

## **The Cleft Collective Cohort Studies Data Access and Publication Policy**

### Introduction

This document summarises the data access policy for The Cleft Collective Cohort Study. A large DNA backed prospective resource for the study of the genetic and environmental determinants of cleft lip and/or palate and the long-term outcomes in children with cleft lip and/or palate”.

This study involves data collection from a large cohort of children born with cleft lip and /or palate and their families across the UK. It provides detailed longitudinal information relating to the social, emotional, behavioural and cognitive development of the child, and the social, economic, psychological and health status of the mother and her partner. In addition, participants provide questionnaire-based information on pre/peri-natal environmental exposures as well as biological samples including DNA.

Ethical approval for the study was obtained from the South West NRES (REC 13/SW/0064).

This study is being conducted by the Bristol Dental School, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN.

### Study Management

Professor Yvonne Wren is the Chief Investigator of the study. The day-to-day running of the study is the responsibility of the Project Management Group (PMG). This group currently comprises Professor emeritus Jonathan Sandy, Ms Kerry Humphries, Professor Yvonne Wren, Professor Sarah Lewis, Professor Evie Stergiakouli, Professor Kate Northstone, Professor Sam Leary and Miss Amy Davies. If you would like the PMG to consider a proposal for the use of existing data or collection of new data or to approve a paper for submission please follow the guidance below and email your proposal or paper to [cleft-collective@bristol.ac.uk](mailto:cleft-collective@bristol.ac.uk)

An independent advisory panel oversees the running of the study. This group will be consulted in the event of that a potential collaborator lodges an appeal against a decision made by the PMG.

### Sharing data with researchers

#### *General issues*

The data collected within this study are set up as a supported access resource rather than as an open access resource. The following sections describe the information available, the process of accessing and sharing different types of the pseudo-anonymised data and the procedures in place to control output.

#### *Information about the available resources*

The study’s website ([www.bristol.ac.uk/cleft-collective](http://www.bristol.ac.uk/cleft-collective)) provides an overview of the cohort study and information on data availability is detailed within the data dictionary ([www.bristol.ac.uk/cleft-collective/professionals/access/](http://www.bristol.ac.uk/cleft-collective/professionals/access/)). Further information about the study can be found in the Cleft Collective protocol paper (<https://doi.org/10.1136/bmjopen-2024-084737>). All questionnaires sent to families can be found on our website ([www.bristol.ac.uk/cleft-collective/professionals/information/questionnaires/](http://www.bristol.ac.uk/cleft-collective/professionals/information/questionnaires/)). We invite researchers to contact The Cleft

Collective team to ask questions about the data collected. Our website lists each member of the team and their specialism. All questionnaires used in the study are available on request. If you require more detailed information on the variables available please send an email to the PMG ([cleft-collective@bristol.ac.uk](mailto:cleft-collective@bristol.ac.uk)) with the title 'Documentation request' expressing your wish to explore the resource for a potential research proposal. Your email should include enough information for the PMG to identify you as a *bona fide* researcher i.e. your institution name, department, your contact information and a brief explanation of your research interests. In all dealings with the materials and data you must also adhere to the Cleft Collective Confidentiality form and your institution's information security policies.

We wish to encourage data sharing to maximise the use of the resource, and also to minimise overlaps between studies and prevent duplication in data analysis. Accordingly, the PMG may put you in touch with other groups working in the same area.

### *Accessing data and costs*

The cost of accessing the data will be charged on a cost recovery basis and the PMG will expect regular updates on the progress of all projects using the data. Any newly generated data must be returned to the Cleft Collective.

### *Output control*

A member of the PMG must see and ratify the final version of the paper to ensure it is accurate and does not bring the Cleft Collective study into disrepute. Copies of all research papers will be made available on the Cleft Collective website.

### *Restricted data access*

New data collected or generated through a specifically funded exercise *may* be made available only to the PI for a period of exclusive access *only* on request to the PMG. Such exclusivity should not be assumed. Up to a year may be requested. If this is agreed, other interested researchers may access these data only after the agreed embargo period.

### *Access to genotyped data and biological samples*

#### *Genotyped data*

DNA from around 7000 individuals (cases, mothers and fathers) within the Cleft Collective have been genotyped using the Illumina Infinium Global Screen Array v3 [Infinium Global Screening Array-24 Kit | Population-scale genetics \(illumina.com\)](#) which includes approximately 650 000 single nucleotide polymorphisms.

#### *Biological Samples*

The Cleft Collective collect four different sample types:

- Cord blood; collected from the study child at birth (Antenatal strand of the birth cohort)
- Blood; collected from the study child (Birth cohort) during their primary surgery
- Lip and/or palate tissue; discarded tissue collected from the study child (Birth cohort) during their primary surgery
- Saliva samples; collected from the study child (Five year cohort), mother, father, mother's partner, sibling (Birth and Five year cohorts)

Further information on the samples collected can be found on our [website](http://www.bristol.ac.uk/cleft-collective/professionals/information/biological-and-environmental-data) (www.bristol.ac.uk/cleft-collective/professionals/information/biological-and-environmental-data) and on the [UKCRC Tissue Directory](https://directory.biobankinguk.org/Profile/Biobank/GBR-1-133) (https://directory.biobankinguk.org/Profile/Biobank/GBR-1-133).

Ethical approval was obtained for all sample collections (details for specific samples available on request). Consent was obtained for the Cleft Collective resource to store data and samples for the purpose of future cleft research which has been ethically approved. Where possible, samples are stored in multiple aliquots to limit the need for freeze thaw cycles and thereby enable as much analysis as possible. However, the number of aliquots varies between sample types and for some samples only one aliquot was produced (i.e. buffy coats and small tissue samples).

If you are planning a proposal to access biological samples, you are strongly advised to discuss your plans with the Cleft Collective PMG (cleft-collective@bristol.ac.uk) and laboratory team prior to submission. The laboratory team will only be able to provide further information about samples or laboratory procedures, please do not contact them about anything else in relation to the application. Bristol Bioresource Laboratories can be contacted at bbl-info@bristol.ac.uk *-please allow 10 working days for a response.*

To use existing biological samples or to carry out specific genetic analysis on Cleft Collective DNA you need to complete a research proposal describing your proposed research. You must ensure you complete the specific section on the biological samples and genetic analysis including details of the type of sample required, amount needed, and, in the case of DNA, the minimum concentration required.

All samples in the Cleft Collective biorepository are finite, i.e. there are limited stocks available. The PMG are responsible for ensuring that samples are used for projects that maximise the amount of research output obtained from available samples and that any derived data/variables/results are subsequently made available to other researchers.

Proposals that request biological samples must satisfy the following criteria to be approved:

- Scientific strength of the proposal must justify use of Cleft Collective cohort samples, i.e. the data obtained from the samples will be analysed in conjunction with other data held by the Cleft Collective.
- There is not already an approved proposal for the same analysis. Requests to repeat or carry out very similar analysis will not be approved unless there are compelling reasons.
- Evidence must be provided to show the methodology is appropriate given the processing history of the samples, e.g. evidence from published literature or pilot data generated on samples processed in a similar manner. Cleft Collective samples will not be released for method development.
- The assay test platform should have proven quality assurance measures in place.
- The volume requested is reasonable and does not seriously deplete the resource.
- The work proposed is within the scope of the consents obtained for the specific samples.

Final aliquots will be retained for future global discovery projects. The exceptions to this are those sample types where only one aliquot was originally produced. In addition, white blood cell pellets (buffy coats) are reserved for production of DNA since they were specifically collected for this purpose. RNA and cell lines will be produced in-house and managed by the Cleft Collective team. Reasonable requests for DNA will be approved and an appropriate aliquot provided.

Proposals which might significantly deplete stocks of finite biological samples or have insufficient evidence of the validity of an assay will be subjected to independent peer review to ensure they meet the conditions above and that the amount of material required is acceptable. The Cleft Collective PMG will notify you of this requirement where necessary. Please note this means the approval of requests to access Cleft Collective samples will take considerably longer than those for data alone and we will try to complete the process within 6 weeks.

If two or more researchers are requesting the same samples either to carry out the same analysis or a different assay, the proposals will be referred to the independent strategy group who will make the decision regarding which proposal (if any) offers the best use of the material.

If a request is approved samples will be supplied with the following conditions:

- Costs incurred providing samples will be covered by the applicant. These will include costs for retrieval, additional processing necessary for the specific project, shipping costs (both out and return) and linking data. Costs will be provided on a case-by-case basis depending on the work involved and may be subject to VAT.
- All derived data, variables and results generated from biological samples must be returned to The Cleft Collective for inclusion in the data resource and will be made available to other researchers.
- The Cleft Collective reserve the right to specify the laboratory where analysis will be carried out to ensure results obtained are comparable to existing data.
  - o Illumina bead chip analysis on Cleft Collective samples will be run in house (e.g. HumanMethylationEPIC BeadChip).
  - o We have preferred laboratories for outsourcing biochemical and metabolomic analysis, please contact the laboratory to discuss such requests.

Samples, including DNA, are provided under the terms of a Material Transfer Agreement (MTA) or in the case of University of Bristol staff a Service Level Agreement (SLA). Each agreement will include a project specific appendix detailing the samples. Samples will not be released until an agreement has been completed and signed. Please note that for samples that are classified as relevant material under the Human Tissue Act (HTA) an HTA MTA should be used instead. Once your proposal is approved we will provide word files with the specific appendix for you to complete, sign and return to us.

Please note if your project requires grant funding samples will be reserved until the outcome of the first funding application is known or 12 months whichever is shortest. After this time The Cleft Collective cannot guarantee sample availability as they may be released to other projects. Researchers will be expected to discuss plans for further funding applications with the PMG.

Note if a participant withdraws consent whilst samples are being analysed researchers may be asked to destroy relevant samples and provide evidence that this has been completed. Costs of supplying samples cannot be refunded in these cases.

### *Linkage Data*

The Cleft Collective are in the process of setting up systems with external data sources to enhance the data resource. Plans are in place to link to health, administrative and environmental records from external organisations. To obtain this linkage data the Cleft Collective must enter into data usage agreements with the relevant data owners. These data usage agreements will specify the conditions under which the Cleft Collective can share these data with third parties

(e.g. researchers). The data access conditions will differ for each data set we link to and are likely to change over time. Updates regarding different types of linkage data and rules surrounding access will be published within the data access policy in due course.

### Submitting a proposal

If you decide that you would like to access the resource, you should contact the PMG ([cleft-collective@bristol.ac.uk](mailto:cleft-collective@bristol.ac.uk)), describing your proposed research, and complete the project proposal form. The PMG will reply within four weeks to inform you of their response to the request and will provide advice on the next stages. The PMG will also estimate any costs of sharing data if applicable. The PMG also reserves the right to impose additional restrictions as appropriate.

Once your project has been approved you will be issued with a project number. The scientific outline of your proposal will be published on our website, [www.bristol.ac.uk/cleft-collective/professionals/access/projects/](http://www.bristol.ac.uk/cleft-collective/professionals/access/projects/). All applicants and co-authors must be listed on the proposal form and must complete a confidentiality form before the project can commence. Should you request biological samples a Materials Transfer Agreement (MTA) must also be completed. The PMG may request further contracts are in place before the project can commence dependent on the nature of your proposal. If you are undertaking the analysis for your project an encrypted pseudonymised dataset will be supplied to you with a collaborator ID via email once all contracts and ethical approval are in place. The dataset provided is strictly for the use of the outlined project associated with your project number and must not be used for any other purpose.

### PR policy

All press releases on research arising from data collected as part of the Cleft Collective study should be seen and approved by the PMG ([cleft-collective@bristol.ac.uk](mailto:cleft-collective@bristol.ac.uk)). We may decide to issue press releases relating to certain articles. In these cases, we expect the lead author of the paper to agree the content of the press release with the PMG and the University of Bristol's public relations team and to be available to deal with media enquiries and interviews. We may also ask authors to prepare a précis of the content of important papers to include in reports to funders and for inclusion in future applications for supporting the extension and maintenance of the resource.

### Authorship and Publication

The Cleft Collective have adopted the criteria of the International Committee of Medical Journal Editors, (<http://www.icmje.org/>). Authorship should be based on those who have made a substantial contribution to a project and who fulfil the ICMJE criteria.

The ICMJE states 'When a large multi-author group has conducted the work, the group ideally should decide who will be an author before the work is started and confirm who is an author before submitting the manuscript for publication'.

The ICMJE recommends that authorship be based on meeting **all** four of the following criteria:

1. There were substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data;
2. The person was involved in drafting the article or revising it critically for important intellectual content;
3. They gave final approval of the version to be published.

4. They must agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

All those designated as authors should meet all four criteria for authorship, and all who meet the four criteria should be identified as authors. Those who do not meet all four criteria but meet some of the criteria should be acknowledged.

It is important to point out that acquisition of the funding, recruiting participants, collection of data, caring for study patients or general supervision of the research group alone does not constitute authorship.

All full papers must be sent to the PMG for approval ([cleft-collective@bristol.ac.uk](mailto:cleft-collective@bristol.ac.uk)) along with a completed papers checklist prior to journal submission. This includes any research output being placed in the public domain. The PMG expect to process all papers within four weeks of receipt. The PMG will scrutinise all papers to check the confidentiality of data is protected and to ensure that the paper will not bring The Cleft Collective study into disrepute. The PMG reserves the right to require that any paper which could potentially breach the confidentiality of any participant(s) be withheld from submission for publication. The PMG will work with the authors to overcome such breaches. If the researcher submits such publications regardless, the PMG will attempt to prevent publication.

The PMG may also offer advice and feedback to authors if we feel this may be helpful. However, our role is not to provide formal peer review.

The Cleft Collective request that publications include the project number assigned to you at approval. This allows The Cleft Collective to track publications produced from individual research projects. The PMG suggest the inclusion of the following wording:

*“This research has been conducted using The Cleft Collective Resource under Application Number xxx.”*

The PMG request all findings to be returned to The Cleft Collective within 6 months of the publication date or 12 months after the end of the project, whichever comes first. This includes all key derived data-fields, a description of the methods used to generate them, the underlying syntax/code (e.g., SPSS, SAS, R, STATA) used to generate the main results of the paper, a copy of the final manuscript and a short and simple paragraph that summarises your published findings, so that we can keep participants informed about research achievements.

A checklist of requirements for papers together with accompanying notes explaining these requirements and containing appropriate text to use when referring to the study is available ([www.bristol.ac.uk/cleft-collective/professionals/access/](http://www.bristol.ac.uk/cleft-collective/professionals/access/)). A completed checklist must be included with each paper submitted for approval. Researchers should let the PMG know whether a paper is accepted for publication. The authors should provide the PMG with an electronic copy of the final published version.

### Describing The Cleft Collective

Any research output being placed in the public domain that describes The Cleft Collective must be reviewed by the PMG to ensure the description is accurate. This includes publications which have not utilised the Cleft Collective resource. The following wording is a suggested description of the study which may be used by researchers wanting to describe The Cleft Collective in their research outputs.

“The Cleft Collective Study is a large prospective cohort study of children born with cleft lip and/or palate, investigating causes of cleft, the best treatments and the impact of cleft on those affected and

their families.<sup>1</sup> The resource comprises biological samples, speech audio recordings and medical records and parent and child completed questionnaires. The resource is available for clinical and academic communities to access and use to address a range of cleft related research questions. More information on the study and how to access the dataset is available at [www.bristol.ac.uk/cleft-collective/professionals/access/](http://www.bristol.ac.uk/cleft-collective/professionals/access/). Initial funding for The Cleft Collective was provided by the Scar Free Foundation; additional funding was provided by The Underwood Trust and the Vocational Training Charitable Trust (VTCT).”

When describing The Cleft Collective, we request that you reference “The Cleft Collective: protocol for a longitudinal prospective cohort study”, citation and DOI can be found below:

1. Davies, A. J. V., Humphries, K., Lewis, S. J., Ho, K., Sandy, J. R., & Wren, Y. (2024). The Cleft Collective: protocol for a longitudinal prospective cohort study. *BMJ Open*, 14(7), e084737–e084737. <https://doi.org/10.1136/bmjopen-2024-084737>

Please submit any description of The Cleft Collective for approval by the PMG to: [cleft-collective@bristol.ac.uk](mailto:cleft-collective@bristol.ac.uk)

### Acknowledgements

All contributors who do not meet the criteria for authorship should be listed in an acknowledgements section.

People who contributed materially to the paper but whose contributions do not justify authorship may be listed under such headings as “clinical investigators” or “participating investigators”, and their function or contribution should be described, for example, “served as scientific advisors”, “critically reviewed the study proposal”, “recruited participants”, or “collected data”.

The following is a standard acknowledgement section that should be included in all papers:

“This publication involves data derived from independent research funded by The Scar Free Foundation; additional funding was provided by The Underwood Trust and the Vocational Training Charitable Trust (VTCT) (REC approval 13/SW/0064). We are grateful to the families who participated in the study, the UK NHS cleft teams, and The Cleft Collective team, who helped facilitate the study. The views expressed in this publication are those of the author(s) and not necessarily those of The Scar Free Foundation, The Underwood Trust, the Vocational Training Charitable Trust or The Cleft Collective Cohort Study team.”

### Intellectual property

The University of Bristol owns The Cleft Collective resource: any data generated and the biological samples collected. As such any requests to access the data must be made through the PMG. Any data generated through an approved project must be returned to the resource, alongside the initial data provided, to encourage ongoing use by the research community, this includes derived variables.

**The Cleft Collective Team  
July 2024**