlap with other current projects, by searching the list of [current data requests](#).
- Search the [TEDS publications](#) to discover whether similar projects may have been completed in the past. Overlap with earlier, completed projects may be permitted in some circumstances, for example for replication or meta-analysis or for use of novel analysis methods.

3. Governance of access

Before submission of the data request, an applicant is required to obtain the agreement of a member of the core TEDS research team to collaborate. The reason for this requirement is to ensure that the project is conducted in accordance with the conditions described in this document. The list of core TEDS team members is available on the downloadable [data request form](#) (Word document: right-click the link and select Save Link As ...). The core TEDS team member must be consulted over the completion of the data request form and must be a named co-author on

any publication resulting from the project. The core TEDS team member may also be able to advise on analysis methods but need not be involved in actual analysis of the dataset.

On submission, the data request will be reviewed by a TEDS steering committee comprising the TEDS Director (Prof. Robert Plomin), the TEDS Deputy Director (Prof. Thalia Eley) and the TEDS Data Manager (Andrew McMillan). The data request may then be accepted if it meets the conditions outlined in this data access policy. If the data request is not accepted, then the grounds for rejection, based on the principles above, will be made clear by the committee.

If the application is successful, then access to the dataset will be for a pre-agreed period of time, not longer than 24 months, as described in section 4 below.

In the event of rejection, the applicant may appeal against the decision. Independent and external oversight of the process will be provided by Alison Pike, Professor of Child and Family Psychology at the University of Sussex. Prof. Pike will rule on the appeal and will inform the applicant and the steering committee of the outcome.

A data request, once accepted, may be terminated by the steering committee at any stage if the applicant fails to follow the guidelines detailed below. For example, an application may be terminated if the applicant changes the focus of the research from the agreed aims, or if the dataset is shared with any third party not named in the agreement, or if data security rules are breached. In the event of termination, the applicant will be required to destroy the dataset and may not publish any results obtained by analysis of the dataset – these terms will be set out in the data sharing agreement (see below). An applicant who breaches the guidelines is unlikely to be successful in any subsequent data request.

4. Time frames for processing data requests

The steering committee may request amendments or clarifications in a submitted data request form. Once the final version has been received, the steering committee will usually be able to give an initial decision of acceptance or rejection within two weeks. The process may occasionally take longer than two weeks, for example at busy periods or if the data request has some unusual requirements.

When a data request has been accepted, and if the analysis is to be done outside KCL, the next stage will be to set up a data sharing agreement - see section 5 below for further details. This agreement will be initiated with the contracts department at KCL, after which it will be the applicant's responsibility to secure completion by their own university's contracts department.

In addition, after the data request has been accepted, the applicant will be required to pre-register their intended analysis with the Open Science Framework (<http://help.osf.io/m/registrations/>) and return a private link to the steering committee. The link need not be made public during the project but must be included in any publication resulting from the project. Pre-registration should take very little time although the onus for this is on the applicant.

Once the data sharing agreement has been finalised and received by the steering committee and the analysis has been pre-registered, we will aim to provide a dataset for the applicant's project within a further two weeks. For a KCL project that does not require a data sharing agreement, we will aim to provide a dataset for the applicant's project within two weeks of the completion of OSF pre-registration.

This target of two weeks may be increased if the applicant has unusual data requirements that require additional processing beyond selecting variables that are listed in the data dictionary.

The period of authorisation of the data request will start from the date on which the data request was agreed. The expected date of completion must be submitted by the applicant in the data request form and will be written into the formal data sharing agreement when confirmed by the steering committee. This end date must be no more than **24 months** from the start date, but may be shorter if preferred by the applicant. The reason for this time limit is to avoid data being sequestered and then not used. This allowance of 24 months is designed to allow 12 months for analysis and drafting a manuscript, followed by up to 12 months more to submit and get acceptance by a journal.

Applicants with objectives other than publication, for example for a grant application or for a non-published dissertation, will normally be expected to agree an end date within 12 months of the start date.

It is expected that analysis of the data and drafting of a manuscript will be completed within the first **12 months** of this period. Therefore, after 12 months, the TEDS steering committee will request a progress report, and at this stage we would normally expect to see a manuscript ready for submission to a journal. The steering committee will terminate a project not showing sufficient progress at this stage.

If, after analysis, the applicant decides not to publish the results, then this decision should be reached within the first 12 months. In this event, the end of the project will be brought forward and the applicant will be asked to delete the dataset.

If the objective of the data request was an output other than publication, then we would normally expect to see evidence of completion by this 12-month stage.

If the objective was publication then we will expect all steps to have been completed before the agreed end date (normally 24 months after the start). This will include submitting the manuscript to a journal, responding to reviewers' comments and making necessary revisions, and acceptance for publication by the journal.

When the end date has been reached, regardless of other factors, the project will be brought to an end. The applicant will be required to delete the TEDS dataset and any copies that may have been made, and to return a form confirming in writing that this has been done. Summary statistics from analysis can be retained, of course, provided that they do not contain any data at the level of individual TEDS subjects. If the applicant wishes to archive the dataset for potential future reference, with any new variables created and syntax files used to process the data, then this archiving can be done within KCL if the files are returned securely to us (using an encrypted data transfer).

There will be **no extensions** to a data request beyond the agreed end date. The end of the project, and the deletion of the dataset, will be legal obligations stipulated in the data sharing agreement.

We recognise that acceptance for publication can sometimes take an unexpectedly long time, and that in some cases re-submission to another journal may be necessary. After the agreed end date, although the applicant will have been obliged to delete the dataset, they may of course continue to seek publication of their manuscript. If further analysis were then required before re-submission, then this would necessitate a fresh data request and a new data sharing agreement (normally lasting at most 12 months not 24 months). In this event, if the applicant had sent their original working dataset for archiving at KCL, then it would still be available for a new data request for this purpose.

5. Data sharing agreements

It is the policy of KCL and of the MRC (our funding body) that access to datasets by collaborators who are not employed at KCL must be governed by data sharing agreements. Such an agreement is required whether the collaborator is sent a copy of the dataset to analyse at their own university, or whether the collaborator visits KCL (as an affiliate) for the purpose of accessing the data.

A data sharing agreement is a formal and legally-binding contract between KCL (which governs TEDS) and the university at which the applicant is employed. Such a contract, when used, will formalise the agreement of the applicant to comply with the guidelines that are listed on this page (see for example section 7 below). The agreement must be issued and signed by the appropriate authorities before any data are released and before any analyses are performed by or on behalf of the applicant.

This agreement will be initiated with the contracts department at KCL, after which it will be the applicant's responsibility to secure completion by their own university's contracts department.

If the applicant wishes the dataset to be analysed by collaborators from more than one university (other than KCL), then a data sharing agreement will be required with each respective university. This will include core TEDS collaborators, if they require access to the dataset and if they are not employed by KCL.

If the analysis is to be carried out by a research student (typically a PhD student or master's degree student), then the data sharing agreement must be arranged in the name of the student's supervisor, who will thereby agree to ensure that the student complies with the TEDS guidelines. This is necessary because research students are typically not university employees.

Collaborators at other universities who do not require access to the dataset for analysis do not need a data sharing agreement. Typically, for example, an applicant may name collaborators who will assist with interpreting findings, writing and reviewing the manuscript without themselves carrying out analysis of the data. This will often be the case for the core TEDS collaborator who is named on the data request. The applicant may share summaries of analysis results with such collaborators, as long as these results do not contain any data at the level of individual TEDS participants.

6. Pre-registration of analysis

When a data request has been authorised by the steering committee, and before a dataset is released, the applicant will be asked to pre-register their intended analysis with the Open Science Framework (OSF): <http://help.osf.io/m/registrations/>. A private link to the pre-registration must then be provided so the steering committee can confirm that the analysis has been agreed. The purpose of this condition is to promote good research practice and to help discourage 'mission creep', i.e. changes to the agreed objectives of the project.

In exceptional cases, where the purpose of the data request is not primarily to test a hypothesis for publication, or where the purpose is for writing an unpublished thesis, an OSF pre-registration may not be appropriate. This might be the case, for example, in a data request for a grant application. If the applicant wishes to make a case for omitting OSF pre-registration, they should do so in their data request.

The pre-registration link need not be made available to the public, until such time as the results are published. If the project is successful and leads to publication, then we require that the link to the pre-registration is included in the publication.

7. Guidelines for use of the data

The applicant for a data request, if successful, must agree to comply with the following conditions. These conditions are designed to ensure the security of the data and the integrity of the research. For example, the guidelines aim to ensure that there is no overlap in the research being carried out by different collaborators, and to prevent collaborators from sequestering TEDS data without actively using it. Some of these conditions will be made legally binding in the formal data sharing agreement. The steering committee will reserve the right to terminate the project, and to withdraw permission to publish, if the dataset is used in any unauthorised way.

- The data request must be submitted in accordance with the procedure outlined above. For example, the project must involve a member of the core TEDS team, who has agreed to be named on the data request, and the project must be completed before the stated end date.
- Every effort will be made by the TEDS data manager to ensure that the data are anonymous. For example, the participant identifiers will be irreversibly encrypted, and potentially identifying variables such as birth dates and postcodes will be removed. It will not be possible to link the dataset with other TEDS datasets.
- Even though every effort will be made to ensure anonymity in the data, collaborators must agree not to make any attempt to identify any TEDS participant in the dataset.
- The dataset will be sent to the collaborating researcher(s) via the KCL file transfer service or by other secure means.
- During any subsequent transfer of data, including the return of datasets to TEDS for archiving, the dataset must be encrypted. This can be done, for example, by placing the dataset within a password-protected 7-zip archive (<https://www.7-zip.org/>) before transfer.
- The only researchers given access to the dataset will be those named as requiring this access in the data request form and confirmed by the data sharing agreement. These collaborators will all agree to abide by the same conditions governing the use of the dataset. The dataset will not be shared with any third parties.

- The dataset will be customised so that it contains only the variables agreed in the data request, consistent with the objectives of the research. Collaborators must agree not to attempt any unauthorised linkage with other datasets, especially as such linkage would increase the risk of identification of participants.
- The collaborator(s) will permanently delete the dataset when the data request reaches its end date, with the option of returning a copy to TEDS for archiving. The applicant will be required to sign and return a form confirming deletion. For any subsequent project, the applicant must request a new dataset and must not attempt to re-use a dataset previously supplied.
- During use of the data for analysis, the collaborator(s) must store the dataset securely where it cannot be accessed by unauthorised users. Suitable examples would be secure network drives, maintained by the applicant's university IT department, and with access given only to authorised users; a hard drive on a desktop computer within secure university premises and with appropriate authorised access; and encrypted drives on portable devices. **The dataset must not be stored on portable devices (such as laptops and USB drives) unless these are encrypted.**
- Any external collaborator, i.e. a researcher who is not a member of the TEDS research team, may have no more than one concurrent data request. If an applicant wishes to plan more than one publication based on the TEDS data, then this should be done sequentially such that the first data request is concluded before the second data request is submitted. Hence, each data request should produce only one publication.
- A data request for a grant application is a special case, described in section 2 above. This need not necessarily require the transfer of a dataset, if analysis results are not needed within the application. The grant application may also refer to more than one prospective publication, to be covered by subsequent data requests if and when the grant were to be funded.
- The collaborator must not deviate from the aims and focus of the research project as stated in the data request.
- At the end of the project, and prior to submission for publication, the manuscript must be sent to louise.webster@kcl.ac.uk for checking (see section 8).

8. Guidelines for publication of results

Any publication arising from the data request must comply with these conditions:

- The core TEDS team member named on the data request must be named as a co-author and must be involved in the usual way in the writing and review of the manuscript.
- The paper must include an acknowledgment of appropriate MRC grants that were used to fund TEDS data collections relevant to the dataset used. We can supply the necessary wording and grant codes.
- It is a requirement of our MRC funding that any publications resulting from analysis of TEDS data must be made open access. If the journal requires a payment for open access, then the cost must be met by the applicant and not by TEDS. However, most journals will allow you to deposit the accepted manuscript (but not the final published version) on *Europe PubMed Central* (<https://europepmc.org/>) six months after publication, free of charge. This method of providing open access is approved by the MRC.
- The main TEDS reference paper (DOI: [10.1017/thg.2019.56](https://doi.org/10.1017/thg.2019.56)) must be cited in the reference list of the publication.
- The publication should include your OSF link, for the analysis that was pre-registered at the start of the project.
- **Before submission**, we require you to send your manuscript to Louise Webster (louise.webster@kcl.ac.uk) to ensure the funding is appropriately acknowledged and to clarify your publication plans, in order to verify that the publication will be open access.

Some journals request that datasets used in a publication are deposited in a publicly available resource. This is not permitted under this Data Access Policy, and a collaborating researcher may not agree to publish results under these terms. Instead, in such cases, the following statement may be made in the submission: "Data used for this submission may be made available on request to the Twins Early Development Study (TEDS), through their data access mechanism (see <http://www.teds.ac.uk/researchers/teds-data-access-policy>). We will then consider requests for sharing data for appropriate research purposes."

9. Costs

At the present time, no financial charges are made for data access. This is currently under review and we may in future introduce financial charges to recover the administrative costs of providing access to the data.

10. Analysis of genetic data

The TEDS resource includes genotypic data derived from DNA samples that have been collected from many of our participating twins. For practical and ethical reasons, the raw genotypic dataset cannot be shared outside KCL. Individual level genotypic data are regarded as sensitive/personal information under GDPR, and our TEDS twins have not provided consent for us to share such data beyond the core TEDS team.

The raw genotypic data may be analysed for the purpose of collaboration but only under the following circumstances. Firstly, the purpose of analysis must not simply be for a GWAS (genome-wide association study) because such analyses, based solely on the TEDS dataset, would be underpowered. GWAS analysis will only be considered for consortia collaborations where replication is inherent. Secondly, the data can only be analysed locally at KCL, on our secure network, and we will not share copies of the data externally. Thirdly, as a consequence, it would be necessary to secure agreement between the applicant and an internal TEDS researcher, who would act as a collaborator and help carry out analysis of the data. Such an agreement would necessarily involve a significant cost in terms of working time and effort for the given TEDS researcher. As such, it would only be feasible if the aims of the project were consistent with the research aims of the individual TEDS researcher. A data request involving analysis of the TEDS genotypic data may therefore be rejected on the grounds that no TEDS researcher is available or willing to collaborate on the project. Any applicant wishing to explore this possibility should make an informal enquiry first, in order to establish feasibility. In due course we hope to find a way to fund an analyst that could undertake such projects.

Data derived from the TEDS genotypic dataset may be shared with external researchers in some circumstances. GWAS summary statistics related to published results may of course be shared by the publishing author. Polygenic scores, derived from the raw genotypic data for a range of phenotypic traits, are becoming available in the TEDS dataset (see section 12). As these are derived variables and do not comprise raw genotypic data, they will be shared with external researchers in due course.

11. Linked data

Some TEDS research has been carried out by linking data collected in TEDS with data from third parties. In the recent past, this has included *National Pupil Database (NPD)* data linked to TEDS twins by the *Department for Education (DfE)*. In future, we hope also to create links to NHS medical data. The suppliers of these linked data have stipulated that they are only to be used for agreed purposes by named researchers with TEDS, and as such they may not be shared with any external collaborators. Therefore, until such time as we obtain permission to share these linked datasets, they will not be shared with collaborators (and they are not described in the data dictionary).

Some of the TEDS twins and their families have in the past agreed to participate in *spin-off studies*, also called *sub-studies*. Examples of these sub-studies are listed in the data dictionary at <https://www.teds.ac.uk/datadictionary/studies/spinoffs.htm>. These sub-studies have involved relatively small, often highly selected, sub-samples of the TEDS families (typically tens or hundreds of families). In most cases, with the consent of the families, these sub-studies have collected and managed their own datasets independently of TEDS. Therefore, such datasets are not available for sharing with collaborators through TEDS, and they are not fully documented in the data dictionary.

Some TEDS research has involved linking participant postcodes to geographical context data. Full postcodes are not available for sharing with external collaborators, because in many cases the postcode could easily lead to identification of TEDS twins and their families. In the past, where the research aims have been suitably justified, we have occasionally agreed to allow linkage to the first part of the postcode, which covers a broader geographical area

and which is less likely to permit identification of participants. Any applicant wishing to explore this possibility should make an informal enquiry first, in order to establish feasibility.

12. Exclusive use of data by the TEDS research team

The TEDS research team is funded by the MRC to deliver a specific programme of research. The TEDS data collections have been designed to enable our researchers to meet the aims of this programme. Each data collection has been completed only after a significant input of intellectual and managerial effort directed at acquiring the necessary funding, forming scientific hypotheses, designing appropriate data instruments, administration of the actual data collection, processing and cleaning the data, documentation, and so on.

Therefore, the TEDS research team reserves the right to a limited period of exclusive use of newly collected data measures. This is necessary for the TEDS team to prepare, process and document the data, analyse the data and publish results. For example, at the time of writing, these arrangements apply to measures newly collected in the TEDS age 21 data collection, and to polygenic scores newly created from genotypic data.

After each new data collection has ended, a period of approximately 12 months is needed to transform and clean the data into a form suitable for analysis, and to document the data collection fully in the TEDS data dictionary. This period of time is variable depending on the complexity of the given study. Overlapping with this but extending beyond it is the time needed by TEDS researchers to fulfil their initial plans to analyse the new measures and publish results. Therefore, our policy is that new measures will be made generally available for sharing no more than 3 months following the first publication of findings based on the data from those measures (this includes online pre-print publications). Furthermore, whether results are published or not, all new data will become available for sharing 2 years after the end of data collection.

Our age 21 “TEDS-21” data collection ended in February 2019. Further post-21 data collections, including the “g-game” and “covid” studies, were commenced during 2020. The processing and documentation of these data are underway. TEDS researchers are analysing the age 21 measures and have plans for publication. Applying the rule as stated above, each age 21 measure will become available for sharing when 3 months have elapsed after publication of results from that measure by TEDS researchers. Regardless of publication, the “TEDS-21” measures will become generally available for sharing in February 2021, when 2 years have elapsed since the end of data collection.

13. The TEDS data dictionary

Metadata describing the TEDS data collections, measures used and dataset variables can be found in the data dictionary at www.teds.ac.uk/datadictionary. This is a web site of linked html pages and pdf documents that can be viewed in any web browser (with appropriate software for reading pdfs). The data dictionary can be navigated by means of links from its home page.

The data dictionary is a comprehensive description of the TEDS data. It contains, among other things:

- Descriptions of the administrative methods used for each data collection, including the selection of the subsample, timing of the study, consent mechanisms and types of contacts made with participants.
- A summary of data returns for each collection, providing an approximate N for each dataset.
- A list of measures used in each data collection, with references where appropriate.
- A summary measures page illustrating the patterns of measures collected across different ages.
- Copies of the items used in all data collection instruments including questionnaires (usually in pdf format) and web tests.
- Annotated copies of questionnaires to show the variable names used for items and the numeric codes used for categorical responses.
- A variable list for each data collection, including both item and derived variables and showing properties such as variable names, brief descriptions, variable type and coding.
- A description of how derived variables were created for each data collection.

- More general pages including a description of standard analysis exclusions used by TEDS researchers, a glossary of terms used within TEDS, comparisons of data returns at different ages, and broad descriptions of how data files have been processed in the construction of datasets.

An applicant searching for measures and variables, for whatever purpose, is advised to start with the summary measures page: <https://www.teds.ac.uk/datadictionary/studies/measures/measures.htm>.

From this page, there are links to more detailed measures pages for specific data collections. Item variables are often most easily identified by means of the annotated booklets, and relevant derived variables may be identified from the measures pages and the derived variables pages for each data collection. Full variable lists, for each data collection, are also included for reference but, as the number of variables collected is very large, a variable list may not be the best place to start.

14. Summary: how to request data

Researchers wishing to use TEDS data should download and complete our [data request form](#), which is provided in Microsoft Word (.docx) format. To download the form, right-click the link and select "Save link as ..."

The form contains guidelines for the procedure: please follow these closely when completing the form. The procedure is also outlined below.

The applicant must be a bona fide researcher affiliated to a recognised academic research institution. If the research is to be carried out by a PhD student or other junior researcher, then the application must be submitted by the student's supervisor, with the student named in the application. The proposed research must be consistent with the broad aims of TEDS, KCL and the MRC (who fund TEDS) of conducting research to promote scientific understanding and to improve public health.

The applicant should formulate research aims that are compatible with the available variables, as documented in the [data dictionary](#).

The applicant should check our [list of current data requests](#) to ensure that the proposed project does not obviously overlap with research currently in progress. The applicant should search our [searchable publications list](#) to gain an understanding of what has already been achieved by analysis of the TEDS data and to avoid simply repeating analysis that has previously been done (although projects involving meta-analysis or new analytical methods, for example, may of course overlap with previous research). The applicant should also consider, with advice from the core TEDS team member, whether the proposed project is feasible using the available variables and measures as described in our data dictionary; and whether the proposed research makes appropriate use of the unique nature of the TEDS twin data.

Having formulated a plan for research, the applicant must proceed by obtaining the explicit agreement of a member of the core TEDS team to collaborate on the project and to be a co-author on any resulting publication. The names and email addresses of the core TEDS team members are listed in the data request form.

The next step is to complete the data request form itself. The required details include the following:

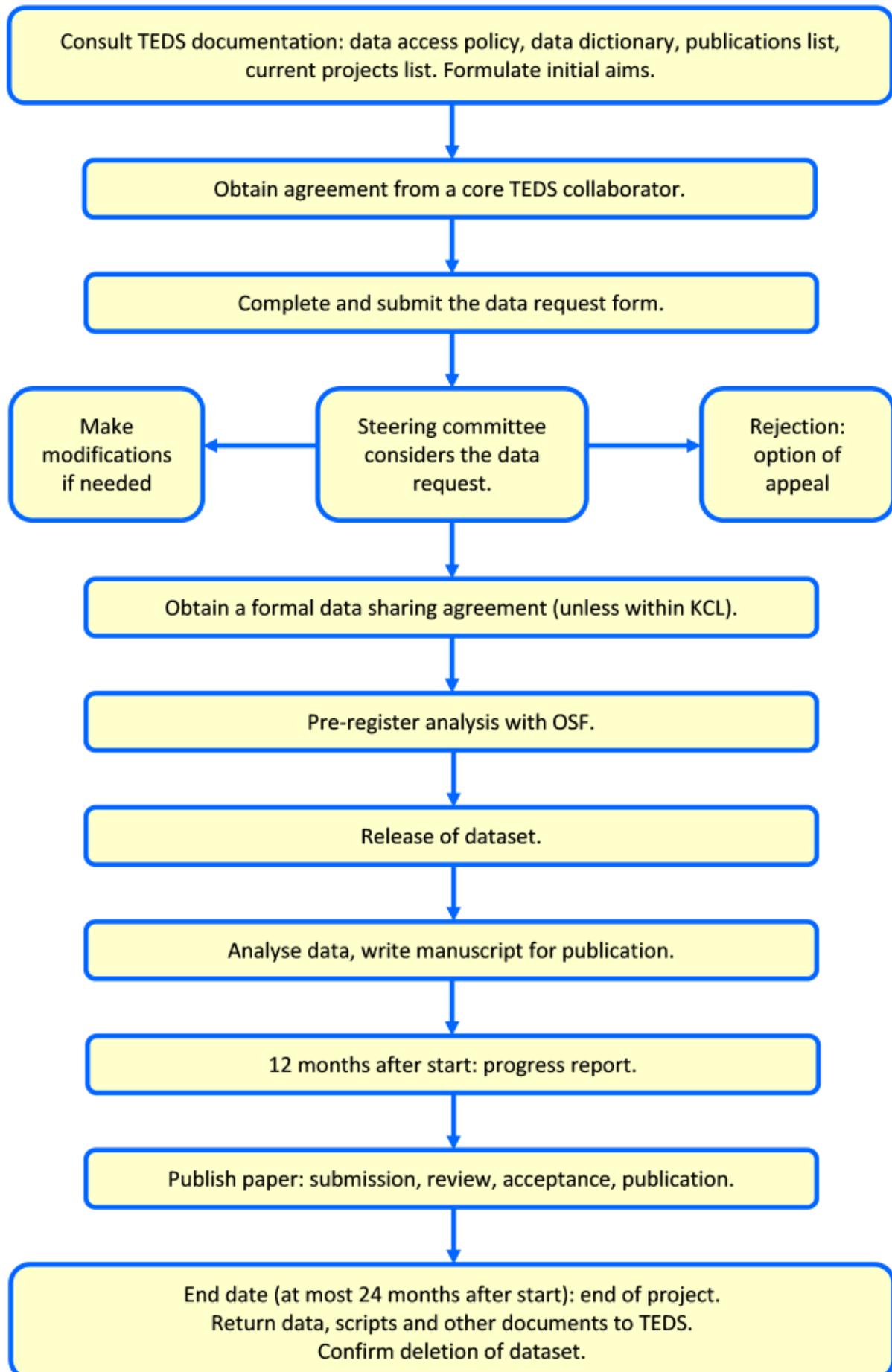
- a descriptive title for the project
- a more detailed description of the project including aims and objectives, specific analysis methods and potential implications
- the name of the core TEDS collaborator who has agreed to collaborate on the project
- the name and email address of the applicant
- the names and email addresses of any other collaborating researchers, specifying which (if any) of these will require access to the dataset for analysis
- a completion date for the project
- details of anticipated publication plans
- an unambiguous list of the measures/variables required in the dataset for your project

If the data request is accepted, certain responses in the data request form will be made publicly available in the [list of current data requests](#). This is designed to help other potential researchers to see what research is already in progress and to help avoid overlapping aims in new data requests. The published fields are marked on the data request form.

Once the data request form is completed, and before it is submitted to TEDS, it must be sent to the named core TEDS collaborator for approval and agreement. It should then be sent as an email attachment to andrew.mcmillan@kcl.ac.uk, with the named core TEDS collaborator copied in as an email recipient.

The data request will then be considered by the TEDS steering committee, as described above. Clarifications and alterations may be required before a final decision is made by the committee. Additional details of the process are described in the sections above.

Flow chart to illustrate the TEDS data request life cycle
For a typical 2-year researcher collaboration aimed at publication



15. Resources

Informal enquiries about sharing TEDS data should be addressed to Andrew McMillan: andrew.mcmillan@kcl.ac.uk.

[TEDS data request form](#) (Microsoft Word format – right-click the link to download). To request data, return the completed form as an email attachment, to andrew.mcmillan@kcl.ac.uk.

Before submission for publication, manuscripts should be sent to Louise Webster: louise.webster@kcl.ac.uk.

[TEDS data access policy](#) (pdf format – right-click the link to download). This is a downloadable version of the policy as described on this web page.

The TEDS data dictionary: www.teds.ac.uk/datadictionary. This provides a full and complete description of the TEDS data collections and datasets, including variable lists.

Searchable list of current TEDS data requests: www.teds.ac.uk/datarequests. This provides details of research projects currently underway.

Searchable list of TEDS publications: www.teds.ac.uk/publications. This provides a full list of papers published by TEDS researchers and collaborators, using results from analysis of the TEDS data.

The TEDS reference paper, to be cited in publications, is [Rimfeld et al \(2019\)](#).

OSF guides and help pages for pre-registering projects are linked from this page: <http://help.osf.io/m/registrations/>.