The Journal of Child Psychology and Psychiatry

Journal of Child Psychology and Psychiatry 65:7 (2024), pp 991–994 doi:10.1111/jcpp.13974

Editorial Perspective: The paradox of precision health in early development – building large samples to yield individual-level measures

Mark H. Johnson,[1,2] Pasco Fearon,[1] Andrew Pickles,[3] and Emily J.H. Jones[2]

[1]Department of Psychology, University of Cambridge, Cambridge, UK; [2]Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK; [3]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

Precision health refers to the use of individualised biomarkers within predictive models to provide more tailored information about an individual's likely prognosis, providing greater opportunities for planning the supports or interventions most likely to have a favourable risk/benefit ratio for that individual (Schüssler-Fiorenza Rose et al., 2019). Although most existing applications of precision health are in adult medical settings, this approach could have greatest potential in early development. Predicting developmental trajectories from early in life is important for planning the services, new opportunities or adjustments a child may need as they grow and develop. Identifying children with an elevated chance of developing specific later developmental or mental health difficulties also presents opportunities for early intervention during sensitive periods across educational, clinical and quality of life outcomes. In early childhood, the complex and heterogeneous nature of neurodevelopmental processes, along with the unclear mapping onto distinct developmental outcomes (Gillberg, 2010; Sonuga-Barke, 2016), creates further motivation for characterising individual trajectories. Thus, deploying resources towards generating individual-level prediction models for use in early childhood could provide significant societal and individual benefit, particularly if an individual child's response to treatment can also be predicted. In order to achieve these goals, we argue that (a) there is a dire need to further develop neurocognitive measures as predictors of later behavioural outcomes and (b) these neurocognitive studies need to be conducted at large scale. We then outline two strategies for developing such cohorts; facilitating links between existing cohort studies, and establishing new national cohort studies that incorporate neurocognitive measures.

Why do we need to focus on neuro-cognitive processes? Collecting and assimilating large-scale developmental data is nontrivial, and usually involves a trade-off with the depth of individual profiling. For this reason, large-scale population samples in human neurodevelopmental science are usually assessed with questionnaires with some limited additional behavioural testing in older age groups (Boyd et al., 2013; Connelly & Platt, 2014). However, prospective studies in populations enriched for varied developmental outcomes (such as infants with a family history of autism) suggest that the earliest predictors of later developmental status may be observed in neurocognitive assessments rather than parent reports of child behaviour (e.g., Hazlett et al., 2017; Johnson, Charman, Pickles, & Jones, 2021; Piven, Elison, & Zylka, 2017). Over the first three years, infants also have a relatively limited behavioural repertoire, and so the effects of genetic or other risk processes on early brain development are most evident with recording modalities that do not depend on inference from observable differences in behaviour (Gui et al., 2021). Further, what cannot be discerned by mere observation are the underlying neurobiological, learning and information processing mechanisms causing behaviour and its development over time. These processes underlying behaviour are likely closer to proximal causes of atypical trajectories and thus are potentially better predictors of later outcome (Johnson, 2015). Finally, the rapid pace of brain development in the first years of life raises significant scientific and societal questions concerning the impact of varied environments on early brain health, which are likely best addressed by directly measuring the brain itself. A range of methods for assessing brain and cognitive development are now widely available with appropriate modifications to make them robust and reliable for use in early development and in home settings, including electroencephalography (EEG), functional near infrared spectroscopy (fNIRS) and eyetracking. Given recent developments in making these measurement methods more potentially suitable for large samples, and data science tools required for large-scale analysis, we argue that the time is right for amassing the sample sizes required to test the utility of neurocognitive markers for tracking the brain development of individuals in order to predict meaningful outcomes.

Why do we need large sample sizes? Precision health requires information that is robust and predictive at the individual level. However, the more the focus is on the individual participant, the larger the sample size required to properly characterise

Conflict of interest statement: No conflicts declared.

population variation and an individual's precise location within it. The importance of large samples sizes has recently been demonstrated in a number of fields. For example, within genetics, a wave of replication failures for candidate genes led to a shift towards genome-wide association studies, which require sample sizes in the tens or hundreds of thousands to yield robust results (Sullivan et al., 2018). It is now abundantly clear that data-driven strategies and large samples were essential not only in yielding robust and reproducible genetic associations but also in generating increasingly predictive polygenic scores and related prediction models. Further, recent simulation approaches have indicated that nontheory driven brain-wide association studies (BWAS) similarly require sample sizes of thousands to be replicable (Marek et al., 2022), though more targeted methods may be robust in smaller datasets (Gratton, Nelson, & Gordon, 2022). To avoid further contributing to disparities in healthcare access, datasets also need to explore factors such as ethnicity and socio-economic status. Thus, to yield more rapid progress in precisely identifying robust predictors of later development from infancy, we will require large-scale representative datasets. Importantly, these large-scale studies will still need to be interspersed with smaller-scale hypothesis testing studies with some methodological plurality in the field (Sonuga-Barke, 2022).

To address these issues we recommend two strategies going forward. First, we should include child neurocognitive measures within new national cohort studies. A few medium-scale studies have begun to systematically map EEG and eyetracking markers in hundreds or the low thousands of children, including the Youth cohort in the Netherlands (Onland-Moret et al., 2020) and ongoing Wellcome-LEAP and Gates foundation collaborations. These efforts show that large-scale neurocognitive measurement is feasible. However, these studies remain comparatively modest in size and are often built on convenience sampling. This limits their generalisability and their ability to generate population-level insights, since participating families are unlikely to be fully representative of the broader community. Integrating neurocognitive measurements into cohorts with representative stratified sampling is a critical next step to identifying robust predictors of developmental outcomes, in addition to robustly testing key questions about the effects of environmental variation on early brain development. To this end, the field needs to develop cost-effective experimental platforms that allow for standardised collection of neurocognitive measures across multiple sites and by field workers with comparatively limited training so that data collection can be made nationally scalable. Many standardised behavioural measures are available that have been successful deployed within multiple large-scale studies. However, this approach has been much less common for neurocognitive measurements, where investigators tend to use bespoke tasks, scripting frameworks and analytic pipelines. This heterogeneity allows greater exploration of the space of potential predictors but limits scope for replication and generalisation. One example of a framework designed to enable cross site standardisation is TaskEngine, a stimulus presentation framework that has been used to standardise eye tracking and EEG data collection across multiple European sites (Jones et al., 2019). Another example is the NIH toolbox (https://www.nia.nih.gov/research/resource/nih-toolbox) that provides standardised cognitive assessments from 3 to 85 years old; extending this approach to infancy through a method such as eyetracking might be fruitful, as this would potentially allow for task demands to be equalised across ages. While integration of neurocognitive measures into population samples requires methodological innovation, it will ultimately provide the database necessary for a population neuroscience within which individuals can be characterised, resulting in a step-change in our ability to identify robust causal mechanisms of neurobehavioural development and, ultimately, to predict future outcomes.

The second, and complementary, strategy is to synergise data across currently funded cohort studies in order to further increase their scientific value in terms of power to detect effects, breadth of research questions and linkage to genetic data. There are several large-scale longitudinal cohort studies in the UK alone, each of which is a significant investment that requires investigators to obtain the best scientific value for money. Efforts to link together these cohorts has many potential advantages. Through informal scientific sharing these cohorts share a proportion of similar measures, although these are not consistent and often developed independently. Leveraging this comparability could allow us to build large sample sizes for some measures, whilst providing a reference point for integration of more diverse measures. Several projects are currently underway to build harmonisation of measurement across existing cohorts (see e.g., the work of CLOSER; McElroy et al., 2020) and to strengthen the degree of harmonisation across future cohorts. Further, whilst planned largescale cohorts tend to focus on particular relatively broadly spaced age points, linking cohorts that have sampled different age points with common measures potentially allows for more dense sampling of age points, allowing for better overall assessment of developmental trajectories. Linking cohorts that have shared methodology and measures further allows us to compare across different samples enriched for diversity that may be the focus of particular research programmes, such as with a family history of autism or ADHD and premature infants; or groups of infants growing up in more diverse environments than could

be captured in a single study. Such work facilitates the transdiagnostic approaches that have gained increasing traction (Astle, Holmes, Kievit, & Gathercole, 2022). Similarly, if large representative cohorts used parallel measures to those obtained in studies of discrete exposures then pooled analyses could be undertaken with greater potential for causal inference (e.g., using synthetic controls or propensity score matching). Finally, linking diverse cohorts can provide natural opportunities for replication and external validation of findings, important to generating robust evidence for societal application.

Whilst there are many cost-effective benefits to linking existing longitudinal cohort studies, there are challenges that mean that it is not often pursued. To demonstrate the potential value of this approach and to make the case for future investment, we should survey the assessments and measures of current cohorts and identify existing common measures, or those that could easily be inserted at low cost and effort. Similar efforts are already underway in the field of child and adolescent mental health (https://www.cataloguementalhealth.ac.uk). Bringing together data from these measures could yield early signals, and help to show the potential value of synergising other data types. However, although the presence of common measures is powerful, requiring standardisation of tasks risks freezing development of measures at a point that may not reflect a mature understanding of the optimal way to measure neurocognitive development. There may thus be significant value in integrating data from measures of a similar construct administered slightly differently, which will require statistical methods for cross-calibration. One related approach is to assemble a shared data resource that maps the multiverse of experimental design options more systematically (Almaatouq et al., 2022; Dafflon et al., 2022). Exploiting the relationship between different experimental measures within calibration or reference samples can help to integrate insights from different cohorts that have deployed distinct measures, incorporating these data directly into the analysis. In so doing we recognise our uncertainty as to how well the diverse measures are related (Collishaw, Maughan, Goodman, & Pickles, 2004). This approach has been developed further in the field of Individual Patient Data (IPD) meta-analysis, for example to consider variables entirely missing in one cohort (Quartagno & Carpenter, 2016). There remains, however, much yet to be done in understanding how these data integration solutions perform in prediction modelling.

Finally, we must consider where limited resources should be deployed. Funders are beginning to develop schemes for data mining that can leverage the rapid acceleration of data science tools available to the research community. However, it is rare for funders to seek to bridge existing cohorts to generate datasets suitable for data mining in ways that maximise the value of all the original studies. The admininstrative and ethics challenges of data sharing means that federated approaches in which novel analyses are brought to existing data are necessary, and models of this are available in the mature field of genetic data sharing (e.g., the Autism Sharing Initiative). We recommend specific grant funding schemes for this purpose that can be applied to diverse data types.

In conclusion, we contend that the dual ambitions of integrating neurocognitive measurement into large-scale cohort studies and bringing together existing datasets are critical first steps to mapping the neurocognitive predictors of developmental outcomes across diverse populations in ways that are robust, reproducible and generalisable. Such efforts are also critical to addressing basic science questions concerning the effects of variation in genetic and environmental background on early brain development. Achieving this goal will require investment in scalable neurocognitive assessment methods, sharing and contributing to common data libraries, and in data science approaches that can maximise the value of existing datasets; if this can be achieved, it may allow us to realise the potential for precision health in early human neurodevelopment.

Acknowledgements

The authors are grateful to the Waterloo Foundation for funding the meeting on which this editorial is based and for discussion from those who attended the "Meta-Cohorts" meeting in London in 2022. Pasco Fearon is Deputy Editor-in-Chief for JCPP. Aside from this, the authors have no competing or potential conflicts of interest in relation to this editorial.

Correspondence

Mark H. Johnson, Department of Psychology, University of Cambridge, Cambridge, UK; Email: mj492@cam.ac.uk

References

Almaatouq, A., Griffiths, T.L., Suchow, J.W., Whiting, M.E., Evans, J., & Watts, D.J. (2022). Beyond playing 20 questions with nature: Integrative experiment design in the social and behavioral sciences. *Behavioral and Brain Sciences*, *47*, e33.

Astle, D.E., Holmes, J., Kievit, R., & Gathercole, S.E. (2022). Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry*, *63*, 397–417.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., … & Davey Smith, G. (2013). Cohort Profile: The 'children of the 90s' – The index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, *42*, 111–127.

Collishaw, S., Maughan, B., Goodman, R., & Pickles, A. (2004). Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry*, *45*, 1350–1362.

Connelly, R., & Platt, L. (2014). Cohort Profile: UK millennium cohort study (MCS). *International Journal of Epidemiology*, *43*, 1719–1725.

Dafflon, J.F., Da Costa, P., Váša, F., Monti, R.P., Bzdok, D., Hellyer, P.J., … & Leech, R. (2022). A guided multiverse study of neuroimaging analyses. *Nature Communications*, *13*, Article 1.

Gillberg, C. (2010). The ESSENCE in child psychiatry: Early symptomatic syndromes eliciting neurodevelopmental clinical examinations. *Research in Developmental Disabilities*, *31*, 1543–1551.

Gratton, C., Nelson, S.M., & Gordon, E.M. (2022). Brain-behavior correlations: Two paths toward reliability. *Neuron*, *110*, 1446–1449.

Gui, A., Meaburn, E.L., Tye, C., Charman, T., Johnson, M.H., & Jones, E.J.H. (2021). Association of polygenic liability for autism with face-sensitive cortical responses from infancy. *JAMA Pediatrics*, *175*, 968–970.

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., … & Network, T.I.B.I.S. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, *542*, 348–351.

Johnson, M.H. (2015). Neurobiological perspectives on developmental psychopathology. In A. Thapar, D.S. Pines, J. Leckman, M.J. Snowling, S. Scott, & E. Taylor (Eds.), *Rutter's child and adolescent psychiatry* (6th edn, pp. 107–118). Oxford: John Wiley & Sons.

Johnson, M.H., Charman, T., Pickles, A., & Jones, E.J.H. (2021). Annual Research Review: Anterior modifiers in the emergence of neurodevelopmental disorders (AMEND)—A systems neuroscience approach to common developmental disorders. *Journal of Child Psychology and Psychiatry*, *62*, 610–630.

Jones, E.J.H., Mason, L., Begum Ali, J., van den Boomen, C., Braukmann, R., Cauvet, E., … & Johnson, M.H. (2019). Eurosibs: Towards robust measurement of infant neurocognitive predictors of autism across Europe. *Infant Behavior and Development*, *57*, 101316.

Marek, S., Tervo-Clemmens, B., Calabro, F.J., Montez, D.F., Kay, B.P., Hatoum, A.S., … & Dosenbach, N.U.F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*, 654–660.

McElroy, I., Villadsen, A., Patalay, P., Goodman, A., Richards, M., Northstone, K., … & Ploubidis, G. (2020). *Harmonisation and measurement properties of mental health measures in six British cohorts*. London: CLOSER.

Onland-Moret, N.C., Buizer-Voskamp, J.E., Albers, M.E.W.A., Brouwer, R.M., Buimer, E.E.L., Hessels, R.S., … & Kemner, C. (2020). The YOUth study: Rationale, design, and study procedures. *Developmental Cognitive Neuroscience*, *46*, 100868.

Piven, J., Elison, J.T., & Zylka, M.J. (2017). Toward a conceptual framework for early brain and behavior development in autism. *Molecular Psychiatry*, *22*, Article 10.

Quartagno, M., & Carpenter, J.R. (2016). Multiple imputation for IPD meta-analysis: Allowing for heterogeneity and studies with missing covariates. *Statistics in Medicine*, *35*, 2938–2954.

Schüssler-Fiorenza Rose, S.M., Contrepois, K., Moneghetti, K.J., Zhou, W., Mishra, T., Mataraso, S., … & Snyder, M.P. (2019). A longitudinal big data approach for precision health. *Nature Medicine*, *25*, Article 5.

Sonuga-Barke, E.J.S. (2016). Distinguishing the challenges posed by surface and deep forms of heterogeneity to diagnostic symptoms: Do we need a new approach to subtyping of child and adolescent psychiatric disorders. *Journal of Child Psychology and Psychiatry*, *57*, 1–3.

Sonuga-Barke, E.J.S. (2022). Editorial: 'Safety in numbers'? Big data discovery strategies in neuro-developmental science – contributions and caveats. *Journal of Child Psychology and Psychiatry*, *64*(1), 1–3. https://doi.org/10.1111/jcpp.13723

Sullivan, P.F., Agrawal, A., Bulik, C.M., Andreassen, O.A., Børglum, A.D., Breen, G., … & Psychiatric Genomics Consortium. (2018). Psychiatric genomics: An update and an agenda. *The American Journal of Psychiatry*, *175*, 15–27.

Accepted for publication: 18 January 2024