









ORIGINAL ARTICLE

Common data elements to standardize genomics studies in cerebral palsy

Yana A. Wilson^{1,2}  | Hayley Smithers-Sheedy^{1,2}  | Katarina Ostojic^{1,2}  | Emma Waight^{1,2}  |
 Michael C. Krueer^{3,4}  | Michael C. Fahey⁵  | ICPGC Phenotype Working Group* |
 Gareth Baynam^{6,7,8,9,10} | Jozef Géczy^{11,12,13}  | Nadia Badawi^{1,2,14} | Sarah McIntyre^{1,2,9} 

¹Sydney Medical School, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia

²Cerebral Palsy Alliance Research Institute, Specialty of Child & Adolescent Health, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia

³Pediatric Movement Disorders Program, Barrow Neurological Institute, Phoenix, Arizona, USA

⁴Departments of Child Health, Neurology, Cellular & Molecular Medicine and Program in Genetics, University of Arizona College of Medicine, Phoenix, Arizona, USA

⁵Department of Paediatrics, Monash University, Melbourne, VIC, Australia

⁶Western Australian Register of Developmental Anomalies King Edward Memorial Hospital, Perth, WA, Australia

⁷Faculty of Health and Medical Sciences, Division of Paediatrics, University of Western Australia, Perth, WA, Australia

⁸Institute for Immunology and Infectious Diseases, Murdoch University, Perth, WA, Australia

⁹Telethon Kids Institute, University of Western Australia, WA, Australia

¹⁰Spatial Sciences, Department of Science and Engineering, Curtin University, WA, Australia

¹¹Robinson Research Institute, The University of Adelaide, Adelaide, SA, Australia

¹²Adelaide Medical School, The University of Adelaide, SA, Australia

¹³South Australian Health and Medical Research Institute, Adelaide, SA, Australia

¹⁴Grace Centre for Newborn Intensive Care, The Children's Hospital at Westmead, Westmead, NSW, Australia

Correspondence

Yana Wilson, Cerebral Palsy Alliance Research Institute, Brain and Mind Centre, University of Sydney, 88 Mallet Street, Camperdown, NSW, 2050, Australia.
 Email: ywilson@cerebralpalsy.org.au

Funding information

PTC; PTC Therapeutics

Abstract

Aim: To define clinical common data elements (CDEs) and a mandatory minimum data set (MDS) for genomic studies of cerebral palsy (CP).

Method: Candidate data elements were collated following a review of the literature and existing CDEs. An online, three-round Delphi survey was used to rate each data element as either 'core', 'recommended', 'exploratory', or 'not required'. Members of the International Cerebral Palsy Genomics Consortium (ICPGC) rated the core CDEs as either mandatory or not, to form the MDS. For both the CDEs and the MDS, a data element was considered to have reached consensus if more than 75% of respondents agreed.

Results: Forty-six individuals from around the world formed the Delphi panel: consumers ($n=2$), scientists/researchers ($n=17$), medical ($n=19$), and allied health professionals ($n=8$). The CDEs include 107 data elements across six categories:

This original article is commented on by Dieterich on pages 1433–1434 of this issue.

Abbreviations: CDE, Common data element; ICPGC, International Cerebral Palsy Genomics Consortium; MDS, Minimum data set; NINDS, National Institute of Neurological Disorders and Stroke; PWG, Phenotype Working Group.

*Members of the ICPGC Phenotype Working Group are listed in the Acknowledgements.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Developmental Medicine & Child Neurology* published by John Wiley & Sons Ltd on behalf of Mac Keith Press.

demographics, diagnostics, family history, antenatal and neonatal details, clinical traits, and CP-specific assessments. Of these, 10 are mandatory, 42 core, 41 recommended, and 14 are exploratory.

Interpretation: The ICPGC CDEs provide a foundation for the standardization of phenotype data captured in CP genomic studies and will benefit international collaborations and pooling of data, particularly in rare conditions.

Cerebral palsy (CP) is a clinically heterogeneous condition of movement, posture, and motor function attributed to a non-progressive and permanent disturbance to the fetal or infant brain.¹ It is well recognized that the causal pathways to CP are complex and, in many cases, undefined. Known causal pathways include preterm birth, congenital anomalies, hypoxia-ischemia, infection, and pre- and perinatal stroke.^{2,3} Recently, it has become widely accepted that genetics also plays a role in the causal pathways to CP,⁴ which has led to the formation of the International Cerebral Palsy Genomics Consortium (ICPGC).⁵

The ICPGC aims to better understand the genome's role in CP through international collaboration via the collection and sharing of phenotypic and genomic data, and the interpretation and dissemination of findings to the broader community. Consequently, the CP Commons data sharing platform was developed to assist collaboration. It is pivotal to the success of genetic data sharing efforts that these data are accompanied by high-quality phenotype information, allowing researchers to identify more precise and reliable genotype–phenotype correlations.⁶ However, data collection is frequently fragmented as it has not been collected systematically or consistently across all institutions or countries. This variation creates notable barriers to information exchange due to a lack of interoperability and the need for data transformation, thereby diminishing the value of the data.

Common data elements (CDEs) are a practical method to standardize data collection across numerous sites and studies. Their primary purpose is to ensure reproducibility and interoperability among different data sets for data sharing and to improve the quality of data collection.⁷ CDEs are precisely defined data collection units comprised of one or more questions together with a set of valid responses. CDEs are usually defined through an iterative process from various stakeholders to obtain a broad consensus.⁷ The US National Institute of Health has been developing CDEs⁸ for over 20 years across many diagnostic groups, including CP.⁹ Similarly, population-based CP registries, such as the Australian Cerebral Palsy Register¹⁰ and the Surveillance of Cerebral Palsy in Europe,¹¹ also have well-established CDEs. Despite these high-quality resources, these CDEs were designed for the purposes of tracking epidemiological trends, measuring the impact of preventative strategies, streamlining outcome measures collected in clinical trials and intervention studies, as well as providing a framework for collaborative research. The CDEs described herein are

What this paper adds

- A set of 107 common data elements (CDEs) for genomics studies in cerebral palsy is provided.
- The CDEs include standard definitions and data values domains.
- The CDEs will facilitate international data sharing, collaboration, and improved clinical interpretation of findings.

designed to build upon these existing frameworks while focusing on aetiology, antenatal risk factors, and deep phenotyping, as is required for genomic studies of CP.

The first aim of this study was to define a set of CDEs for genomic studies in CP using the Delphi method to achieve consensus. The Delphi method consists of iterative rounds of surveys, whereby a group of experts can achieve broad consensus via the revision of their responses based on other participants' responses.¹² As survey responses are anonymous, it creates the opportunity for group-based decision-making as all voices are equally heard.¹² The Delphi method is widely used in many areas of health care for an array of research purposes,¹³ including the development of CDEs.^{14–17}

The second aim was to define a minimum data set (MDS) from the CDEs for genomic studies in CP. The MDS is here defined as the absolute minimum amount of data required to ensure the utility and value of data is achieved. The MDS aims to promote data collection and wider data sharing while minimizing the administrative burden for researchers and health care providers. Establishing an MDS can also promote participation as it can allow the inclusion of retrospective cohorts that may not have captured all the data defined in the CDEs.

METHOD

The development of the ICPGC CDEs for genomic studies of CP consisted of three phases: (1) a scoping review of the literature and development of the draft CDEs; (2) a modified, three-round Delphi process to achieve consensus on CDEs important for genomic studies of CP and to assign each data element to a hierarchical level based on their importance; and

(3) development of the MDS (Figure S1). The reporting of this study follows the Conducting and Reporting of Delphi Studies recommendations,¹⁸ which aim to improve the quality of Delphi study reporting.

Phase 1: identification and prioritization of a candidate set of CDEs

During this phase, a steering committee (YW, SMc, HSS) was established to carry out a scoping review of the literature, existing CDEs, and other standardization efforts. Clinical data elements were extracted from published studies (2005–2017) and the Australian Cerebral Palsy Biobank¹⁹ and imported into Microsoft Excel for processing, collating, and prioritization with existing CDEs from the Australian Cerebral Palsy Register,¹⁰ Surveillance of Cerebral Palsy in Europe,¹¹ the Childhood Cerebral Palsy Integrated Neuroscience Discovery Network,²⁰ National Institute of Neurological Disorders and Stroke (NINDS) CDEs for CP,⁹ the Australian and New Zealand Neonatal Network,²¹ the European Platform for Rare Disease Registries,²² and the National Centre for Advancing Translational Sciences Global Rare Disease Patient Registry Data Repository.²³

The clinical data elements were aggregated to show where commonalities existed across different studies or dictionaries and the differences in the data value domains. The list was reviewed by the steering committee, which used a risk-factor systematic review³ to guide the appropriate inclusion of data elements pertaining to CP aetiology. Data elements were removed if they (1) focused on interventions, treatments, and therapies of CP; (2) were considered of little interest in genomic studies of CP; or (3) the data element was infrequently collected.

The ICPGC Phenotype Working Group (PWG) reviewed the draft clinical data set. Data elements were prioritized for inclusion if the data element was considered important for genetic studies of CP; had an internationally standardized collection method, reporting variables, or ontologies; and had broad utility. Data elements with non-standardized collection methods reporting variables or ontologies (including free text options) could be included if the variable was important to CP aetiology.³

The steering committee collated a final selection of potential data elements for review in the modified Delphi process. The data elements were sorted into the following categories: demographics, diagnostics, clinical traits, CP-specific assessments, family history, and antenatal and neonatal details. The survey was piloted by KO and EW.

Phase 2: modified Delphi consensus process to select CDEs

The steering committee (YW, SMc, HSS) facilitated the three-round Delphi process and did not participate in the

Delphi rounds. A purposive sample of 76 panellists were invited via email to participate in Phase 2 (Appendix S1). The invited panellists included members of the ICPGC ($n=50$); professionals recommended by members of the ICPGC-PWG with expertise in data management, phenotyping, CP genetics, clinical genetics, and CP aetiology ($n=24$); and consumers ($n=2$). The invited panellists and members were from Australia, Canada, China, Finland, France, Israel, Japan, Sri Lanka, Sweden, the UK, and the USA. An invitation was also distributed to members of the Australian Academy of Cerebral Palsy and Developmental Medicine. Participant anonymity was maintained throughout the Delphi process. Participants were required to respond to the prior round of the Delphi survey to be invited to the subsequent stage.

During Phase 2, an online, modified Delphi consensus process was conducted via REDCap.²⁴ The purpose of this process was to (1) reach a consensus on the data elements to be included in the ICPGC CDEs, and (2) attribute each data element to the appropriate hierarchical level. These levels were (1) 'core': essential information required for all genetic studies of CP; (2) 'recommended': important information for most genetic studies of CP; (3) 'exploratory': somewhat important, but unlikely to be required in most genetic studies of CP; or (4) 'not required': not important for genetic studies in CP and would be removed from the list.

During round 1 (Appendix S2), respondents indicated which hierarchical level the data elements should be assigned to and were able to propose new data elements that had not been included or modifications to the existing data elements.

During round 2 (Appendix S3) and round 3 (Appendix S4) of the Delphi process, participants could see the results of all data elements that achieved consensus in the prior round. For data elements that did not reach consensus, participants could view the responses as a percentage of each hierarchical level (core, recommended, exploratory, not required) from the previous round. The results from the previous round were embedded within the survey to provide context for each data element and to highlight whether the data element was new, had reached consensus, or was yet to reach consensus.

After the third round of the Delphi survey, any data elements that did not reach consensus were reviewed and assessed independently by two of the members of the steering committee (YW and SMc). Disagreements were discussed and resolved by either consensus or consultation with a third reviewer from the steering committee (HSS). The data elements were scored against two criteria initially posited by the ICPGC-PWG as key requirements in the development of these CDEs: (1) is the data element important for genomic studies of CP; and (2) does the data element have an internationally standardized collection method, reporting variables or ontologies. A 'yes' response was given a value of 1, and a 'no' response was given a value of 0. Data elements were assigned to core if they had a score of 2, to recommended if they had a score of 1, and to exploratory if they had a score of 0.

Phase 3: define the MDS

During this phase, a team from the ICPGC-PWG was established (HSS, NB, SMC, MF, MCK, YW) to draft the preliminary MDS from the CDEs designated as core during the Delphi process. The proposed MDS was sent to the ICPGC-PWG (Appendices S5 and S6) via a REDCap survey to achieve agreement from the Working Group on what data elements should be 'mandatory'. Agreement was considered reached if 75% or more respondents indicated that the data element should be mandatory. If the data element had less than 75% support, it would remain as 'core'. During this survey, respondents could make additional suggestions for other data elements to be included in the MDS. These data elements had to come from the core data element, as identified during the Delphi process. Any suggestions received during the first round were reviewed and redistributed to the ICPGC-PWG.

Statistical analysis

Descriptive statistics were used to describe participants' demographics and survey responses. Data elements were considered to have reached consensus when more than 75% of Delphi respondents agreed that a data element should be 'core', 'recommended', 'exploratory', or 'not required'. This level of agreement has been considered appropriate in other Delphi studies.²⁵ Analyses were performed using IBM SPSS software, version 24 (IBM Corp., Armonk, NY, USA).

Ethics

The project had approval from the University of Sydney Human Research and Ethics Committee, a National Health and Medical Research Council accredited Human Research and Ethics Committee (2020/468).

RESULTS

Phase 1: identification and prioritization of a candidate set of CDEs

Data elements and variables were extracted from relevant data dictionaries and published CP genetic studies. The number of data elements extracted from each resource varied from 10 to the thousands. There was significant diversity in the data elements collected and reported in the literature. Only two data elements were consistently common among all resources – sex and age – albeit with some variability in their definitions and value domains. Several of the resources included data elements that collected country- or organization-specific variables that were not suitable for inclusion in an international data collection tool. Where possible, these data elements were given a broader range of values, more representative of the international cohorts (i.e. ethnicity).

For review in the modified Delphi survey, a final list of 82 data elements was organized into four categories: participant information, CP-specific assessments, family history, and pregnancy and birth history.

Phase 2: modified Delphi consensus process to select CDEs

Forty-six respondents participated in round 1 of the Delphi: 37 of the 78 invited panellists participated (47.4% response rate), and nine members from Australian Academy of Cerebral Palsy and Developmental Medicine. Of the 46 participants, 27 were female (58.7%) and 19 were male (41.3%); they included physicians ($n=19$, 41.3%), researchers/scientists ($n=17$, 37%), allied health practitioners ($n=8$, 17.4%), and two consumers (4.3%).

Table S1 shows the summary results of all data elements from the modified Delphi survey.

During round 1, 18 of the 82 data elements (22.2%) reached consensus. These included 17 core and one recommended data element. Additionally, 65 suggestions or modifications were made, which resulted in 20 data elements being added to the data elements list and reorganized into six categories: demographics, diagnostics, clinical traits, CP-specific assessments, family history, and antenatal and neonatal details.

Round 2 was distributed to the 46 participants who completed round 1. Thirty-three individuals participated (72% retention rate). Thirty-two (38.1%) of the 84 data elements reached consensus during this round, 20 core and 12 recommended.

Round 3 was sent out to the 33 participants who completed round 2. Twenty-seven completed the Delphi (82% retention rate). Twenty-five (48.1%) of the 52 data elements reached consensus, 10 core and 15 recommended. Twenty-eight data elements did not reach a consensus during the Delphi process and were reviewed by the steering committee. Upon reviewing the raw scores for the data elements, no data element received more than 15% agreement as not required. Therefore, the steering committee included all the data elements in the review. Of the 28 residual data elements, one was assigned to core, 14 to recommended, and 13 to exploratory (Table S1).

Phase 3: define the MDS

The MDS team reviewed the 52 core data elements, and a preliminary MDS of 10 data elements was formed. During this phase, the team introduced two new data elements and modified one existing data element in an effort to operationalize the consensus definition of CP.¹ This MDS was sent to the ICPGC-PWG and Governance Council ($n=23$ members) for review and consensus on the inclusion or exclusion of the data elements. Seventeen individuals responded (73.9% eligibility rate). Of the 10 data elements, eight had more than 75% of respondents agree that it should be considered mandatory and part of the MDS. Four additional data elements were recommended for the possible inclusion in the MDS by

TABLE 1 ICPGC CDE's minimum data set

Minimum data set	Data value domains
What is the individuals' year of birth?	YYYY
What is the sex of the individual?	1, Male 2, Female 3, Intersex 99, Unknown
What country was the individual born in?	International Organization for Standardization-3166 2-alpha code
Does the individual have a permanent (non-paroxysmal) movement disorder?	0, No 1, Yes
Is the individual's clinical course degenerative?	0, No 1, Yes
Please list positive Human Phenotype Ontology traits (minimum 3)	e.g. HP:0100277
What is the predominant motor type of the individual's CP?	1, Spastic 2, Dyskinetic - dystonia 3, Dyskinetic - choreoathetosis 4, Ataxic 5, Hypotonic ^a
What is the laterality of the individual's predominant motor type?	1, Unilateral 2, Bilateral
GMFCS level	1, GMFCS level I 2, GMFCS level II 3, GMFCS level III 4, GMFCS level IV 5, GMFCS level V 99, Unknown
Does the individual have epilepsy?	0, No 1, Yes

ICPGC, International Cerebral Palsy Genomics Consortium; CDE, common data element; GMFCS, Gross Motor Functioning Classification System.

^aNot all CP classification schemes recognize a hypotonic type.

the respondents. Of these four, two data elements had more than 75% agreement from the respondents, making the final MDS a total of 10 data elements (Table 1).

The ICPGC CDEs for genomics studies in CP

The ICPGC CDEs include 107 data elements, of which 10 comprise the MDS, the remainder are split across six categories: demographics (two core, four recommended); diagnostics (four core, two recommended, one exploratory); family history (five core, 10 recommended); antenatal and neonatal details (16 core, 15 recommended, eight exploratory); clinical traits (14 core, one recommended, four exploratory); and CP-specific assessments (one core, 10 recommended) (Table S2).

Where possible, data value domain options were generally based on existing standards. The CDEs are comprised of 74 enumerated and 33 non-enumerated data value domains (Table S2). The non-enumerated values include 16 measurement values (including both integer and numeric options, where all units are metric only), one date value (Gregorian calendar year), and 16 descriptive free text options. Of the enumerated

data values 28 are 'yes/no' values, and the remaining 46 have a defined list of permissible values, of which 32 use values from established functional classification systems (i.e. the Gross Motor Function Classification System), disease classifications (i.e. the International Classification of Diseases), ontologies (i.e. Human Phenotype Ontology), or other global standards (i.e. International Organization for Standardization).

DISCUSSION

Through a review of the literature, existing data dictionaries, and a modified Delphi method, a final set of CDEs for routine collection of data in genomics studies of CP have been developed. The CDEs recommended here will help harmonize phenotype data collected across clinical and research centres worldwide, enable de-identified data sharing, and minimize the need for data transformation. This study represents the first step in the standardization of phenotypic data for genomics studies in CP.

As these CDEs were explicitly designed for phenotypic data in genomic studies of CP, data elements pertaining to phenotype, family history, and aetiology were vital. CDEs deliver more value when aligned with accepted data standards and terminologies.²⁶ Therefore, where possible, we adopted data elements from relevant pre-existing CP data sets including CP registries^{10,11} and the NINDS CDEs for CP,⁹ to ensure interoperability with these resources. We also included established classification systems (i.e. Gross Motor Function Classification System and the International Classification of Diseases) and global standards (i.e. International Organization for Standardization), where possible.

Similar to the NINDS CDEs for CP, our CDEs include a structured hierarchical approach. Interestingly, most of the 'mandatory' and 'core' data elements are focused first and foremost on phenotype and family history. Whereas the 'recommended' and 'exploratory' data elements are focused on aetiology or environmental factors. These findings highlight the current state of CP genomics research, whereby genotype-phenotype studies are more predominant than genotype-environment studies – which is likely to be the next frontier.

Concise phenotypic data is essential for driving data sharing and is fundamental to the clinical interpretation of genetic data. Deep phenotyping requires machine-readable, granular phenotypic information from which trends can be inferred. The Human Phenotype Ontology²⁶ allows for deep phenotyping of patients and disease characteristics that are both human and machine-readable, making it incredibly amenable to computational phenotype analyses, which are paramount to the success of genomics in CP. The Human Phenotype Ontology has become the standard for reporting phenotypes in genomics projects by the Undiagnosed Diseases Program and Network, Genomics England, the Database of Genomics Variation and Phenotype in Humans using Ensemble Resources, and RD-Connect, and is available in several languages.^{27,28} The inclusion of the Human Phenotype Ontology as a mandatory data element is paramount to the interoperability with these other projects and many others.

These CDEs are designed to be incorporated into all CP genomics studies. However, they should not be considered as a finite set of data elements that must be collected. Instead, they should be used as the foundation one can build on, depending on the research question. As comprehensive as these CDEs are, some research teams will need to collect additional data points not currently included in the ICPGC CDEs. Researchers investigating the genome's role in interindividual differences in treatment outcomes (e.g. Diaz Heijtz et al.²⁹) may wish to use a combination of the ICPGC CDEs and the NINDS CP CDEs.

Strengths and limitations

The use of the Delphi process is a strength of this study. The Delphi methodology is a proven method to achieve consensus through an iterative process.¹² Furthermore, our panel comprised 46 individuals from around the world, including physicians, researchers/scientists, allied health practitioners, and consumers. There are currently no CDEs for CP that focus on aetiology, so these CDEs complement the existing NINDS CDEs for CP. In addition, these CDEs may also serve as a foundation for other (non-genetic) CP aetiological research studies upon which to build. The MDS aims to promote data collection while minimizing the administrative burden for researchers and health care providers.

The current study has some limitations. We could not achieve consensus on all the data elements, which meant that the steering committee assigned some hierarchy levels. In addition, these CDEs do not include any data elements dedicated to the many interventions, therapies, and treatments that an individual with CP may access. As intervention studies increase in frequency, recommendations for harmonizing these data will become necessary.

Future

These data elements are intended to be a dynamic first-step to harmonizing clinical data in genomics studies of CP. The next version of these CDEs will focus on comprehensive definitions, semantic annotations, and continued integration with existing global standards.²⁶ To expand the interoperability of these CDEs with other disorders with shared phenotypes that are frequently comorbid with CP, we aim to work with subject matter experts and data managers of these other disorders. We look forward to feedback and recommendations from the ongoing use and implementation of these CDEs. This discussion may lead to new data elements to be added (e.g. the new Visual Function Classification System³⁰), highlight better ways to capture critical clinical characteristics, or identify possible barriers to implementation. As a result of this project, we have now established a methodology for how the ICPGC-PWG will continue to refine, modify, and amend these CDEs into the future.

CONCLUSION

These CDEs offer the first step towards standardization and harmonization of phenotype data in genetic studies of CP to facilitate international data sharing, collaboration, and improved clinical interpretation of findings. For the most up-to-date version of the ICPGC CDEs, please refer to our website, [ICPGC.org](https://www.icpgc.org).

ACKNOWLEDGMENTS

Members of the ICPGC Phenotype Working Group are as follows: David Amor, Catherine Arnaud, Gareth Baynam, Hilla Ben-Pazi, Jesia Berry, Michael Fahey, Darcy Fehlings, Kate Himmelmann, Michael Kruer, Manju Kurian, Richard Leventer, Brian Marshall, Sarah McIntyre, Pamela Ng, Bethany Norton, Maryam Oskoui, Hayley Smithers-Sheedy, Siddharth Srivastava, Xiaoyang Wang, Linda Watson, Yana Wilson, Changlian Zhu.

The following people participated in at least one round and agreed to be identified: David Amor, Nadia Badawi, Gareth Baynam, Jesia Berry, Marie Blackmore, Jimmy Breen, Mark Corbett, Jeffrey Craig, Darcy Fehlings, Catherine Gibson, Prue Gollard, Anna Gubbay, Brandon Guida, Mikko Hallman, Kelly Harper, Kate Himmelmann, Ingrid Honan, Hao Hu, Carolyn Hunt, Anne Kawamura, Alison Kirkman, Adam Kirton, Michael Kruer, Sara Lewis, Alastair MacLennan, H. Thilini Madushika, Brian Marshall, Bethany Norton, Maryam Oskoui, Luis Perez-Jurado, Dinah Reddihough, Heather Scott, Michael Shevell, Siddharth Srivastava, Kirsty Stewart, Yusuke Takezawa, Anna te Velde, Marelle Thornton, Vivienne Travlos, Clare van Eyk, David Winlaw, Richard Wintle, Ellen Wood. The authors are grateful to all the Delphi members who participated in the study. Their participation does not imply that they agree with the results and conclusions from the study.


MF was supported by the Fulbright Foundation. MCK serves as a consultant for Neurocrine, Merz, Aeglea, and PTC Therapeutics, and receives grant support from PTC Therapeutics and Medtronic. The authors have stated that they had no interests that might be perceived as posing a conflict or bias. Open access publishing facilitated by The University of Sydney, as part of the Wiley -The University of Sydney agreement via the Council of Australian University Librarians. [Correction added on 14 May 2022, after first online publication: CAUL funding statement has been added.


DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Yana A. Wilson  <https://orcid.org/0000-0002-6878-1806>

Hayley Smithers-Sheedy  <https://orcid.org/0000-0002-0082-2413>

Katarina Ostojic  <https://orcid.org/0000-0001-6436-4936>

Emma Waight  <https://orcid.org/0000-0002-8027-6183>

Michael C. Krueer  <https://orcid.org/0000-0002-1373-7891>

Michael C. Fahey  <https://orcid.org/0000-0003-0811-4785>

Jozef Gécz  <https://orcid.org/0000-0002-7884-6861>

Sarah McIntyre  <https://orcid.org/0000-0002-0234-1541>

REFERENCES

- Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl* 2007;109:8–14.
- Himmelmann K, Ahlin K, Jacobson B, Cans C, Thorsen P. Risk factors for cerebral palsy in children born at term. *Acta Obstet Gynecol Scand* 2011;90(10):1070–81.
- McIntyre S, Taitz D, Keogh J, Goldsmith S, Badawi N, Blair E. A systematic review of risk factors for cerebral palsy in children born at term in developed countries. *Dev Med Child Neurol* 2013;55(6):499–508.
- Fahey MC, MacLennan AH, Kretzschmar D, Gecz J, Krueer MC. The genetic basis of cerebral palsy. *Dev Med Child Neurol* 2017;59(5):462–469.
- MacLennan AH, Krueer MC, Baynam G, Moreno-de-Luca A, Wilson YA, Zhu C et al. Cerebral palsy and genomics: an international consortium. *Dev Med Child Neurol* 2018;60(2):209–210.
- Petrovski S, Goldstein DB. Phenomics and the interpretation of personal genomes. *Sci Transl Med* 2014;6(254):254fs35.
- Sheehan J, Hirschfeld S, Foster E, Ghitzu U, Goetz K, Karpinski J et al. Improving the value of clinical research through the use of Common Data Elements. *Clin Trials* 2016;13(6):671–676.
- Grinnon ST, Miller K, Marler JR, Lu Y, Stout A, Odenkirchen J et al. National Institute of Neurological Disorders and Stroke Common Data Element Project - approach and methods. *Clin Trials* 2012;9(3):322–9.
- Schiarit V, Fowler E, Brandenburg JE, Levey E, McIntyre S, Sukal-Moulton T, et al. A common data language for clinical research studies: the National Institute of Neurological Disorders and Stroke and American Academy for Cerebral Palsy and Developmental Medicine Cerebral Palsy Common Data Elements Version 1.0 recommendations. *Dev Med Child Neurol* 2018;60(10):976–986.
- Australian Cerebral Palsy Register. Report of the Australian Cerebral Palsy Register Birth Years 1995–2012. Cerebral Palsy Alliance, 2018;86–88. Available from: <https://cpregister.com/wp-content/uploads/2019/02/Report-of-the-Australian-Cerebral-Palsy-Register-Birth-Years-1995-2012.pdf>
- Surveillance of Cerebral Palsy in Europe. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Dev Med Child Neurol* 2000;42:816–824.
- Hsu CC, Sandford BA. The Delphi Technique: Making Sense of Consensus. *Prac Assess Res Eval* 2007;12(1):10.
- Niederberger M, Spranger J. Delphi Technique in Health Sciences: A Map. *Front Public Health* 2020;8:457.
- Barber CEH, Mosher DP, Ahluwalia V, Zimmer M, Marshall DA, Choquette D et al. Development of a Canadian Core Clinical Dataset to Support High-quality Care for Canadian Patients with Rheumatoid Arthritis. *J Rheumatol* 2017;44(12):1813–1822.
- Davey CJ, Slade SV, Shickle D. A proposed minimum data set for international primary care optometry: a modified Delphi study. *Ophthalmic Physiol Opt* 2017; 37(4): p. 428–439.
- Stone E, Rankin N, Phillips J, Fong K, Currow DC, Miller A, et al. Consensus minimum data set for lung cancer multidisciplinary teams: Results of a Delphi process. *Respirology* 2018;23(10):927–934.
- Ward SL, Flori HR, Bennet TD, Sapru A, Mourani PM, Thomas NJ, et al. Design and Rationale for Common Data Elements for Clinical Research in Pediatric Critical Care Medicine. *Pediatr Crit Care Med* 2020;21(11):e1038–e1041.
- Junger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DELphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. *Palliat Med* 2017;31(8):684–706.

- The Robinson Institute, University of Adelaide. The Australian Collaborative Cerebral Palsy Research Group [Internet]. 1970; Available from: <https://www.adelaide.edu.au/robinson-research-institute/research/research-groups/cerebral-palsy>.
- CanChild, McMaster University. About CP-NET. [Internet] 2021; Available from: <http://cpnet.canchild.ca/en/about-cp-net>.
- Australian and New Zealand Neonatal Network. ANZNN 2017 Data Dictionary. 2016: Sydney, Australia.
- Taruscio D, Mollo E, Gainotti S, Posada de la Paz M, Bianchi F, Vittozzi L. The EPIRARE proposal of a set of indicators and common data elements for the European platform for rare disease registration. *Arch Pub Health* 2014;72:35–42.
- Rubinstein YR, McInnes P. NIH/NCATS/GRDR(R) Common Data Elements: A leading force for standardized data collection. *Contemp Clin Trials* 2015;42:78–80.
- Harris PA, Taylor R, Minor BL, Elliot V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software partners. *J Biomed Inform* 2019;95:103208.
- Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 2014;67(4):401–9.
- Kush RD, Warzel D, Kush MA, Sherman A, Navarro EA, Fitzmartin R et al. FAIR data sharing: The roles of common data elements and harmonization. *J Biomed Inform* 2020;107:103421.
- Kohler S, Gargano M, Matentzoglou N, Carmody LC, Lewis-Smith D, Vasilevsky NA, et al. The Human Phenotype Ontology in 2021. *Nucleic Acids Res* 2021;49(D1):D1207–D1217.
- Kohler S, Carmody LC, Vasilevsky NA, Jacobsen JOB, Danis D, Gourdine J-P, Gargano M, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res* 2019;47(D1):D1018–D1027.
- Diaz Heijtz R, Almeida R, Eliasson AC, Forsberg H. Genetic Variation in the Dopamine System Influences Intervention Outcome in Children with Cerebral Palsy. *EBioMedicine* 2018;28:162–167.
- Baranello G, Signorini S, Tinelli F, Guzzetta A, Pagliano E, Rossi A, et al. Visual Function Classification System for children with cerebral palsy: development and validation. *Dev Med Child Neurol* 2020;62(1):104–110.

SUPPORTING INFORMATION

The following additional material may be found online:

Figure S1: Study flowchart.

Appendix S1: Emails to Delphi participants.

Appendix S2: Modified Delphi round 1 questionnaire.

Appendix S3: Modified Delphi round 2 questionnaire.

Appendix S4: Modified Delphi round 3 questionnaire.

Appendix S5: Correspondence.

Appendix S6: Mandatory data elements for the CP Commons.

Table S1: Summary of data element ranking during the three Delphi rounds.

Table S2: The complete ICPGC CDEs for genomics studies of CP (v1).

How to cite this article: Wilson YA, Smithers-Sheedy H, Ostojic K, Waight E, Krueer MC, Fahey MC, ICPGC Phenotype Working Group. Common data elements to standardize genomics studies in cerebral palsy. *Dev Med Child Neurol*. 2022;64(12):1470–1476. <https://doi.org/10.1111/dmcm.15245>

Elementos de dados comuns para padronizar estudos genômicos em paralisia cerebral

Objetivo

Definir elementos de dados clínicos comuns (DCC) e um conjunto mínimo de dados obrigatórios (CMDO) para estudos genômicos de paralisia cerebral (PC).

Método

Os elementos de dados do candidato foram coletados seguindo uma revisão da literatura e através dos DCC existentes. Uma pesquisa on-line de três rodadas Delphi foi usada para classificar cada elemento de dados como 'essencial', 'recomendado', 'exploratório' ou 'não obrigatório'. Os Membros do Consorcio Internacional de Genoma na Paralisia Cerebral (MCIGPC) classificaram os DCC do núcleo como obrigatórios ou não, para formar o CMDO. Tanto para os DCC quanto para o CMDO, um elemento de dados foi considerado como tendo chegado a um consenso se mais de 75% dos respondentes concordassem.

Resultados

Quarenta e seis indivíduos de todo o mundo formaram o painel Delphi: consumidores (n=2), cientistas/pesquisadores (n=17), médicos (n=19) e profissionais de saúde aliados (n=8). Os DCC incluem 107 elementos de dados em seis categorias: demografia, diagnóstico, história familiar, detalhes pré-natais e neonatais, características clínicas e avaliações específicas de PC. Destes, 10 são obrigatórios, 42 essenciais, 41 recomendados e 14 são exploratórios

Interpretação

Os DCC do MCIGPC fornecem uma base para a padronização de dados de fenótipo capturados em estudos genômicos de PC e beneficiarão colaborações internacionais e agrupamento de dados, particularmente em condições raras.