The CP Commons Launch Report

INTERNATIONAL CP GENOMICS CONSORTIUM

CP COMMONS

Cerebral Palsy ALLIANCE RESEARCH INSTITUTE

Acknowledgments

The International Cerebral Palsy Genomics Consortium and the CP Commons Data Custodians sincerely thank all families who continue to participate and support genomics research in CP.

Special thanks to:

- Mr Rob White and Prof Nadia Badawi for their commitment to the work of the CP Commons.
- The International Cerebral Palsy Genomics Consortium Governance Council for their ongoing support and commitment to the CP Commons.
- Ms Yana Wilson, Ms Georgina Henry, Ms Juanita Graham, and Dr Sarah McIntyre for developing, coordinating, and writing this report.
- Dr Hayley Smithers-Sheedy and Dr Shona Goldsmith for their editing.
- Ms Georgina Henry for their work on the design and publishing of the report.
- Kablamo Pty Ltd for the development and ongoing support and maintenance of the CP Commons infrastructure.
- Research personnel at the Cerebral Palsy Alliance Research Institute (CPARI), The University of Sydney who manage the CP Commons.
- Special thanks to Cerebral Palsy Alliance Research Foundation, Perpetual Trustees, and the Ian Potter Foundation for their generous support.

Foreword

Cerebral palsies (CP) are a heterogeneous group of disorders that affect body movement, muscle control, coordination or tone, and posture and balance. CP are permanent disorders, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain.

Cerebral palsies have many and diverse causes. Current revolution in our capability to read and understand human DNA sequence also dramatically impacted our understanding of the causes of CP. The aggregate results of a dozen of recently reported CP genomics studies point to a mean 33% contribution of major genetic determinants, i.e. single genes. CP is being redefined as another neurodevelopmental disorder similar to and overlapping with autisms, intellectual disabilities or epilepsies.

Today we are proudly launching CP Commons to provide a unique, custom designed platform for international CP genomic and clinical data sharing and collaboration.

Sharing is caring and we care to drive new discoveries with the CP Commons, to better understand the full spectrum of genomics causes of CP and the underlying complexity of its neurobiology as well as to realign state-of-the-art medical care for individuals with CP with genomics-empowered precision medicine.

This report describes the opportunity that the CP genomics community can leverage through data sharing and collaboration via the CP Commons. Thanks also to the Cerebral Palsy Alliance who have funded the development of the CP Commons – the centerpiece of the ICPGC.

Contact us about contributing to the CP Commons. Together we can drive forward genomics medicine for CP.

Jozef Gecz PhD, FAA, FAHMS, FFSc(RCPA)

Executive Summary

Part I:
Overview on the ICPGC and the CP Commons

- The ICPGC is working to advance our understanding of the genome's role in cerebral palsy through a number of key projects.
- The CP Commons aims to accelerate scientific discovery and understanding of the genome's role in cerebral palsy through collaboration and data sharing.
- The primary function of the CP Commons is a clearinghouse for clinical and genetic data from people with CP and their families.
- Researchers from around the world will be able to seamlessly collaborate with others while maintaining data ownership; access data to test novel hypotheses or supplement cohorts; reuse data to interrogate areas of the genome not yet investigated; and identify similar cases with matching clinical and genetic features.

Part II:
User Guide for the CP Commons

- This section provide researchers and clinicians with instructions for how to use the CP Commons.
- These instructions are accompanied by screen shots of the CP Commons.
- The instructions include how to set-up an account, how to upload data to the CP Commons and how to request access to datasets in the CP Commons.

Index

Part I

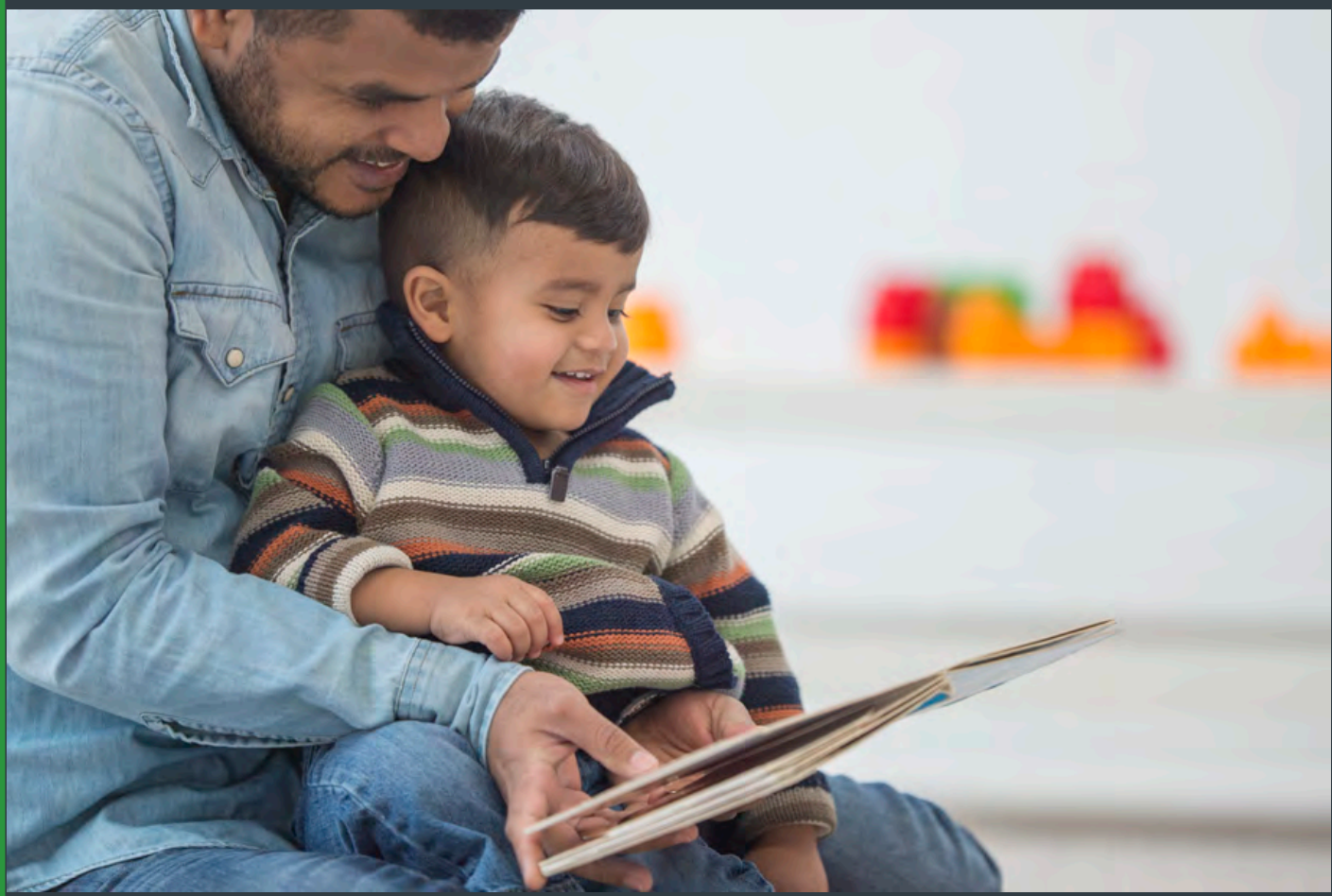
Introduction to the CP Commons and the International CP Genomics Consortium (ICPGC)

Mission Statement of the CP Commons

The mission of the CP Commons is to accelerate scientific discovery in cerebral palsy genomics through collaboration and data sharing.

Impact Statement

Sharing clinical and genomics data will increase research efficiency and expedite knowledge translational efforts, whilst maximising the utility of existing data.



About Cerebral Palsy

Cerebral palsy (CP) is a clinically diverse condition of movement, posture and motor function that is attributed to a non-progressive and permanent disturbance to the fetal or infant brain[1]. Although the hallmark features of CP are motor and posture disorders, many individuals with CP also present with other comorbid conditions - including intellectual impairments, epilepsy, speech and language difficulties, vision and hearing impairments, congenital anomalies, and behavioural disorders[1].

Cerebral Palsy:

1. Is an umbrella term for a group of disorders
2. Is a condition which is permanent but not unchanging
3. Involves a disorder of movement and/or posture and of motor function
4. Is due to a non-progressive interference, lesion or abnormality
5. The interference, lesion or abnormality originates in the immature brain

For more information, please see this Fact Sheet available on our website ICPGC.org.

About Cerebral Palsy Genomics Research

Aetiology
The cause of a condition.

Genome
Is the complete set of genetic instruction in an individual. Each genome contains all the information needed to build, grow and repair our bodies.

Gene
Genes are sections of DNA that contains the instructions for a specific molecule, usually a protein. There are more than 20,000 genes in the human genome.

Sequencing
is a technology that allows researchers and clinicians a means to examine genetic information. Sequencing can mean an entire genome or a specific area of the DNA.

Exome sequencing
focuses on the portion of the genome that encodes for genes only. This is about 2% of the genome.

Genome sequencing
involves sequencing all the genetic information in the entire genome. This is about 3 billion letters long.

CP is caused by a combination of antenatal and perinatal factors such as prematurity, low birth weight, congenital infections, congenital anomalies, and neonatal encephalopathy [2,3]. However, for most individuals with CP, the specific causal pathway to their brain injury remains unclear. More recently, there has been considerable interest in the role of genetic factors in CP aetiology[4].

Over the last twenty years, the research landscape of CP genetics has evolved significantly. Gone are the days where a genetic link to CP would have been criticized, now, dozens of research teams around the world continue to investigate the genome's role in CP. One thing has become clear, there is no single CP gene, but rather hundreds of genes. As of August 2021, in excess of 3000 individuals with CP have undergone exome or genome sequencing world-wide with about one in three (~33%) having a meaningful genetic variant [unpublished, collective work from members of the ICPGC].

Despite the huge gains over these last two decades in CP genomics research, there has been a continued bottleneck for both discovery and translational research in this area. Whilst there have been many individual efforts to examine the role of the genome in CP, these continue to occur in predominantly “siloed” projects, in non-standardised formats, prohibiting easy exchange.

As the first international CP genomics research database, the CP Commons aims to revolutionize the CP research landscape by co-ordinating "siloed" research efforts to accelerate scientific discovery.

International Cerebral Palsy Genomics Research (ICPGC)

INTERNATIONAL CP GENOMICS CONSORTIUM

What is the ICPGC?

The International Cerebral Palsy Genomics Consortium is a multidisciplinary, international, open consortium that is dedicated to accelerating the understanding of the genome's role in CP.

The ICPGC is comprised of ~200 members from academic and medical research centers from around the world, spanning across more than 20 countries. Since its inception, the ICPGC meet every year to discuss breaking research, challenges in the CP genomics space, and approaches that can drive research momentum.

ICPGC Annual Meetings

2017 - Adelaide Australia

The inaugural meeting led to the formation of the consortium that was announced in a manuscript entitled, "*Cerebral palsy and genomics: an international consortium*"[5].

2018 - Zhengzhou, China

During this meeting, the importance of genetic findings not dictating a change in the clinical diagnosis of CP was a common theme. This resulted in the consortium preparing a consensus piece that was published in the Journal of Child Neurology, "*Genetic or Other Causation Should Not Change the Clinical Diagnosis of Cerebral Palsy*"[6].

2019 - Anaheim, United States of America

This meeting was held in conjunction with the American Academy of Cerebral Palsy (AACPDM) and Developmental Medicine and the International Alliance of Academies of Childhood Disability (IAACD). This was our largest meeting to date and the consortium grew significantly in membership.

2021 - Toronto, Canada (virtual)

This virtual meeting is kindly supported by CanChild, the McLaughlin Centre and CP-NET.

Attend our 2022 meeting!

In March 2022, the ICPGC will have a virtual meeting as part of the Australasian Academy of Cerebral Palsy and Developmental Medicine meeting. For more information, click here.

ICPGC Governance Council

The ICPGC Governance Council is a group of clinicians, researchers and advocates who manage the overall scientific direction and goals of the ICPGC, including the objectives of the CP Commons. This council is fulfilled by three-year, rotating positions. Members of the Governance Council (2021):

Prof. Gareth Baynam
(WA, Australia)

Prof. Michael Fahey
(Vic, Australia)

Prof. Jozef Gecz (Chair)
(SA, Australia)

A/Prof Andres Moreno De Luca (PA, USA)

Dr. Charles Steward
(UK / Germany)

Yana Wilson
(NSW, Australia)

Dr. Richard Wintle
(ON, Canada)

Dr. Changlian Zhu
(Sweden / China)

What is the CP Commons

The CP Commons was established by the ICPGC to tackle the siloing, fragmentation, and inaccessibility of data and to promote collaboration and data sharing.

It aims to centralise information and resources necessary to drive the discovery, interpretation, and characterisation of the genome's role in CP.

At its core, the CP Commons serves as a clearinghouse for clinical and genomics data. Authorised researchers from around the world will have the opportunity to collaborate with others; pool data to create larger cohorts; reuse data to interrogate areas of the genome not yet investigated; and identify similar cases with matching clinical and genetic features.

Register for an account here

Find the CP Commons here:
https://cpcommons.icpgc.org/

Engagement with People with CP and their families

Although the CP Commons has been developed for the research community, its innovation is underpinned by the lived experiences of people with CP and their families.

We know that people with CP and their families have identified '*genetics*' as a research priority[7]; however, the success of this research will rely heavily on families' willingness, trust and ongoing support.

Recently, we asked people with CP and their families how they felt about genetic research, including international data sharing[8]. What we found that families are willing to participate in genetic studies of CP, and those willing to participate are also supportive of international data sharing. Importantly, even though many families were supportive of international data sharing, this was contingent upon transparency and ongoing communication.

Although the CP Commons does not directly recruit people with CP and their families, all research supported through data provision from the CP Commons, as well as the overall research findings, will be publicly available on the ICPGC.org website in plain language with a family audience in mind.

If you have CP, or are a family member of someone with CP, and are interested to join the CP Commons Reference Group, please email: info@icpgc.org

Data management in the CP Commons

Genomics data, even when de-identified, remains unique to the individual and thus requires significant security around the storage, distribution, and use of these data. The management of these data are important to maintain the trust of people with CP and their families[8] and is a shared responsibility between the data owners, the data custodians, and authorised users.

The CP Commons data portal authenticates researchers and clinicians at multiple points in the system. The data is divided into tiered levels of access: Open data includes aggregate data, and Controlled data includes comprehensive data at the individual-level. All authorised researchers and clinicians can review the Open data. Whereas, Controlled data can only be access through an approved project. Projects may include data that has been uploaded by a user, data that has been uploaded by another user that they are authorised to view, or a combination of both. Users may have one project, or they can have many.

When the CP Commons releases de-identified data to authorised researchers or clinicians, these data are encrypted. It is the responsibility of the lead researcher and their home research institution)to use a secure computing facility for any local use of the data.

De-identified data
Is data that has undergone the removal or alteration of information that could potentially be used to re-identify an individual.

Phenome
is the set of all phenotypes observed in an individual.

Phenotype
is the set of observable traits or characteristics or clinically relevant abnormalities, including signs, symptoms and abnormal findings of laboratory analyses, imaging studies, physiological examinations, as well as behavioural anomalies.

Human Research Ethics Approval

The CP Commons has approval for the collection, sharing, and processing of de-identified phenome and genome data from people with CP and their families from the University of Sydney's Human Research Ethics Committee [2021/448].

This project adheres to the Australian Code for the Responsible Conduct of Research and the National Statement on Ethical Conduct in Human Research (2007) - Updated July 2018.

The CP Commons only accepts data from ethically approved research studies. During the data submission process, researchers or clinicians are required to attach evidence of their local human research ethics approval.

Data Ownership

Researchers who upload their data will remain the Data Owner of all the original data shared with CP Commons. Cerebral Palsy Alliance Research Institute (CPARI) are the Data Custodians of CP Commons data and will manage the data with the highest level of care. CPARI does not claim ownership of any data deposited into the database.

Collaboration

The CP Commons has been designed to easily facilitate collaborative projects between researchers and clinicians from different sites. All authorised CP Commons members are considered independent researchers. Teams form under each project through the lead researcher listing all other collaborators on the project application form.

If a listed collaborator from another institute intends to upload or download any data to/from the CP Commons, the collaborator and their home institute will need to execute a Data Transfer Agreement (see **Appendix 2: Section 5**) or Data Access Agreement (see **Appendix 3: Section 6**), respectively.

This ensures that data ownership is correctly recognized, and that data management and access is handled according to the informed consent collected from the research participant.

Research Participant Consent

Respect for, and protection of the interests of, research participant data is fundamental to the CP Commons stewardship of human genomics data. All research participants must have given informed consent for their de-identified data to be shared internationally.

The informed consent under which data were originally collected forms the basis for Data Owner's to determine the appropriate level of data access and any additional restrictions applied to the data.

For older projects, where Data Owners did not collect informed consent for international data sharing, we would recommend recontact with participants, where possible. Alternatively, you may wish to seek advice from your local human research ethics committee as to whether it is possible to obtain a consent waiver.

Data Access

All registered members of the CP Commons can view the Open data, whereas access to the Controlled data is only granted following approval of a project request.

Open data

The Open Data can be browsed through a search interface allowing authorised users to search study metadata, phenotype variables, and experimental features. IIt is hoped that these data will stimulate new research questions and help investigators identify datasets suitable for their research. Users will see an overview of all available data, across all projects, that fulfil their search criteria. User will also be able to navigate to see an overview of a specific project and learn more about what was carried out in that study.

Controlled data

Access to individual-level data files is granted following approval of a new Project Application: Request Access (**Appendix 3**). All data access requests undergo a preliminary review by the CP Commons Data Custodians to ensure that the project adheres to the ICPGC and CP Commons goals and mission. The requests are also reviewed to ensure that the stated research purposes are compatible with participant consent. This step affords discrete secondary use restrictions, as stipulated by the research participants, to be considered. Once approved, researchers will be able to download the de-identified, individual-level data file for which they have been approved. All data access approvals are granted for a 12-month period, with the option of extension.

The CP Commons recommends the adoption of the Global Alliance for Genomics and Health **Data Use Ontology (DUO)** to describe the secondary data use restrictions and conditions on the research participant data. This ontology gives explicit data use restrictions, which may be more amenable with participants individual preferences than broad consent for international data use.

Informed consent
A process of getting permission prior to the commencement of a research project, and allow the results of the study to be used in some way, such as for additional research or health care activities or for sharing with others in a publication or database.

Metadata
Is 'data about data', refers to information that accompanies other data and explains their context or provenance.

Ontology
Is a structured vocabulary that has clear categories and classifications. In this case the DUO gives clear options for secondary data use (i.e., not-for-profit research only, or must have ethics approval).

Annotation
Is the process of assigning biological information and/or function to sections of the genome.

Communication

The CP Commons was designed to accelerate discovery, annotation, and characterisation of genes and genomics variants contributing to CP. As such, upon completion of the data access period, researchers will be required to submit a final project summary. These summaries are used to update our resources, as well as to be shared via our ICPGC.org website for families and other researchers to see.

As part of responsible sharing, we also ask that if researchers find any medically actionable, clinically valid, or highly-penetrant pathogenic, or likely pathogenic variants, related to cerebral palsy, or other neurodevelopmental disorders, to report these in the final project summary, where possible. These variants will subsequently be relayed to Data Owners who can refer to their Ethically Defensible Plan and individual participant consent's regarding 'Return of Results'.

To promote collaboration and collegiality, we ask that Data Owners allow us to make their contact information available via the CP Commons to other members, so that it is possible for other members to contact them about their original study from which the data were derived.

Publication

CP Commons members that access data for their research, should acknowledge the CP Commons in all written and oral publications.

ICPGC / CP Commons Project

The CP Commons and ICPGC.org aim to be a centralised knowledgebase for CP genomics. This means that researchers won't have to search multiple databases from around the world to find the resources they require. All information on ICPGC.org is available to the public, whereas the CP Commons is restricted to authorised researchers and clinicians. Here are some ICPGC project highlights to date.

Phenotyping in CP Genomics

Rich phenotypic data are essential for guiding our interpretation of genomic data. However, the capture and analysis of phenotypic data are often fraught with challenges given its semi-subjective nature. For example, phenotypic data are frequently collected in a manner specific to a project, hospital, or organisation. Furthermore, there is an almost infinite spectrum of clinical characteristics that an individual may present with, many of which are descriptive rather than quantitative. Pivotal to the success of the CP Commons, is the ability to harmonise phenotypic data from different projects. Therefore, it helps when researchers adhere to a common set of high-quality data collection standards.

Common Data Elements for Genomics Studies in CP

A practical method used to standardise data collection across multiple sites is through the development of common data elements (CDEs). CDE's enhance data quality and interoperability, so that data can be easily shared. CDEs are comprised of a precisely defined question that is paired with a specific set of responses for that question. This project was carried out by the ICPGC Phenotype Working Group and is completed. The first version of the ICPGC CDEs are available online at ICPGC.org.

Advancing the Human Phenotype Ontology (HPO) for CP

The Human Phenotype Ontology (HPO) is a comprehensive, structured vocabulary that logically organises phenotypic data to enable computational inference and sophisticated analyses. The HPO has become the global standard to record, exchange and analyse phenotypic data from individuals undergoing genomic sequencing. This project, which is a collaboration with the Jackson Laboratory (HPO), the Monarch Initiative, the ICPGC and others, aims to enhance the HPO for CP. This will increase the accuracy with which we can describe phenotypic features specific to everyone with CP and support advanced computational analyses. Hyperlinks to any published work arising from this project will appear on ICPGC.org.

Phenotypic data
Are clinical information regarding an individuals signs and symptoms, as well as relevant demographic data, such as age, ethnicity and sex.

Interoperability
The ability to readily exchange and use different datasets.

Ontology
Is a structured vocabulary that has clear categories and classifications. In this case the HPO describes traits that are different from the expected normal trait (i.e., strabismus is an abnormality of the eye)

Genomics in CP

We are only just starting to explore the genome's role in CP, but importantly, genomics data represents a stable, persistent source of information that can help us better understand the underlying biology of CP. The ability to reanalyse existing genomics data sets simultaneously via the CP Commons, paired with recent scientific breakthroughs in our understanding of the genome, will help to unravel previously unexplored relationships the genome may have with CP. These studies may help us better understand biological pathways involved in CP that may direct future therapies or interventions. Alternatively, our genome may also dictate why some individuals respond better to an intervention compared to others. The CP Commons truly represents an important driver in the progress of biomedical research for CP around the world. Members of the ICPGC are already working on ground-breaking projects that will involve combining and reanalysing genomic data. These projects are made possible via the CP Commons and include:

Aggregation Project

By combining and analysing a large genomic dataset, it will be possible to uncover patterns and relationships that would not otherwise be evident. The enormous value of data sharing has previously been seen in aggregation studies in cancer[10], autism[11, 12] and developmental disorders[9].

Polygenic Risk in CP

The majority of the CP genetic research to date implicates that genetic variant's that are linked to CP are rare and highly pathogenic. However, it is also likely that combinations of genetic variants that are more common in the general population could also confer a high risk for CP. Similar combined genetic effects (polygenic risks) are known to contribute to other neurodevelopment disorders[13, 14] that are frequently co-morbid with CP.

Multi-omics

Sequencing the genome is just one way that we can learn more about how genes and genetic change can contribute to CP. There are many different technologies that can also provide new insights into the genome's role in CP. By combining data obtained from these different 'omics' technologies, it may help researchers identify novel CP genes and modifier genes, discover biomarkers that could be used in early detection, and even lead to novel therapeutic targets.

Polygenic
An observable trait that is produced from the cumulative effects of multiple genes.

Modifier gene
Are genes that affect the phenotypic and/or molecular expression of other genes.

Biomarker
a naturally occurring molecule, gene, or characteristic by which a particular pathological or physiological process, disease, etc. can be identified

Knowledge Translation

It has become increasingly clear that researchers and clinicians must work together to make the most of the large, complex datasets being created through genomics endeavors. This is vital to achieve the ultimate goal of translating genomics data into clinically actionable results to benefit people with CP. With the creation of genomics data becoming increasingly easier and cheaper, the interpretation of these data, and the identification of the relevant environmental factors, will remain the biggest challenges of the years to come – a challenge which these two projects outlined below are tackling head on:

ClinGen

ClinGen was established in 2013 (supported by the National Institute of Health, USA) and is a central resource that defines the clinical relevance of genomics variants for use in precision medicine. In 2021, the ICPGC is forming a ClinGen Gene Curation Expert Panel for CP, with the aim to classify the genes that are linked CP. Hyperlinks to ClinGen and published works from this study will be available via ICPGC.org.

Precision medicine
Medical care that has been optimised for an individual based on their genome finding.

CP Gene

The genetic architecture of CP is complex. There are several types of genetic variations which have been identified to date. These include small mutations that affect the function of a single gene, or structural changes that can impact sections of the genome and dozens of genes. Genetic changes can be inherited or can spontaneously occur in the individual. CP Gene will be an online, dynamic resource created for clinicians and researchers for the ongoing collection, curation, and visualisation of genes linked to CP. This resource will list all genes identified in individuals with CP and the current evidence supporting that gene's role in CP. This resource will be available via ICPGC.org.

Clinical Applications

Identification of a genetic variant potentially opens the door for people living with CP and their families to access improved clinical care. This may be through provision of guidance on appropriate treatments and surveillance as well as targeted and advanced therapeutics, that is "precision medicine"[15, 16]. Additionally, the identification of a genetic cause can also impact reproductive autonomy; improve access to support groups; and contribute more widely to patient and family quality of life and psychosocial outcomes[15].

Matchmaker Exchange

In most circumstances when an individual undergoes genetic sequencing, a single causative genetic variant isn't identified. Often dozens, if not more, variants are found in multiple genes, of which the clinical significance may be uncertain. Figuring out which variant plays an important role in a person's condition can be challenging. The Matchmaker Exchange[17] was established in 2013, and provides the infrastructure to allow clinicians and researchers the opportunity to "match" individuals based on the genetic and phenotypic features, across a federated network of databases. This resource will be available through the CP Commons.


William

Part II
CP Commons User Guide for Researchers and Clinicians

Who can join:

All *bona fide* CP investigators are eligible to apply for an account with the CP Commons. Track record is taken into consideration when reviewing applications. We also ask for groups where multiple individuals require access to submit a single application under their Laboratory Head or Group Leader. Members can be either:

- **Data contributor and data user:** If you are a researcher or clinician and have collected genomics data from people with CP, consider sharing your data on the CP Commons (appropriate consent essential). You will be able to combine your data with other data contributed to the CP Common by other researchers and clinicians.
- **Data user only:** Any researcher or clinician can become an authorised member of the CP Commons. Data users are not required to have their own data to share.

How to join:

To register for an account, researchers and clinicians can complete the registration form on the ICPGC.org website, under Data Access. Click **here** to complete your registration now.

Alternatively, researchers and clinicians can simply email info@icpgc.orgto express their interest. A CP Commons Data Custodians will send a word version of the sign-up form [see **Appendix 1**], which will need to be completed.

Once reviewed and approved, you will be notified via email of the outcome.

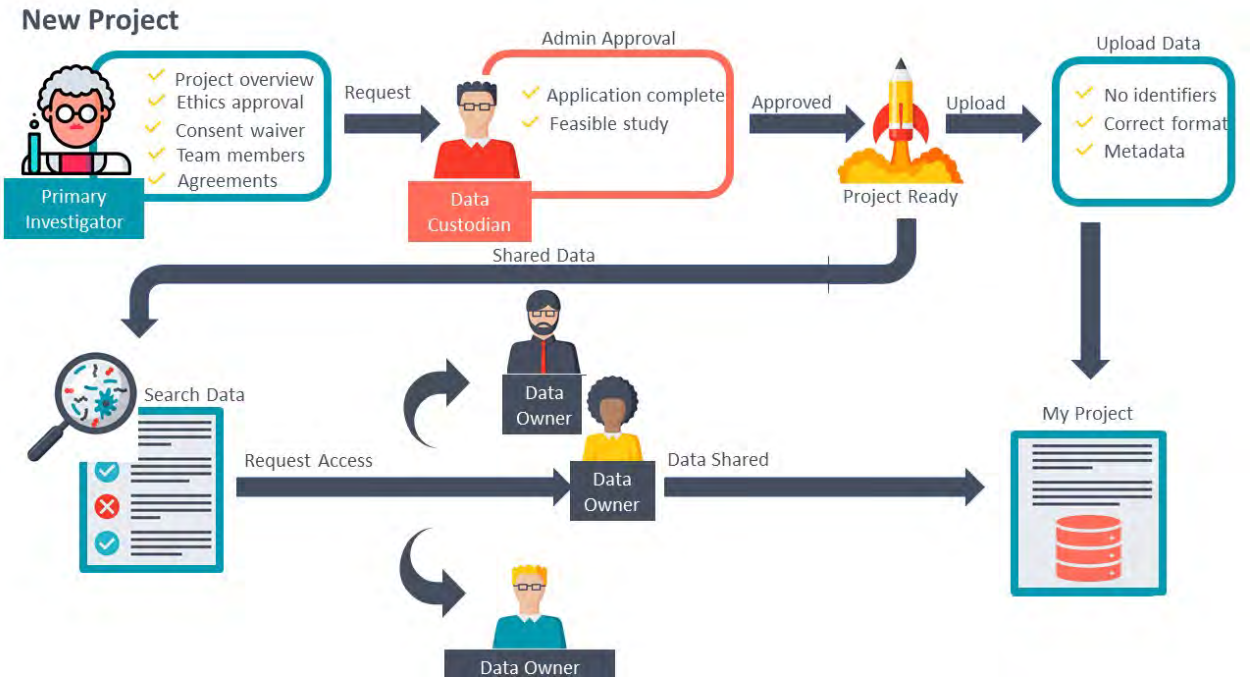
Please note:

- As the CP Commons is a new venture, we welcome feedback and suggestions from all members.
  Please contact Yana Wilson via **ywilson@cerebralpalsy.org.au**.

CP Commons Projects:

Projects are the central function of the CP Commons. You can think of a project as a space where you would upload your own data to share, or the place where you can see any data you have been granted access to. A project also allows for easy collaboration between different research groups – your data can be easily lumped or split within the collaboration project. There is no limit on the number of projects you can have, or how you can use them.

Contact a CP Commons Data Custodian today about how you can use a project space in the CP Commons.



How to create a project

This section includes instructions for setting up a new project in the CP Commons. All new project applications are reviewed by the CP Commons Data Custodians. Upon approval of the project application, researchers will be able to view their project in the My Projects space of the CP Commons.

1. The Primary Investigator (PI) will complete a new project application form (**see Appendix 2 and Appendix 3)**, and email to the CP Commons Data Custodians.
2. All applications at a minimum will require:
   a) A scientific abstract and a brief layman summary of the project, the layman summary is shared on **ICPGC.org,**
   b) A list of all collaborators that require access to the data,
   c) A copy of the ethics approval/exemption letter,
   d) An executed copy of either:
      (i) Data Transfer Agreement **(Appendix 2, Section 5)** for researchers uploading their data to the CP Commons, or
      (ii)Data Access Agreement **(Appendix 3, Section 6)** for researchers requesting access to data in the CP Commons.
3. The application will be reviewed by the CP Commons Data Custodians.
4. Once approved, the CP Commons Data Custodians will create the Project in the CP Commons and notify the PI that their project is now available.
5. All authorized collaborators will now be able to view the Project (Figure 1) in the “My Projects” space.

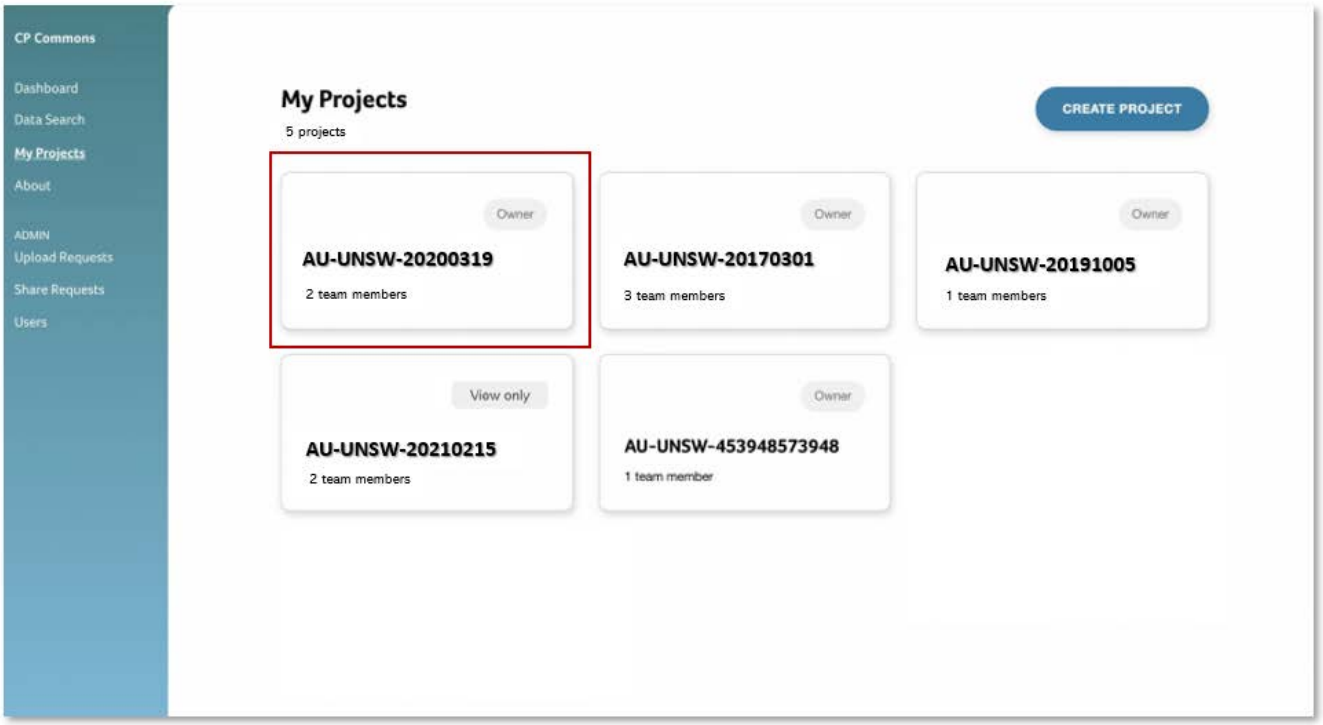


Figure 1: My Projects dashboard in the CP Commons

How to upload data to a project

This section provides instructions for researchers to upload their own genetic and clinical data.

1. Researchers need to ensure that prior to uploading their data:

    a) All Participant ID codes submitted to the CP Commons must have undergone a two-step de-identification process **(Appendix 4).**
    b) All clinical data must be encoded according to the CP Commons Data Dictionary **(Appendix 5).**
    c) The clinical data should contain no personally identifiable information (examples of identifiers are included in **Appendix 6)**. Free text variables should always be checked prior to upload.
    d) All sequencing data must be uploaded in an accepted format **(Appendix 7).**
    e) All sequencing data must be accompanied by an experimental metadata file **(Appendix 8).**

2. When a researcher is ready to upload their data, navigate to the corresponding Project space, and select ‘Upload Data’ (Figure 2).

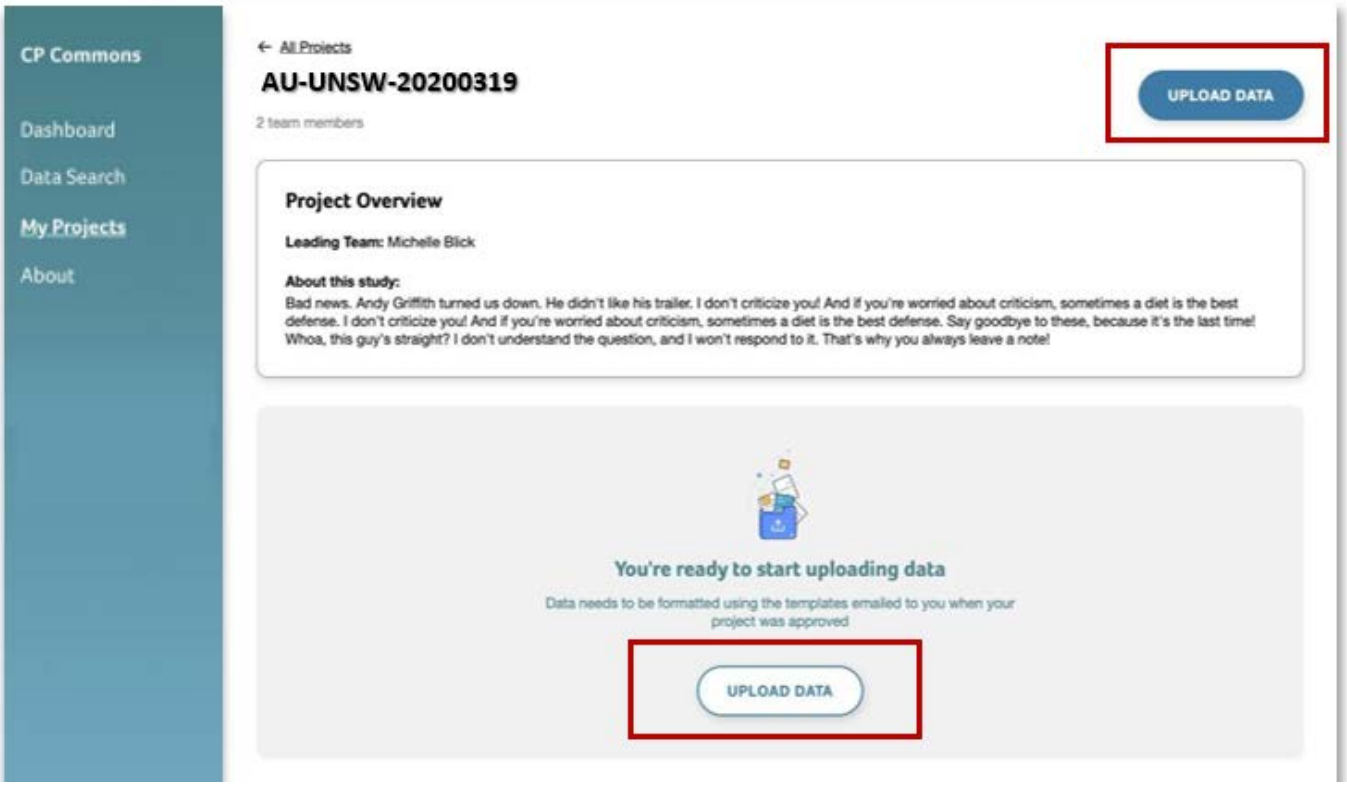


Figure 2: Project overview and data upload.

3. On the next screen, the team member can select which data they are uploading (genomics, clinical or both) and simply drag and drop the data into the corresponding bucket. Users can also browse files if they prefer. The maximum volume of data that can be uploaded at any time is 6TB (Figure 3). When all files have been uploaded, you may submit your data for review. Please note, you do not need to submit all data at once.

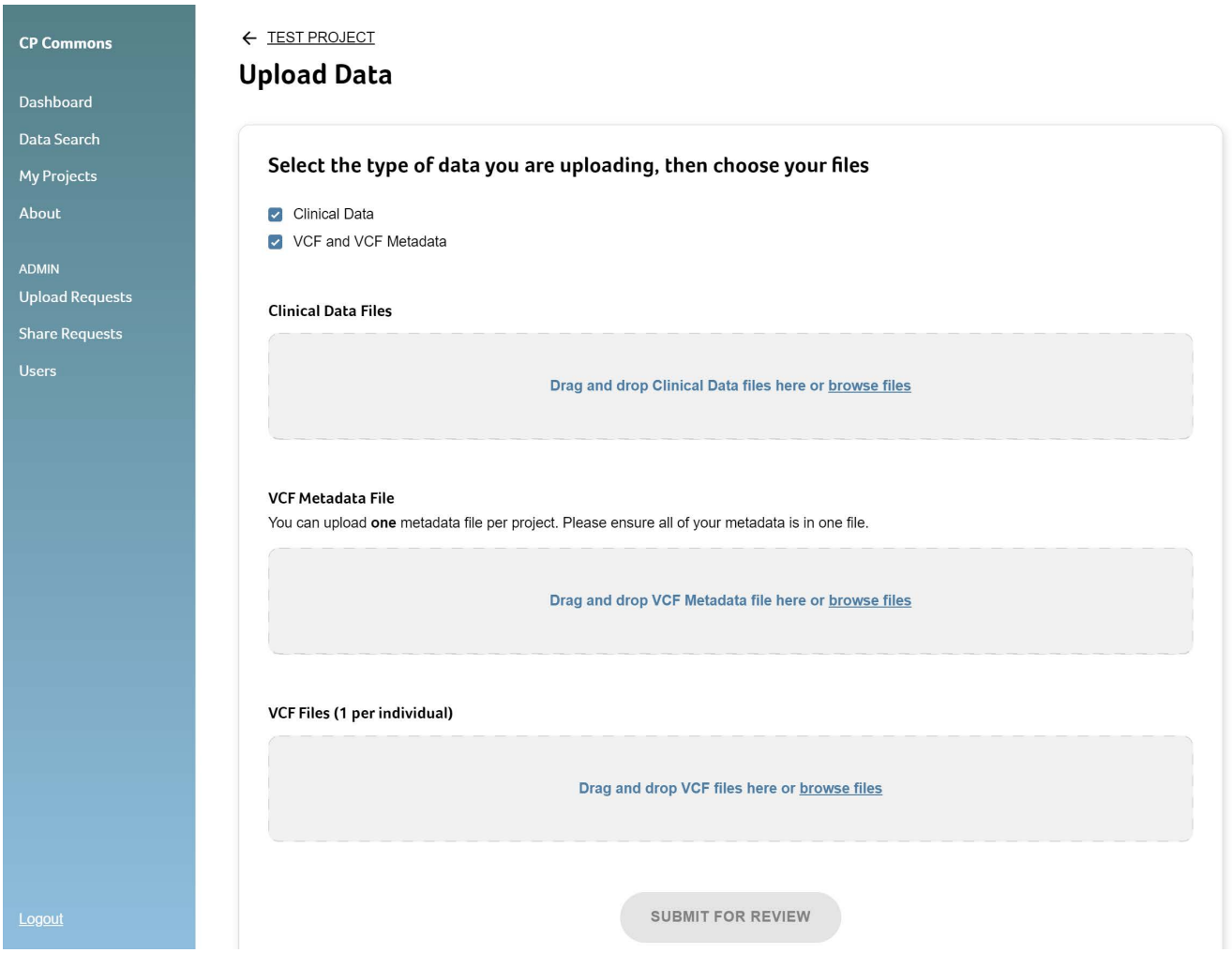


Figure 3: Data upload file drag and drop

4. The CP Commons Data Custodians will review the data within 48hrs. This review is to confirm completeness of your data set and to verify that there is no personally identifiable information included in the clinical data. If any personally identifiable information is identified, we will remove these data from the dataset, and request that you upload the missing data again, without the personally identifiable information.

5. Once approved, the data will appear in the Project workspace under ‘My data’ (see Figure 4) and can also be found by other researchers in ‘Data Search’.

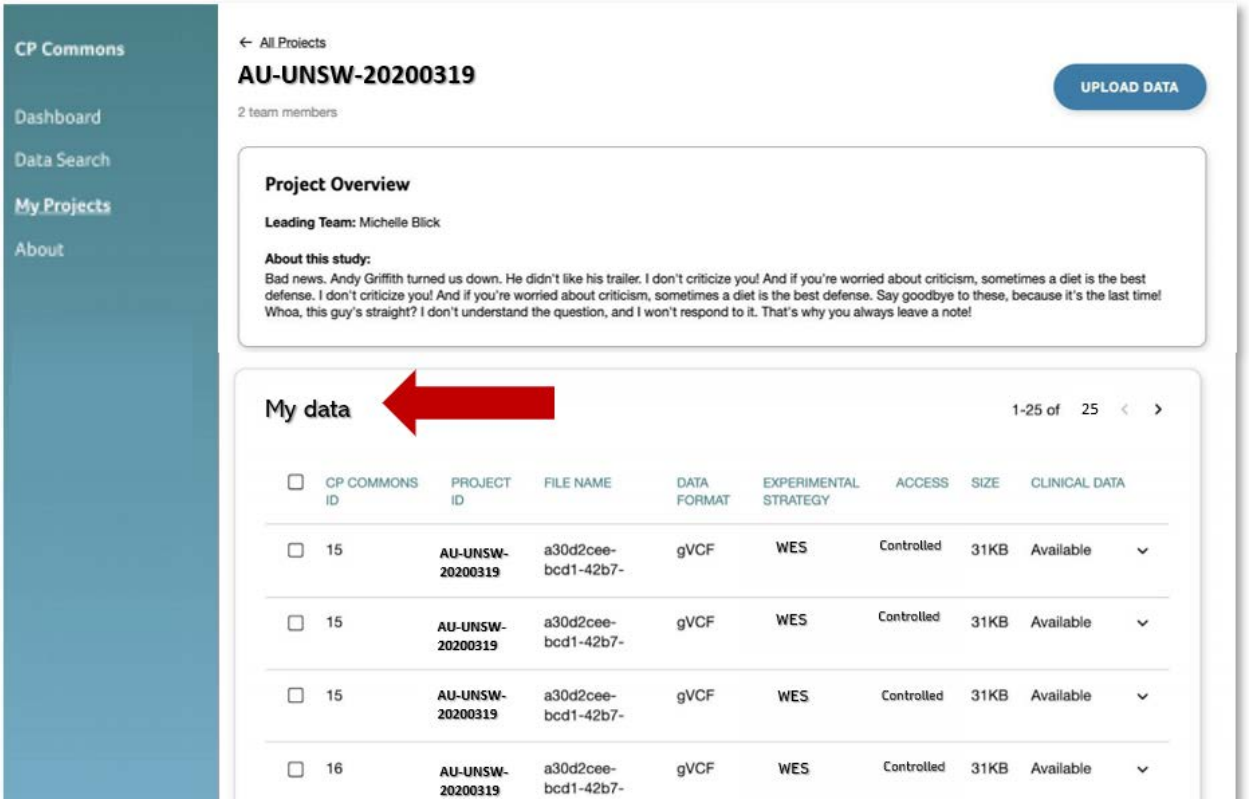


Figure 4: My Project overview with ‘My Data’

How to find other data for my project

This section provides instructions for how researchers can find data in the CP Commons and how to request access to them.

1. To search available datasets in the CP Commons, researchers can navigate to the ‘Data Search’ page (Figure 5). From here researchers can search and create a cohort from the datasets available using the search filters (Figure 5).

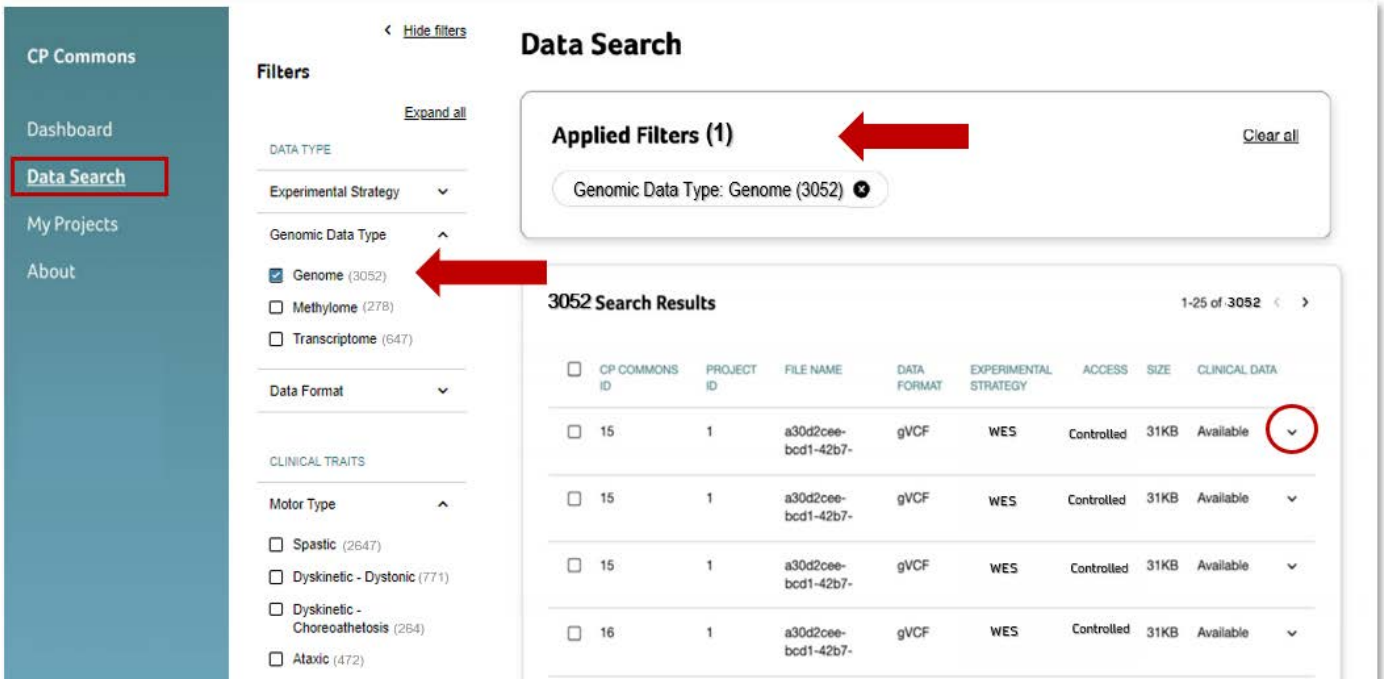


Figure 5: Data Search

2. Each row of data corresponds to a single sequencing data asset from an individual. Some individuals may have multiple sequencing data assets. To view more information about each file simply click the down arrow at the end of the row (Figure 5).

3. The researcher can then select the data or select all data that they wish to access (Figure 6).

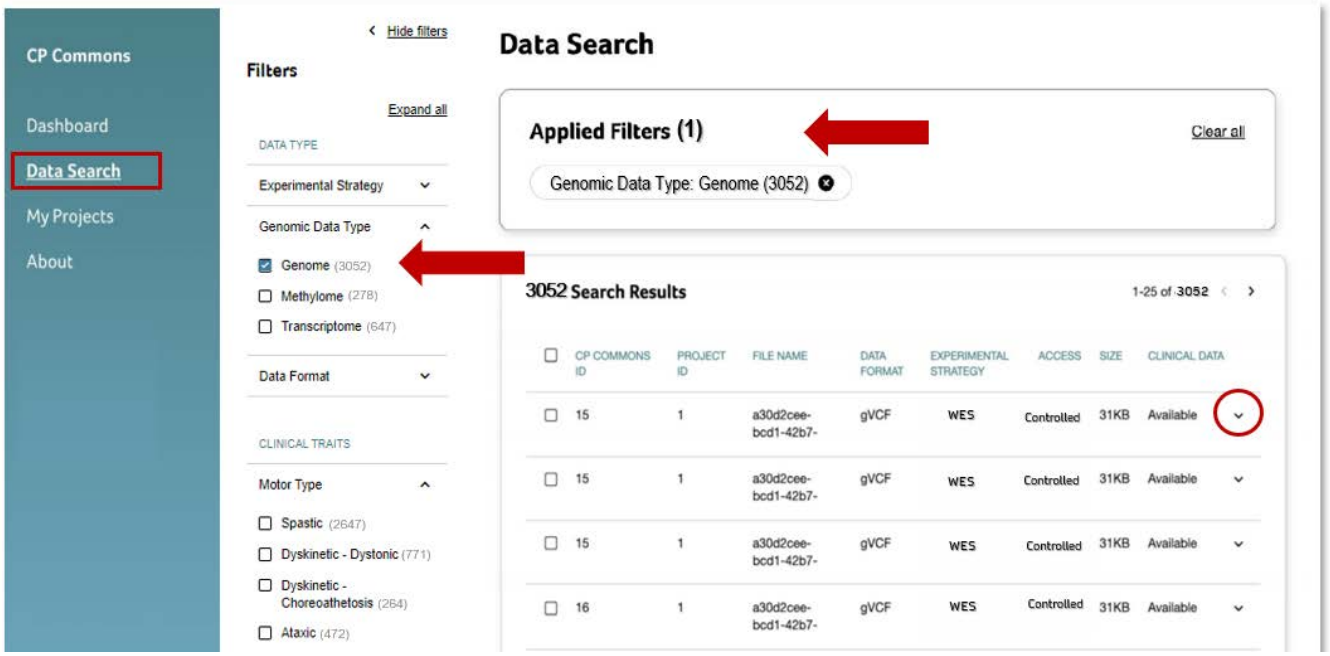


Figure 6: Data search filters

4. The researcher will need to assign these data to one of their Projects (Figure 7)

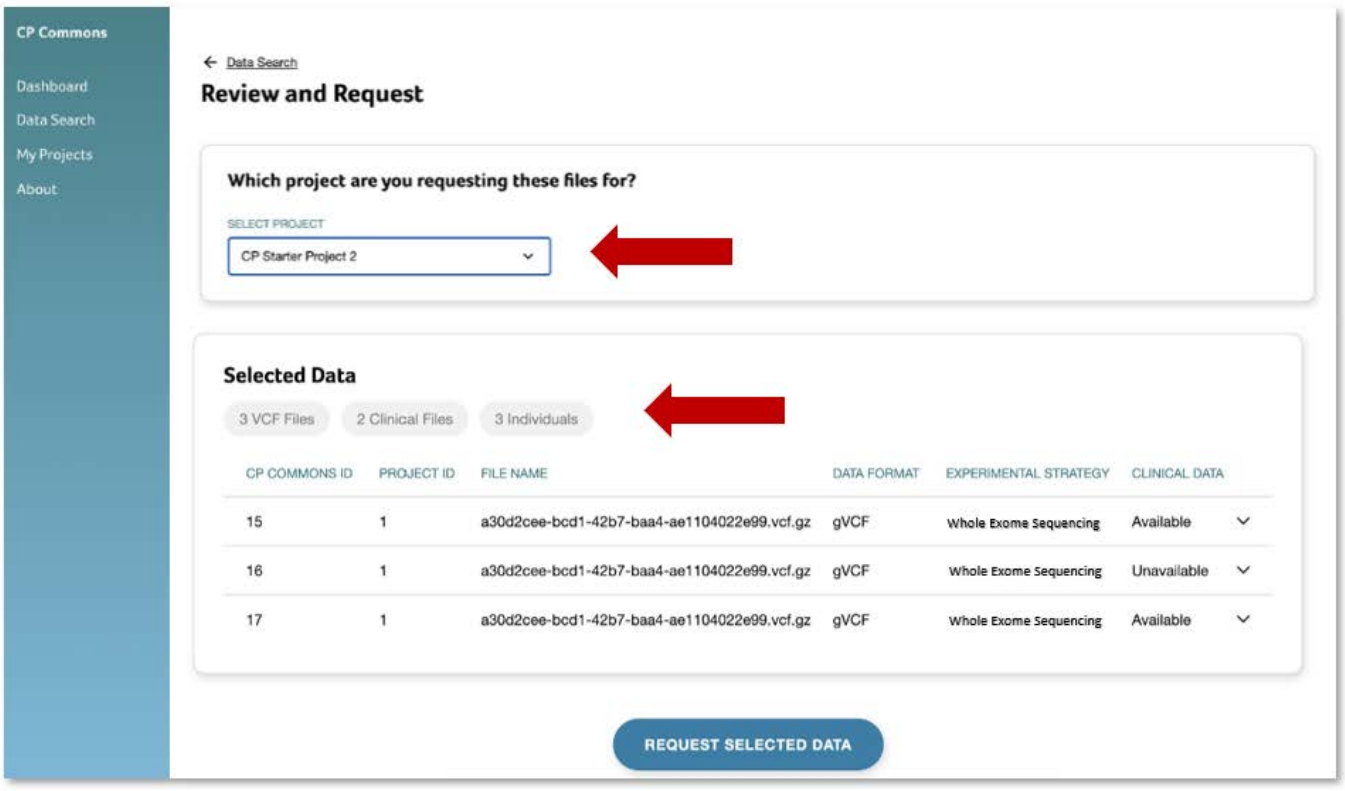


Figure 7: Assign data request to a Project

5. Once the Request has been submitted it will undergo a review process:
    a. Requests will initially be reviewed by the CP Commons Data Custodians to confirm the validity of the data access request.
    b. Once approved by the Data Custodians, the request is sent on to the relevant Data Owners. The purpose of this review is to ensure that any discrete or specific consent requirements associated with the data are met.
6. After the request is submitted, the researcher will be able to view the data in their Project space under the 'Data Shared with Me' space. All pending-approval data will not be able to be selected, the checkbox will be replaced with a pending icon (Figure 8). Once the Data Owners have approved the request, the data will be able to be selected (to move to download or computer).

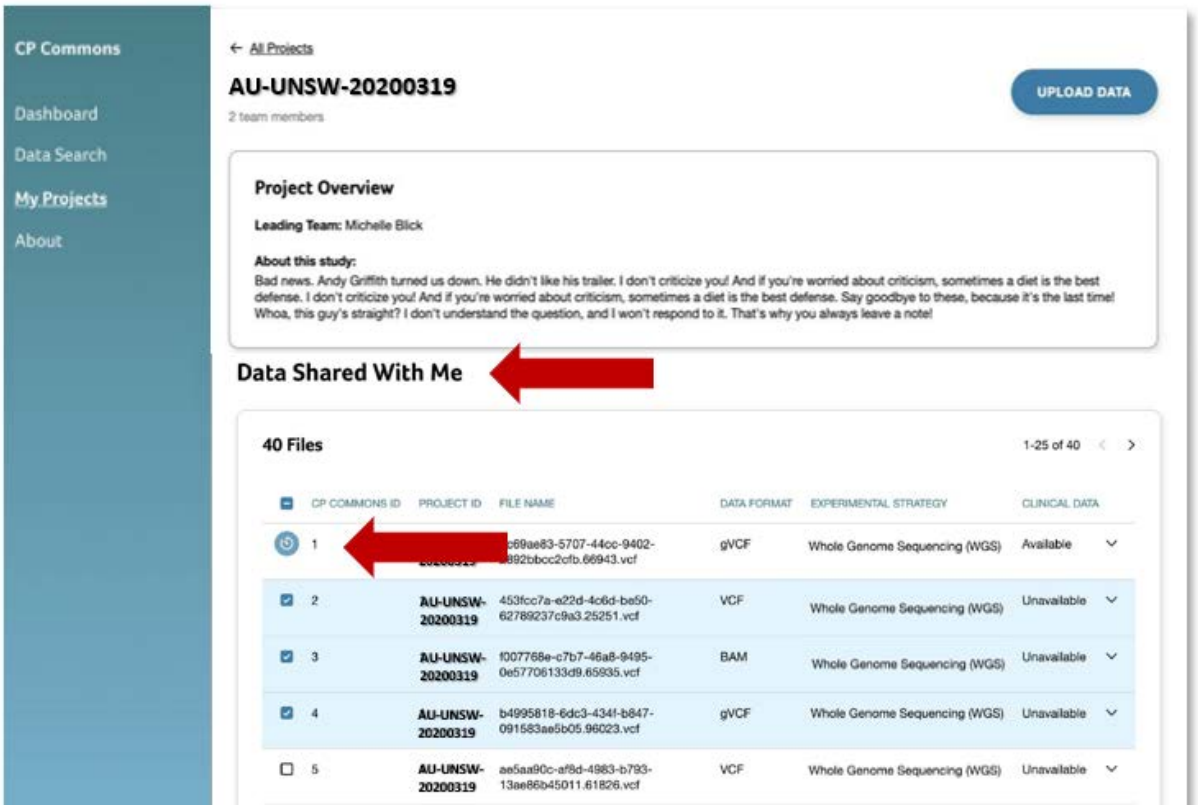


Figure 8: My project overview with "data shared with me"

7. You will have access to the data for 12 months.

Appendices

Appendix 1 - CP Commons Sign-up Form

This form can be completed by individual researchers or clinicians, or by a Laboratory Head / Team Leader if multiple members of their team require access to the CP Commons.

The Laboratory Head, Group Leader, Theme Lead etc will be taking responsibility for all researchers (e.g. PhD students, postdocs or junior staff) working under you that you list in this registration form.

Please provide as much information as possible as this will help to expedite the review process to verify legitimate users.

We will contact you via the email you provide in this form regarding the outcome of your application.

If you have any problems or questions, please contact info@icpgc.org with the subject "CP Commons Registration".

User information

Title		Job Title	
First Name		Last Name	
Institute email address			
Short summary of research interests and professional expertise			

Institution information

Name of institute	
Institute website	
Institute address	

International Cerebral Palsy Genomics Consortium

Are you a member?	☐ Yes ☐ No
If not, do you know an ICPGC member that can vouch for you?	☐ Yes ☐ No
If yes, please provide their name and email:	Name:_____ Email:_____

Note: The information you provide in this form will be stored in Australia. By submitting this application, you consent to submit your personal information to a 'data processor' located in Australia.

Please list any additional team members that will require access to the CP Commons

Name	Email	Read and accept the Terms of Use
		☐ Yes
		☐ Yes
		☐ Yes

Additional Information

Please provide three recent publications PubMed IDs or DOIs	
ORCID ID	
Research Gate Profile	
LinkedIn Profile	

CP Commons Terms of Use

CP Commons is a shared resource developed for the advancement of studies into Cerebral Palsy. The information on this site is intended for authorised users only.

By accessing this site, you agree to

- Not attempt to attribute, link, connect, or associate any Data with any natural person.
- Not attempt to gain access to areas of the CP Commons or the data, which they are not authorised to access.
- Not to share data with unauthorised users outside of the CP Commons system

CP Commons is for research purposes only, and is not intended to diagnose, treat, cure, mitigate, or prevent any disease or condition.

All User platform requests and actions are securely logged for audit purposes. The Data will be stored in perpetuity, even if a User withdraws from the CP Commons.

I have read and accept the Terms of Use	☐ Yes ☐ No

Note: The information you provide in this form will be stored in Australia. By submitting this application, you consent to submit your personal information to a 'data processor' located in Australia.

NEW PROJECT APPLICATION FORM: UPLOAD DATA

Preamble

This application form must be signed and completed by the Primary Investigator and the legal entity with which they are affiliated ("Data Owner") prior to being granted permission to uploading data to the CP Commons. All sections are integral components of this application.

For projects with multiple institutes submitting data: each institute will be required to submit Section 3-6 of this form to ensure proper provenance, attribution, and data ownership.

Your Research Project (as defined below) will be checked for conformity with the goals and policies of the International Cerebral Palsy Genomics Consortium (ICPGC) including, but not limited to, policies concerning the purpose and relevance of the research, quality of the participant data, and the protection of the participants' data. The terms you accept in this application, form an agreement between the Data Owner and the Cerebral Palsy Alliance ("CPA") which is the legal entity that administrates the CP Commons on behalf of ICPGC member institutions. CPA includes its employees, officers, directors, contractors, subcontractors, and agents (including the CP Commons Data Custodians, as defined immediately below).

Once the CP Commons approves your application, you may upload the Data outlined in this Project Application to the approved Project space.

1. Project Overview

This form should be filled out by the Lead Investigator (PI).

Project Contact:

Full Name	
Project Title	

Scientific Abstract:

This section should describe the **background**, **objectives**, and **methodology** of your research project in no more than 500 words.

Project Lay Summary:

Lay summaries are posted onto the ICPGC.org website and used in miscellaneous publications. These summaries are for the public, including people with CP and their families, please describe your project how you would to a friend that is not an expert. Scientific terminology such as “germline”, “non-coding regions”, “heterogenic” should be described or defined in lay terms. In addition to explaining the **background and objectives** of your research project, please also describe any **results or findings**.

Publications that include the use of these data:

Consent to the use of your Project Information:

☐	Primary Investigators agree that the application information may be included in a registry containing the Primary Investigators name, institution, the scientific abstract, any publications arising from the data, and PI contact information. This information will be made available to other members of the CP Commons.
☐	Primary Investigators agree that the application information be included in a registry containing the PI’s name, institution, publications, and the lay summary of the scientific abstracts. This registry will be publicly available at ICPGC.org, and on occasion in any miscellaneous reports or publications that we may prepare about the CP Commons.

2. Collaborators

Please include the names of all **investigators, collaborators, research staff (including post-docs) and students (including graduate students)**, who will have access to the Data in the Project in the CP Commons. Please ensure that a **valid institutional email address** is included for each collaborator and **declare whether they will be uploading data** or not.

Full Name	**Institutional email address**	**Uploading data?** (Yes/No)

***Any co-investigators, collaborators, or students at other institutes that are uploading data will be required to complete and submit Section 3-6 of the Project Application – Upload Data form. This ensures data provenance is correctly attributed and Data Access Requests are submitted to the appropriate Data Owner. All investigators on the project will have immediate access to the data in the Project space under, "Data Shared with Me".*

3. Ethics

The CP Commons only collects data from ethically approved research studies. The CP Commons and the ICPGC are not responsible for the ethics approval/monitoring of individual Research Projects and bear no responsibility for the applicant's failure to comply with local/national ethical requirements.

Respect for, and protection of the interests of, research participant data is fundamental to the CP Commons stewardship of human genomic data. All Research Participants must have given informed consent for their de-identified data to be shared internationally. Consent may be broad or specific. If you have not consented for international data sharing, we recommend re-contact with Participants. Alternatively, you may wish to seek advice from your local human research ethics committee as to whether it is possible to obtain a consent waiver.

☐	You represent and warrant that your Project has been approved* by a research ethics committee formally designated to approve and/or monitor research involving humans. Approval number: _____ **Please attach a copy of your approval letter*
☐ ☐	You represent and warrant that you have sought informed consent for the purposes of international data sharing from your Research Participants. If not, You represent and warrant that your Project has been approved* for consent waiver by a research ethics committee formally designated to approve and/or monitor research involving humans. **Please attach a copy of your approval letter*

4. Institution Representative

An institutional representative is a qualified representative of a legal entity who has the **administrative power to legally commit that entity to the terms and conditions of the data access agreement**. Examples of institutional representatives include, but are not limited to: a Vice-President Research, a Research Director, or a Contracts Officer for the entity.

The Institutional Representative will have final approval of this application before being reviewed by the CP Commons Data Custodians.

Institutional Representative Information

Title	
First Name	
Middle Name	
Last Name	
Position Title	
Primary Affiliation	
Institutional Email Address	
Institution/Company Mailing Address	
City and State / Province / Territory	
Postcode / Zip Code	

5. Data Transfer Agreement (CP Commons)

Agreement Details

CPA	Cerebral Palsy Alliance ABN 45 000 062 288 187 Allambie Road, Allambie Heights, NSW, 2100, Australia
Data Owner	[Insert name of Primary Investigator's Institution] ABN (as applicable) [Insert ABN] Of [Insert address]
Primary Investigator	[Insert name of Primary Investigator] [Insert Address] [Insert Email]
Commencement Date	[Insert]
Data	*Insert description of the data including number of individuals, types of data to be shared*
Purpose	In line with the Human Research Ethics Committee approval for the CP Commons (2021/448)

Terms and Conditions

This Agreement governs the terms of transfer of the Data (as defined in the Agreement Details above) from the Data Owner to CPA. In signing this Agreement, the Parties agree to be bound by the terms and conditions of access set out in this Agreement.

1. Definitions

Agreement means this Agreement, including any Schedules.

Authorised User means a person who is approved to access the Data in the CP Commons and is party to a Data Access Agreement.

CP Commons means the cloud-based database managed by CPA that holds de-identified genomic and clinical data collections of people with cerebral palsy (Research Participants) and their families

Data means the data so described in the Agreement Details.

Data Access Agreement means a data access agreement in the approved form from time to time between CPA and a third party to grant access to certain data in the CP Commons.

Data Owner means the Institution so described in the Agreement Details, and at which the Primary Investigator is employed, affiliated or enrolled.

ICPGC means the International Cerebral Palsy Genomics Consortium.

Intellectual Property means all present and future industrial and intellectual property rights, including without limitation:

(1) inventions, patents, copyright, trade business, company or domain names, rights in relation to circuit layouts, plant breeders rights, registered designs, registered and unregistered trademarks, know how, trade secrets and the right to have confidential information kept confidential, and any and all other rights to intellectual property which may subsist anywhere in the world; and

(2) any application for or right to apply for registration of any of those rights.

Primary Investigator means the person so described in the Agreement Details.

Purpose means the purpose so described in the Agreement Details.

Research Participant means an individual having contributed their de-identified data to a research project that has been uploaded to the CP Commons.

2. Interpretation

In this Agreement, the following rules apply:

2.1 headings are for convenience only and do not affect interpretation;

2.2 another grammatical form of a defined word or expression has a corresponding meaning;

2.3 a reference to time is to time in Sydney, Australia;

2.4 a reference to legislation (including subordinate legislation) is to that legislation as amended, re-enacted or replaced, and includes any subordinate legislation issued under it;

2.5 a reference to a person includes a natural person, partnership, body corporate, association, governmental or local authority or agency or otherentity; and

2.6 the meaning of general words is not limited by specific examples introduced by *including, for example* or similar expressions.

3. Ownership and Use of Data

3.1 CPA acknowledges that the Data is the property of the Data Owner.

3.2 The Data Owner grants CPA a non-exclusive, perpetual, transferable, fee-free licence to use the Data for the Purpose.

3.3 CPA will not use the Data other than for the Purpose unless required by law, or the Discloser consents in writing to that use.

3.4 The Data Owner agrees that CPA can provide access to the Data to Authorised Users for research purposes.

3.5 The Data Owner certifies to the best of its knowledge and belief that the Data is accurate.

3.6 The Data Owner acknowledges that the Data was collected in a manner consistent with all applicable laws and regulations, as well as institutional policies. The Data Owner further acknowledges that the Data was collected pursuant to an informed consent that is consistent with this data submission to CPA under this Agreement.

4. Confidentiality

4.1 CPA agrees to treat as confidential the Data and to use reasonable endeavours to keep the Data secure and protected from any use, disclosure or access which is inconsistent with this Agreement.

4.2 CPA undertakes not to use or attempt to use the Data to compromise or otherwise infringe the confidentiality of information on Research Participants.

4.3 CPA agrees not to transfer or disclose the Data, in whole or part to anyone that is not an Authorised User, except as otherwise required by law.

4.4 CPA will notify the Data Owner if it becomes aware of or suspects any unauthorised use, disclosure or access of the Data.

5. Reports and Outcomes

5.1 CPA will keep and maintain accurate records in connection with third party access to the Data.

5.2 If requested, CPA will produce a report for the Data Owner to verify that CPA is complying with the terms of this Agreement.

5.3 CPA will acknowledge the Data Owner and the Primary Investigator in any publication about the CP Commons and the Data.

6. Intellectual Property

6.1 The Data is the property of the Data Owner. In addition, all Intellectual Property subsisting in or in relation to the Data is the property of the Data Owner.

6.2 CPA agrees not to make intellectual property claims on the Data and not to use intellectual property protection in a way that would prevent or block access to, or use of, any element of the Data, or conclusion drawn directly from the Data.

7. Term and Termination

7.1 This Agreement commences on the Commencement Date and will continue for so long as CPA continues to hold the Data.

7.2 Either Party may terminate this Agreement if the other party is in breach of any of its obligations under this Agreement and, if that breach is capable of remedy, does not rectify that breach within 30 days after receipt of a notice to remedy that breach.

7.3 The Parties will attempt in good faith to resolve through negotiation any disputes arising out of or relating to this Agreement.

7.4 In the event that CPA no longer holds any Data, this Agreement will automatically terminate.

8. General

8.1 Each Party must comply with its obligations under all applicable laws in relation to the collection, storage, use and disclosure of any personal information or health information (as defined in any applicable Privacy Laws) which it obtains during the conduct of the Purpose or to which it becomes privy as a result of this Agreement.

8.2 Neither party to this Agreement will be liable for any indirect, incidental, consequential, punitive, or exemplary damages arising out of this Agreement, including damages for loss of profits, goodwill, use, or data from or relating to any breach of this Agreement, even if a party has been advised of the possibility of such damages.

8.3 This Agreement may be executed in any number of counterparts. All counterparts will constitute one instrument. The parties agree that email signatures will be accepted as originals.

8.4 A party must not assign or otherwise transfer any or all of its rights arising out of this Agreement without the written consent of the other party.

8.5 This Agreement constitutes the entire agreement between the parties with respect to the Recipients' access to and use of the Data.

8.6 These terms are governed by the law of the State of New South Wales, Australia. The parties submit to the exclusive jurisdiction of its courts.

6. Execution

Executed as an Agreement

Executed by an authorised person of the **Cerebral Palsy Alliance** ABN in the presence of:

…….. ……..

Signature of authorised person Signature of witness

…….. ……..

Name of authorised person Name of witness

Date: ……..

Executed by an authorised person of the [Insert organisation name] ABN (as applicable) [Insert AB in the presence of:

…….. ……..

Signature of authorised person Signature of witness

…….. ……..

Name of authorised person Name of witness

Date: ……..

Read and acknowledged by :

……..

Signature of Primary Investigator

……..

Name of Primary Investigator

Date: ……..

NEW PROJECT APPLICATION FORM: DATA ACCESS

Preamble

This application form must be signed and completed by you and the legal entity with which you ("Primary Investigator") are affiliated ("Primary Investigator's Institute") prior to being granted access to CP Commons Controlled Data (the "Controlled Data" as further defined in the Data Access Agreement of this application). All sections are integral components of this application.

Your Research Project (as defined below) will be checked for conformity with the goals and policies of the International Cerebral Palsy Genomics Consortium (ICPGC) including, but not limited to, policies concerning the purpose and relevance of the research, the protection of the participants and the security of the participants' data. The terms you accept in this application, form an agreement between you and the Cerebral Palsy Alliance ("CPA") which is the legal entity that administrates the CP Commons Controlled Data on behalf of ICPGC member institutions. CPA includes its employees, officers, directors, contractors, subcontractors and agents (including the CP Commons Data Custodians, as defined immediately below).

If the CP Commons Data Custodians approves your application, access to the Controlled Data will be granted for a one-year period (starting from the date you are approved for access). An Annual Renewal Application must be completed by you in order to access/use controlled data beyond that one-year time period and thereafter as applicable.

If your application is approved, you agree that your application information will be included in a registry containing the applicants' names, institutions and lay summaries of the scientific abstracts of all applicants having been granted access to Controlled Data. This registry will be available for public access at ICPGC.org.

1. Project Overview

This form should be filled out by the Project Lead Investigator.

Project Contact

Full Name	
Project Title	

Scientific Abstract:
This section should describe the **background**, **objectives**, and **methodology** of your research project in no more than 500 words. Please include 1-2 sentences that clearly explain how the **CP Commons Controlled Data** will be used in the research Project.

**** Please note: If you are planning on combining CP Commons Controlled Data with other datasets, please note that, as per Clause 5 of the Data Access Agreement, you agree not to link or combine the Controlled Data to other data available in a way that could re-identify the Research Participants. Please ensure to clarify how the methods you intend to use to combine datasets will not lead to the re-identification of the Research Participants.*

Project Lay Summary:
Lay summaries from CP Commons Data Custodian Approved Projects are posted onto the ICPGC.org website. These summaries are for the public, including Research Participants and their families, please describe your project how you would to a friend that is not an expert. Scientific terminology such as “germline”, “non-coding regions”, “heterogenic” should be described or defined in lay terms. In additional to explaining the **background and objectives** of your research project, please ensure you clearly explain **how the CP Commons Controlled Data will be used**.

Please provide a Lay Summary for your Research Project: *****

Consent to the use of your Project Information:

☐	Primary Investigators agree that the application information be included in a registry containing the PI’s name, institution, publications, and lay summary of the scientific abstracts and study findings. This registry will be publicly available at ICPGC.org, and these information may also be including in other reports or publications that we may prepare about the CP Commons.

2. Collaborators

Please include the names of all **investigators, collaborators, research staff (including post-docs) and students (including graduate students)**, who will have access to the Controlled Data in order to work on the project (see "Research Project"). Please ensure that a **valid institutional email address and job title/function** are included for each collaborator.

Full Name	**Institutional email address**	

****Please Note: Any co-investigators, collaborators or students at other institutions that will be included in the study AND will be requiring a copy of these data (i.e. to download or move to compute) will need to submit an additional Project Application for Data Access. If the collaborators will not host a local copy or will only be working with derivatives of the Data (i.e. summary statistics or aggregated results for visualisations) they can be added to this list.*

3. Ethics

The CP Commons is aware that some countries/regions do not require ethics approval for use of coded data (i.e. use of the Controlled Data). Depending upon the nature of your Research Project, it is possible, however, that such approval is needed in your country. If you are uncertain as to whether your Research Project needs ethics approval to use the CP Commons Controlled Data, we suggest you contact your local human research ethics committee (or country equivalent) to clarify the matter.
Please choose one of the following options:

☐	You represent and warrant that your country/region does not require your Research Project to undergo ethics review.
☐	Your country/region requires your Research Project to undergo ethics review, and therefore, this Research Project has been approved by a research ethics committee formally designated to approve and/or monitor research involving humans. *Please attach a copy of your human research ethical approval letter to this application.*

**** Please note: The CP Commons and the ICPGC are not responsible for the ethics approval/monitoring of individual Research Projects and bear no responsibility for the applicant's failure to comply with local/national ethical requirements.*

**** Please note: That access to some data is contingent upon human research ethics approval (explicitly implied in the informed consent process).*

4. Institution Representative

An institutional representative is a qualified representative of a legal entity who has the **administrative power to legally commit that entity to the terms and conditions of the data access agreement**.
Examples of institutional representatives include, but are not limited to: a Vice-President Research, a Research Director, or a Contracts Officer for the entity. *Please fill out the following details for the authorized institutional representative. Please ensure that a full postal address and a valid institutional email are included. The Institutional Representative will have final approval of this application before being reviewed by the CP Commons Data Custodians.*

Institutional Representative Information

Title	
First Name	
Middle Name	
Last Name	
Position Title	
Primary Affiliation	
Institutional Email Address	
Institution/Company Mailing Address	
City and State / Province / Territory	
Postcode / Zip Code	

5. Information Technology Agreements

To avoid inadvertent disclosure of CP Commons Controlled Data to unauthorized individuals, the CP Commons Data Custodians ask that you observe basic information security practices. If you make local copies of CP Commons Controlled Data, you must minimize the risk that this information might be used and/or disclosed to persons who have not been approved for access to CP Commons Controlled Data.

At a minimum, you agree to the following:

- **Physical security** – CP Commons Controlled Data will be maintained on physically secure computer systems, such as in a locked office. If the data is stored on a laptop computer, it must be encrypted to avoid its disclosure in case of loss or theft.
- **Access security** - Only individuals who are listed in this application will have access to CP Commons Controlled Data. If copies of the CP Commons Controlled Data are stored locally on a shared computer system or a file server, then they must be strong password and/or encryption protected so that only the individuals named in the application have access to it. If the computer that holds CP Commons Controlled Data is backed up, the backup media must be encrypted and/or stored in a physically secure location.
- **Network security** - If CP Commons Controlled Data is stored on a network-accessible computer, there must be controls in place to prevent access by computer "hackers", or contamination by viruses, malware and spyware. Network security is usually implemented by your institution's IT department and will consist of some combination of network firewalls, network intrusion monitoring, and virus scanning software.
- **End of project** - After finishing the Research Project for which you are requesting access or if your access approval is terminated, you must securely destroy all local copies of the CP Commons controlled Data, including any backup copies. However, if necessary, you may still keep the CP Commons Controlled Data for archival purpose in conformity with national audits or other legal requirements.
- **Training** - Everyone who will access and/or use CP Commons Controlled Data must be trained in the responsible use of personal health information, familiarized with the terms and conditions of the Data Access Agreement, and briefed on your security plans.
- **Computer Cloud Use** - You may transfer the CP Commons Controlled Data into your own AWS storage container or another private or commercial cloud environment for analytical purposes. If you do so, you acknowledge that you maintain responsibility for the data and you agree that: you must take care to apply strong encryption to the data while in transit and at rest; restrict access to stored copies of the data to yourself, authorized personnel, students, and authorized collaborators listed on this Project Proposal; use firewall rules to restrict ingress and egress from virtual machines to trusted network address(es); keep virtual machines that host controlled data up to date with security patches; and destroy all copies of the data, including snapshots and backups, at the end of the research Project or if your application is not renewed; and ensure there is an agreement in place with your cloud provider that ensures you can meet these requirements.

  Any use of a private or commercial cloud is between you and the cloud provider. To the extent permitted by law CPA accepts no responsibility for any interaction between

you and the cloud provider and is released from any liability arising out of or in any way connected with such interaction.

Access to CP Commons Controlled Data is a procedure that entails legal and ethical obligations. You and your institution must have modern, up to date, information technology (IT) policies in place that must minimally include the following items:

- Logging and auditing of access to data and to computer network
- Password protection to computer network
- Virus and malware protection to computers on computer network
- Auditable data destruction procedure, when necessary
- Secure data backup procedure, when necessary
- Strong encryption on any portable device which may store or provide access to CP Commons controlled access data
- Privacy breach notification

You MUST agree to the following procedures in order to have access to the CP Commons Controlled Data:

☐	**Yes**, You will keep all computer systems on which CP Commons Controlled Data reside, or which provide access to such data, up-to-date with respect to software patches and antivirus file definitions (if applicable).
☐	**Yes**, You will protect CP Commons Controlled Data against disclosure to and use by unauthorized individuals.
☐	**Yes**, You will monitor and control which individuals have access to CP Commons controlled Data.
☐	**Yes**, You will securely destroy all copies of CP Commons Controlled Data in accordance with the terms and conditions of the Data Access Agreement.
☐	**Yes**, You will familiarize all individuals who have access to CP Commons Controlled Data with the restrictions on its use.
☐	**Yes**, You agree to swiftly provide a copy of both your institutional and Research Project related IT policy documents upon request from a CP Commons Data Custodian.
☐	**Yes**, You will notify the CP Commons Data Custodians immediately if you become aware or suspect that someone has gained unauthorized access to the CP Commons Controlled Data.

6. Data Access Agreement

This agreement governs the terms of access to the **CP Commons Controlled Data** (defined below). In signing this agreement, you agree to be bound by the terms and conditions of access set out therein.

For the sake of clarity, the terms of access set out in this agreement apply to the Primary Investigator and to the Primary Investigator Institution(s) (as defined below). The current agreement is limited to the **CP Commons Controlled Data** (as defined below) and does not cover other data.

Agreement Details

CPA	Cerebral Palsy Alliance ABN 45 000 062 288 187 Allambie Road, Allambie Heights, NSW, 2100, Australia
Primary Investigator Institution	[Insert name of Primary Investigator Institution] Of [Insert address]
Primary Investigator	[Insert name of Primary Investigator] [Insert Address] [Insert Email]
	together Primary Investigator Institution and Primary Investigator are referred to as the **Recipients**
Commencement Date	[Insert date]
End Date	12 months after Commencement date, unless CPA agrees in writing to extend the agreement
Controlled Data	*To be filled out by the CP Commons Data Custodians.* [Insert description of the data including number of individuals, types of data (i.e., WGS, RNA-seq) to be shared from the CP Commons.]
Research Project	*To be filled out by the Primary Investigator.* [Insert description of the Research Project for which the Controlled data is to be used.]

1. Definitions

Collaborators means the individuals that have been listed in Section II of the New Project - Data Access Application.

Consortium means the International Cerebral Palsy Genomics Consortium (ICPGC), which is an international confederation of members pursuing the common mission to better understand the genomic causes of cerebral palsy.

Controlled Data means the data in the CP Commons as described in the Agreement Details.

CP Commons means the cloud-based database managed by CPA that holds de-identified genomic and clinical data collections of people with cerebral palsy (Research Participants) and their families.

CP Commons Data Custodians means the staff or contractors of CPA who oversee the management and day-to-day running of the CP Commons.

Data Access Application means the CP Commons application form which was submitted by the Recipients and approved by CPA for access to and use of the Controlled Data on the terms of this Agreement.

Data Owner means the institution that originally submitted the Controlled Data to the CP Commons.

Data Breach means when information is lost or subjected to unauthorised access, modification, use or disclosure or other misuse. A data breach may be caused by malicious action (by an external or insider party), human error, or failure in information handling or security systems.

External Collaborator means a collaborator of the Primary Investigator who is not a listed Collaborator in Section II of the New Project - Data Access Application.

Primary Investigator means the applicant who is accessing data in the CP Commons for the Research Project and a party to this Agreement.

Primary Investigator Institution means the Institution at which the Primary Investigator is employed, affiliated, or enrolled and a party to this Agreement.

Publications includes articles published in print journal, electronic journal, reviews, books, posters and other written and verbal presentations of research.

Purpose means the purpose for which the Controlled Data will be used by the Recipients as described in the Data Access Application, which must be for the advancement of biomedical science research exclusively in connection with the Research Project.

Recipients means the Primary Investigator Institution and Primary Investigator.

Research Participant means an individual whose data has been uploaded to the CP Commons.

2. Interpretation

In this Agreement, the following rules apply:

a) headings are for convenience only and do not affect interpretation;

b) another grammatical form of a defined word or expression has a corresponding meaning;

c) a reference to time is to time in Sydney, Australia;

d) a reference to legislation (including subordinate legislation) is to that legislation as amended, re-enacted or replaced, and includes any subordinate legislation issued under it;

e) a reference to a person includes a natural person, partnership, body corporate, association, governmental or local authority or agency or otherentity; and

f) the meaning of general words is not limited by specific examples introduced by *including, for example* or similar expressions.

3. Use of Controlled Data

3.1 The Recipients agree to only use the Controlled Data for the Purpose.

3.2 The Recipients agree not to transfer or disclose the Controlled Data, in whole or part, to anyone not listed as a "Collaborator" of the Research Project in the Data Access Application,

except as otherwise required by law. If the Recipients wish to share the Controlled Data with an External Collaborator, the External Collaborator must complete a separate Data Access Application for the Controlled Data.

3.3 The Recipients agrees to use the Controlled Data in compliance with all Consortium policies made available from time to time.

3.4 The Recipients accept that CPA, the Consortium and the member institutions of the Consortium, including the Data Owner, bear no responsibility for the further analysis or interpretation of the Controlled Data, over and above that published by the Consortium.

3.5 The Recipients agree to destroy or discard any Controlled Data held, once it is no longer required for the "Research Project" unless it is required to retain the Controlled Data for legal requirements.

3.6 If requested, the Recipients will allow any applicable documentation to be inspected to verify that they are complying with the terms of this Agreement.

3.7 In addition to the training requirements set out in the Data Access Application, the Recipients agree to distribute a copy of this Agreement and explain its content to any listed Collaborator in the Data Access Application.

3.8 The Recipients must update the "Collaborators" to reflect any changes or departures in researchers, collaborators and personnel within 30 days of the changes made. This update can be sent by e-mail to this address: info@icpgc.org, with the subject: Collaborator Change.

3.9 The Recipients must notify the CP Commons Data Custodians prior to any significant changes to the Research Project's protocol. This update can be sent by e-mail to this address: info@icpgc.org, with the subject: Protocol Change.

3.10 The Recipients must notify the CP Commons Data Custodians (info@icpgc.org) as soon as they become aware of a breach of the terms or conditions of this Agreement.

4. Security and Protection

4.1 The Recipients must keep the Controlled Data secure and protected from unauthorised access, misuse, damage, destruction, unauthorised disclosure, modification, or theft.

4.2 The Recipients must not distribute or release the Controlled Data to any third party.

4.3 The Recipients must not use any Controlled Data to attempt to re-identify a Research Participant and if a Recipient becomes aware of any attempt to do so (including by any of the Recipient's Collaborators), the Recipient must notify the CP Commons Data Custodians immediately.

4.4 In the event that Recipient becomes aware of a Data Breach (or suspected Data Breach) relating to the Controlled Data or any other use or disclosure of the Controlled Data that is inconsistent with this Agreement, the Recipient must notify the CP Commons Data Custodians immediately.

4.5 The Recipients must cooperate with CPA, and provide all reasonable assistance which CPA may request in order to remedy and otherwise manage any Data Breach, whether or not caused by or contributed to by the Recipients.

4.6 The Recipients agree to provide written notice to the CP Commons Data Custodians upon gaining knowledge of the occurrence of any of the following: (1) abnormalities in the Controlled Data, or (ii) evidence that certain Controlled Data is incorrect.

4.7 The Recipient is responsible for monitoring and ensuring that all Collaborators comply with the terms of this Agreement and will be liable for any breaches by Collaborators.

5. Publication and Outputs

5.1 The Recipients agree to recognising the contribution of the Consortium and include a proper acknowledgement in all reports or publications resulting from the Recipients use of the Controlled Data.

5.2 The Recipients agree to provide a project summary at the end of the Research Project. This Project summary must include links to any publications as well as any findings from the Controlled Data that have clinical implications directly related to the condition being investigated (Pertinent Findings) or findings that are unintentionally discovered during the evaluation of the Controlled Data, that have clinical implications for the individual but are not associated with the clinical condition being investigated (Incidental Findings) that the Primary Investigators (and his or her staff or students) have uncovered during their analysis of the Controlled Data. This Project summary is due within 30 days of Project closure.

6. Intellectual Property

6.1 The Recipients agree not to make intellectual property claims on the Controlled Data and not to use intellectual property protection in a way that would prevent or block access to, or use of, any element of the Controlled Data, or conclusion drawn directly from the Controlled Data.

6.2 The Recipients acknowledge that the Controlled Data is the property of the Data Owner. CPA grants to the Recipients a non-exclusive, non-transferable, royalty-free licence to use the Controlled Data for the Purpose. The Recipients must not sell, loan or otherwise provide the Controlled Data to any other party for any purpose without the prior written consent of CPA.

7. Term and Limitation, Suspension and Termination

7.1 This Agreement will remain in effect from the Commencement Date until the earlier of:

(a) the End Date; or
(b) until such time as the Agreement is terminated in accordance with this clause 7.

7.2 CPA may by written notice limit or suspend the Recipients' right to access the Controlled Data, or terminate this Agreement with immediate effect, in any of the following circumstances:

(a) where a Recipient is in breach of this Agreement;
(b) if CPA becomes aware that the security of the Controlled Data or the CP Commons has been compromised;
(c) to comply with any legal requirement or any request or direction by a law enforcement agency; or
(d) where CPA, acting reasonably, determines that access or use of the Controlled Data should be limited, suspended or terminated for any other reason.

7.3 The Recipients may terminate this Agreement with immediate effect on written notice to CPA.

7.4 In the event of termination of this Agreement, CPA will immediately terminate the Recipients' access to CP Commons and the Recipient must follow CPA's instructions in relation to the destruction or return of any local copies of the Controlled Data in the power, possession or control of the Recipients.

7.5 Termination of this Agreement does not affect any accrued rights or remedies whicha party may have.

8. Liability and Indemnity

8.1 To the extent permitted by law:

(a) CPA gives no guarantee, warranty or representation in relation to the Controlled Data, including in relation to its availability, quality, fitness of purpose or security or in relation to the non-infringement of any third party Intellectual Property Rights. Any warranties or guarantees that may be implied or conferred by statute, custom or the general law are expressly excluded;
(b) the Recipients assume all liability for damages which may arise from its use, storage or disposal of the Controlled Data (including any decision made, or action taken, in relation to or in reliance upon the Controlled Data);

(c) CPA will not be liable to the Recipients for any loss, claim or demand made by the Recipients, or made against the Recipients by any other party, due to or arising from the use of the Controlled Data by the Recipients, except where caused by the gross negligence or wilful misconduct of CPA.

8.2 The Recipients agree to indemnify CPA against any loss, claim or demand that the CPA may sustain or incur in connection with:

(a) a breach by the Recipients of this Agreement;
(b) any Data Breach which is caused or contributed to by the Recipients;
(c) any unlawful or negligent act or omission of the Recipients under this Agreement;
(d) any breach by the Recipients of any third party Intellectual Property Rights, except to the extent that CPA caused or contributed to such loss, claim or demand.

9. General

9.1 Clauses 1, 2, 5, 6, 7.4, 8 and 9 will survive termination of this Agreement.

9.2 This Agreement may be executed in any number of counterparts. All counterparts will constitute one instrument. The parties agree that email signatures will be accepted as originals.

9.3 This Agreement may not be amended or varied other than with the written agreement of the Recipients and CPA.

9.4 A party must not assign or otherwise transfer any or all of its rights arising out of this Agreement without the written consent of the other party.

9.5 This Agreement constitutes the entire agreement between the parties with respect to the Recipients' access to and use of the Controlled Data.

9.6 These terms are governed by the law of the State of New South Wales, Australia. The parties submit to the exclusive jurisdiction of its courts.

7. Execution

Executed as an Agreement

Signed by **Cerebral Palsy Alliance** ABN 45 000 062 288 by an authorised signatory in the presence of:

……..……..	……..……..
Signature of authorised person	Signature of witness
……..……..	……..……..
Name of authorised person	Name of witness

Date: ……..……..

Signed by [Insert organisation name] ABN [INSERT] (as applicable) in the presence of:

……..……..	……..……..
Signature of authorised person	Signature of witness
……..……..	……..……..
Name of authorised person	Name of witness

Date: ……..……..

Read and acknowledged by :

……..……..

Signature of Primary Investigator

……..……..

Name of Primary Investigator

Date: ……..……..

Appendix 4 - Two Step De-identification of Participant Data

Important!

- Participants submitted to the CP Commons must have a single unique code that is linked to all data assets shared to the CP Commons. This code is called the SUBMITTER_ID.
- The SUBMITTER_ID should not be linked directly to any personal identifying information from the original study records.
- Submitted SUBMITTER_IDs should be two steps removed from any personal information.
- In the CP Commons, the SUBMITTER_ID is only visible to the Data Owner of those Data and the Data Custodians. Other Users of the CP Commons, will see the CP Commons Unique ID.

Process:

Personal Information removed and replaced with a Unique Individual ID
Unique Individual ID from original study removed and replaced with SUBMITTER_ID for submission to the CP Commons.

Appendix 5 - Clinical Data Dictionary

Background

This data dictionary is based on the common data elements (CDEs) that were developed in 2019-2020 by the International Cerebral Palsy Genomics Consortium Phenotype Working Group. They are designed to be used for genomics research studies in cerebral palsy.

There are **123 CDEs**: 14 data elements make up the Minimum Data Set (Mandatory) and must be included in your data upload. The other data elements include 49x Core, 44x Recommended and 16x Exploratory.

Please use this data dictionary to ensure that your data is correctly coded prior to upload into the CP Commons via the upload tool.
Please note metric values only.

Any problems recoding legacy datasets, please get in touch with the CP Commons Data Custodians on (info@icpgc.org).

Current working version: Version 1.2, updated 1st March, 2021

Unique IDs (2)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	NOTES /LOGIC
submitter_ID	**Mandatory**	Unique ID from submitting team	String, alphanumeric (no special characters)	Unique ID from submitting team, must be two steps removed from any personal information. The SUBMITTER_ID can only be viewed by Data Owners and CPA Data Custodians.
family_ID	**Mandatory**	Family ID	String, alphanumeric, permissible special characters (_)	A unique ID code assigned by the submitting team to a family. FAMILY_ID is not necessary for singleton cases.

Demographics (11)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	NOTES / LOGIC
pedigree	**Mandatory**	What is the relationship of this individual to the proband?	0, Proband/Index 1, Daughter 2, Granddaughter 3, Grandson 4, Half-brother 5, Half-sister 6, Maternal Aunt 7, Maternal cousin 8, Maternal grandfather 9, Maternal Grandmother 10, Maternal Uncle 11, Brother 12, Father 13, Mother 14, Sister 15, Nephew 16, Niece 17, Paternal Aunt 18, Paternal Cousin 19, Paternal Grandfather 20, Paternal Grandmother 21, Paternal Uncle 22, Son 99, Unknown	Any value between 23-98 or text will generate an error
clinical_status	**Mandatory**	Does the individual have any clinically relevant phenotype present?	0, Unaffected 1, Affected	Proband should always=1. NB: All subsequent Mandatory fields are only mandatory for individuals classified as 1
birth_year	**Mandatory**	What is the individuals' year of birth?	XXXX (years)	1900+

sex	**Mandatory**	What is the sex of the individual?	1, Male 2, Female 3, Intersex 99, Unknown	Boolean
birth_country	**Mandatory**	What country was the individual born in?	ISO-3166 2-alpha code	Only 2 alpha codes are permissible
country_residence	Recommended	What country does the individual currently reside in?	ISO-3166 2-alpha code	
personal_ethnicity	Recommended	What ethnicity does the individual identify with?	free text	
genotype_ethnicity	Recommended	What is the individuals' ethnicity; as determined by genotype?	free text	
vital_status	Core	What is the individual's vital status?	1, Alive 2, Dead 3, Unknown 4, Not reported	
death_age	Recommended	If deceased, what was the individual's age at time of death (years)?	1-99	
cause_of_death	Core	If deceased, what was the individuals primary cause of death (ICD-10 code)	3-7 alphanumeric code with a single period (.) that follows the first 3 alphanumeric characters (e.g. F91.9)	https://icd.who.int/browse10/2019/en

Diagnostics (10)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	NOTES /LOGIC
CP_phenotype	**Mandatory**	Does the individual have a permanent (non-paroxysmal) movement disorder?	0, No 1, Yes	
CP_onset	Recommended	Was onset of permanent (non-paroxysmal) movement disorder at age 2 or younger?	0, No 1, Yes	
degenerative	**Mandatory**	Is the individual's clinical course degenerative?	0, No 1, Yes	
CP_dx	Core	Has the individual been diagnosed with CP?	0, No 1, Yes	
CP_dx_year	Recommended	What year was the individual diagnosed with CP?	YYYY	
confirmed_CP_dx	Core	Was the diagnosis of CP confirmed at age 5 or older?	0, No 1, Yes 2, Not yet five years of age 99, Unknown	
timing_CP_injury	Recommended	During what period did the individuals' CP-related brain disturbance occur?	1, Prenatal 2, Perinatal 3, Post-neonatal 99, Unknown	
ORDO	Core	Has the individual been diagnosed with a known syndrome? (ORDO codes)	3-6 numeric code	List all applicable ORDO codes (only use a space between each code, NO commas) "Blank" if patient phenotype is unknown.
OMIM	Core	Has the individual been diagnosed with a genetic disorder? (OMIM)	6-digit integer (minimum 100000)	List all applicable OMIM codes (only use a space between each code, NO commas) "Blank" if patient phenotype is unknown.

ICD10	Core	Please describe any medical conditions the individual may have (ICD-10 codes)	3-7 alphanumeric code with a single period (.) that follows the first 3 alphanumeric characters (e.g. F91.9)	List all applicable ICD-10 codes (only use a space between each code, NO commas) "Blank" if patient phenotype is unknown.

Clinical traits (32)

FIELD	CATEGORY	DESCRIPTION	VALUE	
primary_motor	**Mandatory**	What is the predominant motor type of the individuals CP?	1, Spastic 2, Dyskinetic - Dystonic 3, Dyskinetic - Choreoathetosis 4, Ataxic 5, Hypotonic	
primary_motor _laterality	**Mandatory**	What is the laterality of the individuals predominant motor type?	1, Unilateral 2, Bilateral	
prim_spast_topo	Core	What is the individuals' predominant spastic topography?	1, Left hemiplegia 2, Right hemiplegia 3, Diplegia 4, Quadriplegia	This can only be responded if [primary_motor] has a value of (1).
prim_dyskinesia _type	Exploratory	If dyskinetic is predominant motor type, is the movement:	1, Focal 2, Generalised	This can only be responded if [primary_motor] has a value of (2 or 3).
second_motor	Core	What is the secondary motor type that the individual presents with?	1, Spastic 2, Dyskinetic - Dystonic 3, Dyskinetic - Choreoathetosis 4, Ataxic 5, Hypotonic	
second_motor _laterality	Core	What is the laterality of the individuals' secondary motor type?	1, Unilateral 2, Bilateral	
second_spast _topo	Core	Additional description of dyskinesia presentation	1, Left hemiplegia 2, Right hemiplegia 3, Diplegia 4, Quadriplegia	This can only be responded if [second_motor] has a value of (1).
second_dyskinesia _type	Exploratory	What is the individuals' secondary spastic topography?	1, Focal 2, Generalised	This can only be responded if [second_motor] has a value of (2 or 3).
second_NDD	Core	Does the individual have another neurodevelopmental disability?	0, No 1, Yes	

autism	Core	Does the individual have autism?	0, No 1, Yes	
autism_severity	Core	If the individual has an autism, what is the severity of the autism? Refer to DSM-5	1, Level 1: Requiring support 2, Level 2: Requiring substantial support 3, Level 3: Requiring very substantial support	This can only be responded if [autism] has a value of (1).
ADHD	Core	Does the individual have ADHD?	0, No 1, Yes	
ADHD_type	Core	If the individual has ADHD, what type of ADHD?	1, Primarily Hyperactive-Impulsive ADHD 2, Primarily Inattentive ADHD (formerly called ADD) 3, Combined Type ADHD	This can only be responded if [ADHD] has a value of (1).
ADHD_severity	Core	If the individual has an ADHD, what is the severity of the ADHD? Refer to DSM-5	1, Mild 2, Moderate 3, Severe	This can only be responded if [ADHD] has a value of (1).
epilepsy	**Mandatory**	Does the individual have epilepsy?	0, No 1, Yes 99, Unknown	
epilepsy_type	Core	If the individual has epilepsy, what type of epilepsy?	free text	This can only be responded if [epilepsy] has a value of (1).
ID	Core	Does the child have an intellectual impairment?	0, No 1, Yes	
ID_severity	Core	If the individual has an intellectual impairment, what is the severity of the impairment? *(ICD10 codes F70 to F73)*	1, Mild (IQ50 - 69) HP: 0001256 2, Moderate (IQ 35 - 49) HP: 0002342 3, Severe (IQ 20 - 34) HP: 0010864 4, Profound (IQ < 20) HP: 0002187 5, Impairment unspecified (IQ < 50)	This can only be responded if [ID] has a value of (1).
second_NDD _other	Core	If other, please specify:	free text	

visual_impairment	Core	Does the individual have a visual impairment?	0, No 1, Yes	
visual_severity	Core	If the individual has a visual impairment, please describe:	free text	This can only be responded if [visual_impairment] has a value of (1).
hearing_impairment	Core	Does the individual have a hearing impairment?	0, No 1, Yes	
hearing_severity	Core	If the individual has a hearing impairment, please describe:	free text	This can only be responded if [hearing_severity] has a value of (1).
congenital _anomalies	Core	Does the individual have any congenital anomalies?	0, No 1, Yes	
congenital _anomalies_type	Core	If the individual has a congenital anomaly/ies, which major anatomical system/s are involved?	1, Nervous system 2, Cardiovascular system 3, Respiratory system 4, Gastrointestinal system 5, Genital organs 6, Urinary system 7, Musculature system 8, Skeletal system 9, Integumentary system (skin) 10, Eye 11, Ear 12, Nose	This can only be responded if [congenital_anomalies] has a value of (1). Add all permissible values applicable – NO commas only spaces between each value.
congenital _anomalies_detail	Recommended	If the individual does have congenital anomalies, please describe: [free text]	free text	Add any free text to describe the congenital anomalies. This can only be responded if [congenital_anomalies] has a value of (1).

radiology_text	Recommended	If the individual has any relevant radiological findings, please describe: [free text]	free text	Add any radiology free text. Please watch for personal identifiers
chronic_pain	Exploratory	Has the individual reported experiencing chronic pain?	0, No 1, Yes - self report 2, Yes - proxy report	
chronic_pain_age	Exploratory	At what age was chronic pain reported?	XX	Integer
chronic_pain _measure	Exploratory	What validated outcome measure was used to assess chronic pain?	free text	
positive_HPOs	**Mandatory**	Please list positive HPO traits (preferably minimum of 3):	*e.g. HP:0100277*	Add all applicable HPO that are present in the individual. Spaces between values, NO commas.
negative_HPOs	Exploratory	Please list phenotypic traits that are NOT present	*e.g. HP:0012447*	Add all applicable HPO that are NOT present in the individual. Spaces between values, NO commas.

CP-Specific Assessments (13)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	NOTES / LOGIC
GM	Recommended	Was a General Movements Assessment performed at 12 weeks of age, and if so what was the score?	0, No 1, Yes - normal 2, Yes - Abnormal fidgety 3, Yes - Absent fidgety	
HINE	Recommended	Was the Hammersmith Infant Neurological Examination (HINE) performed?	0, No 1, Yes	
HINE_age	Recommended	If the HINE was performed, at what age was the assessment performed?	MM	This can only be responded if [HINE] has a value of (1). Corrected age in months between 2-24 (integer)
HINE_outcome	Recommended	If the HINE was performed, what was the outcome of the assessment?	0-78	This can only be responded if [HINE] has a value of (1). Global score between 0-78 (Integer)
GMFCS	**Mandatory**	Gross Motor Functioning Classification System (GMFCS)© Score	1, GMFCS I 2, GMFCS II 3, GMFCS III 4, GMFCS IV 5, GMFCS V 99, Unknown	
BFMF	Recommended	Bimanual Fine Motor Function (BFMF) ©Score	1, BFMF I 2, BFMF II 3, BFMF III 4, BFMF IV 5, BFMF V	
MACS	Recommended	Manual Ability Classification System (MACS)© Score	1, MACS I 2, MACS II 3, MACS III 4, MACS IV 5, MACS V	

CFCS	Recommended	Communication Function Classification System (CFCS)© Score	1, CFCS I 2, CFCS II 3, CFCS III 4, CFCS IV 5, CFCS V
VFCS	Exploratory	Visual Function Classification System (VFCS)© Score	1, VFCS I 2, VFCS II 3, VFCS III 4, VFCS IV 5, VFCS V
MRICS	Core	Predominant Brain Pattern (Expanded MRI Classification System)	1, **(A) Maldevelopments** 2, (A.1) Maldevelopments - Disorders of cortical development (proliferation and/or migration and/or organisation) 3, (A.2) Maldevelopments - Other (ex: holoprosencephaly, Dandy-Walker formation, corpus callosum agenesis, cerebellar hypoplasia) 3, **(B) Predominant White Matter Injury (PWMI)** 4, (B.1) PWMI - PVL (mild/severe) 5, (B.2) PWMI - Sequelae of IVH or PVH infarction 6, (B.3) PWMI - Combination of PVL and IVH sequelae 7, **(C) Predominant Grey matter Injury (PGMI)** 8, (C.1) PGMI - Basal ganglia/thalamus lesions (mild/moderate/severe) 9, (C.2) PGMI - cortico-subcortical lesions only (watershed lesions in parasagittal distribution/multicystic encephalomalacia) not covered under C3 10, (C.3) PGMI - Arterial infarctions (middle cerebral artery/other) 11, **(D) Miscellaneous changes** (ex: cerebellar atrophy, cerebral atrophy, delayed myelination, ventriculomegaly not covered under B, haemorrhage not covered under B, brainstem lesions, calcifications) 12, **(E) Normal**
swallowing	Recommended	Participant eating/drinking/swallowing	1, Standard mean for age - no modifications required 2, Requires modified diet

VIKING	Recommended	Viking Speech Scale© Score	1, Score I 2, Score II 3, Score III 4, Score IV	
ICF	Recommended	WHO Disability Assessment Schedule 2.0	0-100	Summary disability profile score according to WHODAS2.0 (integer)

Family History (16)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	
mat_age	Core	Maternal age at time of birth	YY	Age at birth of proband (integer)
mat_cob	Recommended	Maternal Country of Birth	ISO-3166 2-alpha code	
mat_edu	Recommended	Maternal Education	1, Primary school only 2, High school graduate or equivalent 3, Occupational/Technical/Vocational 4, Undergraduate university qualification 5, Postgraduate university qualification 6, Never attended	
pat_age	Core	Paternal age at time of birth	YY (age at birth of proband)	
pat_cob	Recommended	Paternal Country of Birth	ISO-3166 2-alpha code	
pat_edu	Recommended	Paternal Education	1, Primary school only 2, High school graduate nor equivalent 3, Occupational/Technical/Vocational 4, Undergraduate university qualification 5, Postgraduate university qualification 6, Never attended	
consanguinity	Core	Is any consanguinity reported?	0, No 1, Yes	
CP_family_hx	Core	Is there a family history of CP?	0, No 1, Yes	
CP_sibling_hx	Core	Does a sibling have a diagnosis of CP?	0, No 1, Yes	

CP_family_hx_detail	Recommended	If there is a family history of CP, please select all members that have been diagnosed with CP?	List all that apply: 0, None 1, Daughter 2, Granddaughter 3, Grandson 4, Half-brother 5, Half-sister 6, Maternal Aunt 7, Maternal cousin 8, Maternal grandfather 9, Maternal Grandmother 10, Maternal Uncle 11, Brother 12, Father 13, Mother 14, Sister 15, Nephew 16, Niece 17, Paternal Aunt 18, Paternal Cousin 19, Paternal Grandfather 20, Paternal Grandmother 21, Paternal Uncle 22, Son	For every family member that has CP, add the applicable number. SPACES between numbers, if more than one. No commas.
neuro_family_hx	Core	Is there a family history of neurological disorders?	0, No 1, Yes	
neuro_family_hx_detail	Recommended	If there is a family history of neurological disorders, please list all family members that have been diagnosed with a neurological disorder.	0, None 1, Daughter 2, Granddaughter 3, Grandson 4, Half-brother 5, Half-sister	For every family member that has a relevant neurological history, add the applicable number. SPACES between numbers, if more than one. No commas.

		Please separate each family member with a space	6, Maternal Aunt 7, Maternal cousin 8, Maternal grandfather 9, Maternal Grandmother 10, Maternal Uncle 11, Brother 12, Father 13, Mother 14, Sister 15, Nephew 16, Niece 17, Paternal Aunt 18, Paternal Cousin 19, Paternal Grandfather 20, Paternal Grandmother 21, Paternal Uncle 22, Son 99, Unknown	**NB Data Owners:** This is NOT Compulsory but Highly Recommended. In order to link these cases to the original proband, Data Owners should create a Family_ID for family members. You do not need to upload genomic data for these family members. If there is a positive family history of neurological conditions, each family member should be added to the dataset as a new row and their [clinical_status] must = 1, and data must be entered into either: known_syndrome (Diagnostics), OMIM (Diagnostics), med_condition (Diagnostics), or positive_HPO (Clinical symptoms and physical signs).
neuro_family_hx_text	Recommended	If there is a family history of neurological disorders, please describe:	Free text	Please provide free text detail describing the relevant neurological family history.
siblings	Recommended	Does the individual have sibling/s?	0, No 1, Yes	
sib_number	Recommended	If the individual has sibling/s, how many siblings?	XX	This can only be responded if [siblings] has a value of (1) [integer].
sib_phenotype	Recommended	Please describe any relevant clinical conditions or phenotypes present in the sibling(s):	Free text	This can only be responded if [sibling] has a value of (1). Watch for personal information

Antenatal and Neonatal History (39)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	NOTES /LOGIC
mat_ht	Exploratory	What was the maternal height at beginning of pregnancy with this individual?	XXX	Metric [numeric]
mat_wt	Exploratory	What was the maternal weight at the beginning of the pregnancy with this pregnancy?	XXX	Metric [numeric]
gravidity	Exploratory	Total number of confirmed pregnancies prior to this individual	XX	0+ [integer]
preterm_birth	Core	Was there a preterm birth prior to this individual?	0, No previous preterm birth 1, Yes, there was preterm birth	
perinatal_death	Core	Was there a perinatal death prior to this individual? (includes more than 20 completed weeks, still births and death during neonatal period)	0, No previous perinatal death 1, Yes, there was perinatal death	
conception	Core	Was the child conceived through assisted conception?	1, No 2, Yes - Unknown 3, Yes - IVF 4, Yes - ICSI 5, Yes - GIFT 6, Yes - AI 7, Yes - Ovulation Stimulation only 8, Yes - Other	
first_an_visit	Recommended	Total number of completed weeks of the pregnancy at first antenatal visit	XX	1-40 completed weeks [integer]
mat_principal_morbid	Recommended	Pre-existing maternal morbidity, principal diagnosis	free text	
mat_add_morbid	Recommended	Pre-existing maternal morbidity, additional diagnoses	free text	

gestation_principal_morbid	Recommended	Maternal morbidity (pregnancy - and birth-related), principal diagnosis	free text	
gestation_add_morbid	Recommended	Maternal morbidity (pregnancy- and birth-related), additional diagnoses	free text	
G_smoke	Recommended	Did the mother smoke during pregnancy?	0, No 1, Yes	
G_alcohol	Recommended	How many standard alcoholic drinks did the mother consume per week during pregnancy?	XX	integer
G_drug	Recommended	During the first trimester, did the mother use any recreational drugs?	0, No 1, Yes	
teratogen	Recommended	Please describe any teratogen exposure prior to, or during pregnancy	free text	
fetal_abnormality	Core	Was a fetal abnormality identified by ultrasound during pregnancy?	0, No 1, Yes	
fetal_abnormal_detail	Recommended	If yes, please describe	free text	This can only be responded if [fetal_abnormality] has a value of (1).
birth_facility	Exploratory	What facility was the individual born in?	1, Hospital 2, Birth centre attached to hospital 3, Birth centre free standing 4, Home birth planned 5, Home birth unplanned 6, Born before arrival at hospital 7, Born outside home/hospital without medical assistance 8, Other	
mag_sulphate	Exploratory	Was magnesium sulphate administered to the mother during pregnancy?	0, No 1, Yes	

labour_onset	Recommended	Onset of labour during this individual's birth	1, Spontaneous 2, Induced 3, No labour	
labour_ hyperthermia	Exploratory	Did the mother treat with hyperthermia during labour?	0, No 1, Yes	
birth_presentation	Recommended	Presentation at birth	1, Vertex 2, Breech 3, Face 4, Brow 5, Transverse	
delivery	Core	Method of delivery	1, Vaginal non-instrumental 2, Vaginal instrumental 3, Emergency c-section 4, Elective c-section 99, Unknown	
ga	Core	Gestational age (completed weeks)	20-46	
bw	Core	Birth weight (grams)	200-5000	Metric
head_cir	Core	Head circumference of individual at time of birth (centimetre)	10-60	Metric
plurality	Core	Birth plurality	1, Singleton 2, Twin 3, Triplets 4, Quadruplets 5, Quintuplets 6, Sextuplets 7, Other	

birth_order	Recommended	Birth order if a multiple birth	1, First of multiple 2, Second of multiple 3, Third of multiple 4, Fourth of multiple 5, Fifth of multiple 6, Sixth of multiple 7, Other	This can only be responded if [plurality] has a value of (2+).
zygotic	Core	If this was a multiple birth, was it:	1, Monozygotic 2, Dizygotic 3, Other	This can only be responded if [plurality] has a value of (2+).
amniotic	Core	If this was a multiple birth, was it:	1, Monoamniotic 2, Diamniotic 3, Triamniotic 4, Other	This can only be responded if [plurality] has a value of (2+).
chorionic	Core	If this was a multiple birth, was it:	1, Monochorionic 2, Dichorionic 3, Trichorionic 4, Other	This can only be responded if [plurality] has a value of (2+).
newborn_care	Core	Did the individual receive more than routine newborn care?	0, No 1, Yes, NICU or special care	
phototherapy	Recommended	Did the individual receive phototherapy?	0, No 1, Yes	
hypothermia	Core	Did the individual receive therapeutic hypothermia?	0, No 1, Yes	
ventilation	Core	Did the individual receive mechanical ventilation?	0, No 1, Yes	
neo_infection	Core	Was infection present during the neonatal period?	0, No 1, Yes	

placenta_wt	Recommended	Placental weight (trimmed of extra placental membranes and umbilical cord) (grams)	10-3000	Metric, decimal
gross_placental	Exploratory	Gross placental abnormalities detected	free text	
histo_placental	Exploratory	Histological placental abnormalities detected	free text	

Appendix 6 - Examples of Personal Identifiers

These are examples of data that is not accepted about Research Participants.

Identifier	Example
Personal Information	Name All dates, including birth, death, diagnosis, test (except year) Telephone number inc. mobile/cell numbers, fax numbers Geographic data inc. coordinates or addresses (except country) Email addresses Passport, birth certificate, driver's license numbers Banking or credit information Employment record, employer's name, job title Medical record, health insurance, life insurance numbers License plates Biomedical device identifiers or serial numbers Internet Protocol addresses
Unique Codes	These are any national identity numbers that are unique to an individual, such as (but not limited to): Tax File Number Social Security Number National Insurance Number Medicare Number Medicaid Number National Identification Number
Biometric data	Fingerprint Retinal scan Photo or videos ** Magnetic Resonance Imaging or Computational Topography or X-ray imaging: only images that have had identifiers stripped and face masked are accepted

Appendix 7 - Accepted Genomic Data Formats for Submission

Important!

- All samples must be de-multiplexed prior to upload. Reads for different individuals should be submitted using separate files. The only exception is when a BAM or CRAM file contains reads for a large number of samples intended to always be analysed together – this needs to be agreed upon with the CP Commons Data Custodian during the New Project Request stage.
- Files must be compressed: Files that are in a human-readable text format (gVCF) must be compressed before they are uploaded to the CP Commons (use gzip, bzip2 or similar). Files that are not in a human readable text format (BAM, CRAM, SFF) are already in a compact format so additional compression is not required.
- Files must have an MD5 checksum registered in the metadata.

Generic Formats Accepted

Accepted Format	Notes
Genomic VCF format	Each submitted gVCF file must: • be harmonised using the appropriate ICPGC pipeline accessed through the CP Commons Code Share • be submitted as a separate run use the .gvcf file name suffix (e.g. 'filename.gvcf.gz')
Binary Alignment Map files (BAM)	Each submitted BAM file must: • be compatible with the SAM/BAM Format Specification • be readable with Samtools • be submitted as a separate run • use the .bam file name suffix (e.g. 'filename.bam')
CRAM (compressed version of BAM)	Files received in CRAM will be converted to BAM for processing gVCFs/VCFs. Each submitted CRAM file must: • be compatible with the CRAM Format Specification • be readable with Samtools • contain only reference sequences that exist in the CRAM Reference Registry • be submitted as a separate run • use the .cram file name suffix (e.g. 'filename.cram') A CRAM index (CRAI) file is created by the CP Commons for each submitted CRAM file and is available in the same directory as the CRAM file from which it was created. CRAM index file names start with the CRAM file name and end up with the .crai suffix (e.g. 'filename.cram.crai' for CRAM file 'filename.cram').

Appendix 8 - Sequencing Metadata Dictionary

Unique IDs (2)

FIELD NAME	CATEGORY	DESCRIPTION	VALUE	NOTES /LOGIC
submitter_ID	**Mandatory**	Unique ID from submitting team	String, alphanumeric (no special characters)	Unique ID from submitting team, must be two steps removed from any personal information. The SUBMITTER_ID can only be viewed by Data Owners and CPA Data Custodians.
family_ID	**Mandatory**	Family ID	String, alphanumeric, permissible special characters (_)	ID code assigned by the submitting team to a family. FAMILY_ID is not necessary for singleton cases.

Genomic Fields (15)

FIELD NAME	REQUIRED	DESCRIPTION	VALUE	NOTES /LOGIC
filename	**Mandatory**	Name of the genome datafile	String, alphanumeric, permissible special characters (_)	The name of the file
data_access_ restriction	**Mandatory**	Data Access Restrictions	1, Open-access data 2, Controlled-access data	Open – immediately accessible to users Controlled – Users must submit a data access request to access the data.
consent	**Mandatory**	Do you have consent from the participant to share their de-identified data?	0, No 1, Yes	
consent_no	Core	If no, have you received a consent waiver to share these data without the individuals consent?	0, No 1, Yes	Response required if [consent] has value of 0.
restrictions	Recommended	Does this data have additional data use restrictions?	0, No 1, Yes	

pipeline	Core	Please select the CP Commons harmonisation pipeline	1, Pre-processing (BAM only) 2, gVCF pipeline 3, SNPs + indels
file_format	Core	Please indicate the file format for this data file:	1, VCF 2, gVCF 3, BAM 4, TXT
molecular_class	Core	Broad categorisation of the molecular data:	1, Genome 2, Methylome 3, Transcriptome
exp_strategy	Core	Please select the sequencing strategy used to generate the data file:	1, Whole Genome Sequencing (WGS) 2, Whole Exome Sequencing (WES) 3, Array-CGH (CNV) 4, SNP-array 5, Methylation array 6, CpG island array 7, Whole Genome Bisulfite-sequencing (WGBS) 8, Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) 9, Chromatin immunoprecipitation sequencing (ChIP-seq) 10, microRNA sequencing (miRNA-seq) 11, Microarray 12, Whole RNA-seq 13, Whole transcriptome shotgun sequencing (WTSS) 14, Targeted RNA-seq amplicon
instrument	Recommended	Name of platform used for sequencing:	1, Affymetrix 2, Agilent 3, Illumina 4, Ion Torrent

			5, Nimblegen 6, PacBio	
instrument_model	Recommended	Name of instrument model used for sequencing:	String, alphanumeric, permissible special characters (_)	
instrument_centre	Recommended	Where was the sequencing performed?	String, alphanumeric, permissible special characters (_)	
raw_data	Core	Is the raw data stored in an open database?	0, No 1, Yes	
raw_data_detail	Core	If yes, please provide the hyperlink:	String, alphanumeric, permissible special characters (_ , . , / , :)	Response required on if [raw_data] has a value of (1).
md5sum	**Mandatory**	MD5 Sum:	Integer	The 128-bit hash value expressed as a 32 digit hexadecimal number (in lower case) used as a file's digital fingerprint.

References

1. Rosenbaum, P., et al., *Proposed definition and classification of cerebral palsy, April 2005*. Dev Med Child Neurol, 2007. 47: p. 571-576.
2. Himmelmann, K., et al., *Risk factors for cerebral palsy in children born at term*. Acta Obstet Gynecol Scand, 2011. 90(10): p. 1070-81.
3. McIntyre, S., et al., *A systematic review of risk factors for cerebral palsy in children born at term in developed countries*. Dev Med Child Neurol, 2013. 55(6): p. 499-508.
4. Fahey, M.C., et al., *The genetic basis of cerebral palsy*. Dev Med Child Neurol, 2017. 59(5): p. 462-469.
5. MacLennan, A.H., et al., *Cerebral palsy and genomics: an international consortium*. Dev Med Child Neurol, 2018. 60(2): p. 209-210.
6. MacLennan, A. H., et al., *Genetic or other causation should not change the clinical diagnosis of cerebral palsy*. Journal of child neurology, 2019. 34(8): p. 472-476.
7. McIntyre, S., I. Novak, and A. Cusick, *Consensus research priorities for cerebral palsy: a Delphi survey of consumers, researchers, and clinicians.* Dev Med Child Neurol, 2010. 52(3): p. 270-5.
8. Wilson, Y.A., et al., *People with Cerebral Palsy and their Family's Preferences about Genomics Research.* Public Health Genomics, 2021. [accepted manuscript].
9. Lecoquierre, F., et al., *Variant recurrence in neurodevelopmental disorders: the use of publicly available genomic data identifies clinically relevant pathogenic missense variants.* Genet Med, 2019. 21(11): p. 2504-2511.
10. Consortium, I.T.P.-C.A.o.W.G., *Pan-cancer analysis of whole genomes.* Nature, 2020. 578(7793): p. 82-93.
11. Yuen R.K., et al., *Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder.* Nature neuroscience, 2017. 20(4): p. 602-611.
12. Ruzzo, E.K., et al., *Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks*. Cell, 2019. 178(4): p. 850-866 e26.
13. Cable, J., et al., *Harnessing rare variants in neuropsychiatric and neurodevelopment disorders-a Keystone Symposia report*. Ann N Y Acad Sci, 2021.
14. Serdarevic, F., et al., *Polygenic Risk Scores for Developmental Disorders, Neuromotor Functioning During Infancy, and Autistic Traits in Childhood*. Biol Psychiatry, 2020. 87(2): p. 132-138.
15. Baynam, G., et al., *Improved Diagnosis and Care for Rare Diseases through Implementation of Precision Public Health Framework, in Rare Diseases Epidemiology: Update and Overview*, M. Posada de la Paz, D. Taruscio, and S.C. Groft, Editors. 2017, Springer International Publishing: Cham. p. 55-94.
16. Thompson, M.A., et al., *Coordinating an Oncology Precision Medicine Clinic Within an Integrated Health System: Lessons Learned in Year One*. J Patient Cent Res Rev, 2019. 6(1): p. 36-45.
17. Sobreira, N.L.M., et al., *Matchmaker Exchange*. Curr Protoc Hum Genet, 2017. 95: p. 9 31 1-9 31 15.

Images

Page 8
Vectors by pch.vector from FreePik

Page 13
Icon by Freepik from Freepik

Page 17
Icon by Eucalyp from Flaticon

Page 18
Icons by Becris from Flaticon
Icon by Eucalyp from Flaticon

Page 19
Icon by Nikita Golubev from Flaticon

All photos are courtesy of Cerebral Palsy Alliance

For more information,
please get in touch:

Email us at **info@icpgc.org**

Visit **https://icpgc.org/**

INTERNATIONAL
CP GENOMICS
CONSORTIUM

CP COMMONS

Cerebral Palsy
ALLIANCE