

# A framework for neural organoids, assembloids and transplantation studies

<https://doi.org/10.1038/s41586-024-08487-6>

Received: 26 August 2024

Accepted: 4 December 2024

Published online: 9 December 2024



Sergiu P. Pasca<sup>1,2,✉</sup>, Paola Arlotta<sup>3,4</sup>, Helen S. Bateup<sup>5,6</sup>, J. Gray Camp<sup>7,8</sup>, Silvia Cappello<sup>9,10</sup>, Fred H. Gage<sup>11</sup>, Jürgen A. Knoblich<sup>12,13</sup>, Arnold R. Kriegstein<sup>14,15</sup>, Madeline A. Lancaster<sup>16,17</sup>, Guo-Li Ming<sup>18,19,20,21</sup>, Gaia Novarino<sup>22</sup>, Hideyuki Okano<sup>23</sup>, Malin Parmar<sup>24</sup>, In-Hyun Park<sup>25,26</sup>, Orly Reiner<sup>27,28</sup>, Hongjun Song<sup>18,19,20,29</sup>, Lorenz Studer<sup>30</sup>, Jun Takahashi<sup>31</sup>, Sally Temple<sup>32</sup>, Giuseppe Testa<sup>33,34</sup>, Barbara Treutlein<sup>35</sup>, Flora M. Vaccarino<sup>26,36,37,38</sup>, Pierre Vanderhaeghen<sup>39,40</sup> & Tracy Young-Pearse<sup>41,42</sup>

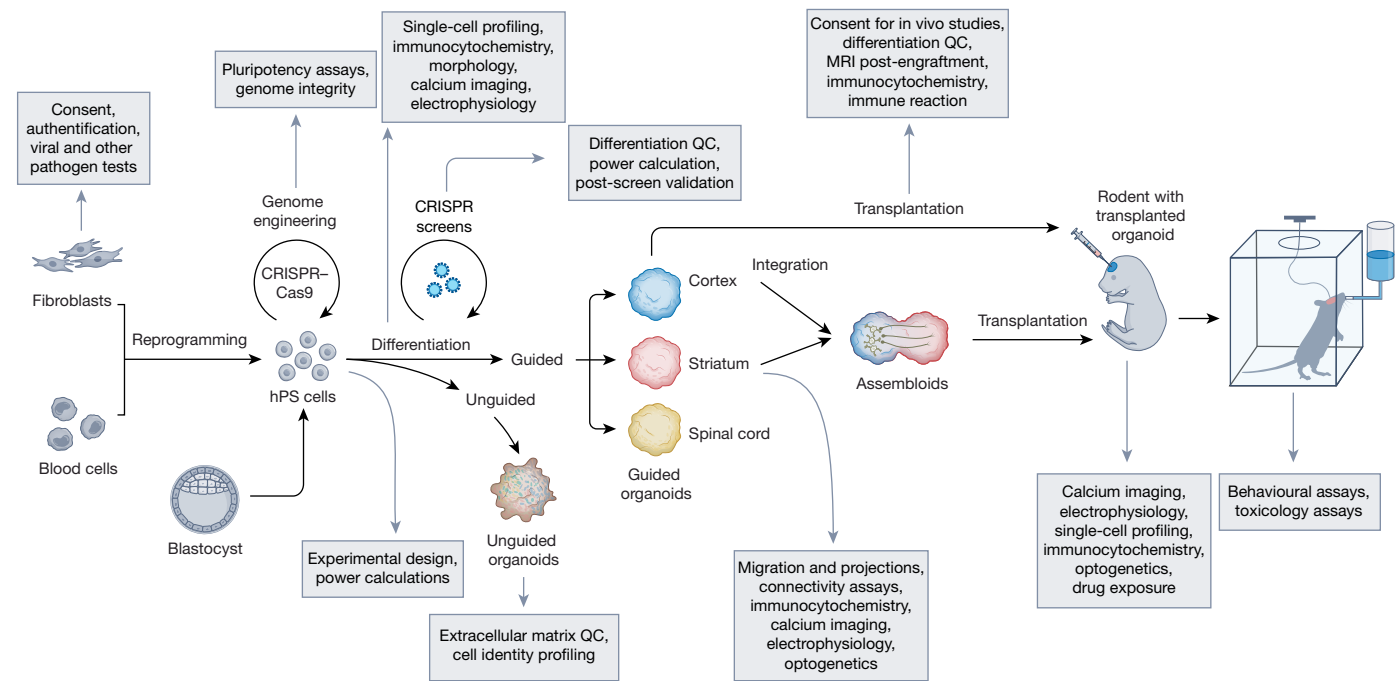
As the field of neural organoids and assembloids expands, there is an emergent need for guidance and advice on designing, conducting and reporting experiments to increase the reproducibility and utility of these models. In this Perspective, we present a framework for the experimental process that encompasses ensuring the quality and integrity of human pluripotent stem cells, characterizing and manipulating neural cells in vitro, transplantation techniques and considerations for modelling human development, evolution and disease. As with all scientific endeavours, we advocate for rigorous experimental designs tailored to explicit scientific questions as well as transparent methodologies and data sharing to provide useful knowledge for current research practices and for developing regulatory standards.

The field of stem cell-based modelling of human development, evolution and disease using organoids and assembloids has seen a tremendous surge in interest, with more than 3,000 articles published annually. Together with 2D cellular models, these 3D preparations, whether utilized in vitro or in vivo after transplantation into animals, leverage the inherent self-organization capabilities of stem and progenitor cells to mimic aspects of physiology and hold the potential to reveal new human tissue biology and pathophysiology<sup>1</sup>. Their application for neuroscience is particularly welcome as inaccessibility has limited understanding of the biology of the human nervous system and the mechanisms underlying neurodevelopmental and neurodegenerative disorders. However, it is becoming increasingly difficult to navigate the expanding literature and identify experimental standards for building reliable 3D microphysiological systems, especially as in vitro studies are often not reported with sufficient experimental detail<sup>2</sup>. Importantly, these complex multi-cellular systems are maintained in long-term cultures (often

for hundreds of days) and assessments require increasingly sophisticated assays that include cellular omics, functional assays and circuit probing in vitro and in animals. Consequently, unlike experimental systems with faster turnarounds and simpler readouts, adjustments based on successful or unsuccessful outcomes in a human cellular model are often delayed, which affects experimental progress, limiting reproducibility across laboratories. As many groups are implementing these multi-cellular systems for development, evolution or disease applications, and as funding, publishing and regulatory agencies become engaged, it is becoming increasingly important to outline critical experimental variables and challenges for this rapidly expanding field.

Following a similar consensus effort to clarify nomenclature for the field<sup>1</sup>, we gathered a group of international researchers to outline an experimental framework that can apply to neural organoids and assembloids and their xeno-transplantation to highlight some of the limitations and to delineate quality control measures and standards (Fig. 1).

<sup>1</sup>Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. <sup>2</sup>Stanford Brain Organogenesis, Wu Tsai Neurosciences Institute and Bio-X, Stanford University, Stanford, CA, USA. <sup>3</sup>Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. <sup>4</sup>Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. <sup>5</sup>Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA, USA. <sup>6</sup>Department of Neuroscience, University of California Berkeley, Berkeley, CA, USA. <sup>7</sup>Institute of Human Biology (IHB), Roche Pharma Research and Early Development, Roche Innovation Center Basel, Basel, Switzerland. <sup>8</sup>Biozentrum, University of Basel, Basel, Switzerland. <sup>9</sup>Department of Physiological Genomics, Biomedical Center (BMC), Faculty of Medicine, Ludwig Maximilian University of Munich, Munich, Germany. <sup>10</sup>Max Planck Institute of Psychiatry, Munich, Germany. <sup>11</sup>Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, CA, USA. <sup>12</sup>Institute of Molecular Biotechnology, Austrian Academy of Sciences, Vienna Biocenter, Vienna, Austria. <sup>13</sup>Department of Neurology, Medical University of Vienna, Vienna, Austria. <sup>14</sup>Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California San Francisco, San Francisco, CA, USA. <sup>15</sup>Department of Neurology, University of California San Francisco, San Francisco, CA, USA. <sup>16</sup>MRC Laboratory of Molecular Biology, Cambridge, UK. <sup>17</sup>Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. <sup>18</sup>Department of Neuroscience and Mahoney Institute for Neurosciences, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>19</sup>Department of Cell and Developmental Biology, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>20</sup>Institute for Regenerative Medicine, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>21</sup>Department of Psychiatry, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>22</sup>Institute of Science and Technology of Austria, Klosterneuburg, Austria. <sup>23</sup>Keio University Regenerative Medicine Research Center, Kanagawa, Japan. <sup>24</sup>Department of Experimental Medical Science, Lund Stem Cell Center, Lund University, Lund, Sweden. <sup>25</sup>Department of Genetics, Yale School of Medicine, New Haven, CT, USA. <sup>26</sup>Yale Stem Cell Center, Yale University, New Haven, CT, USA. <sup>27</sup>Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. <sup>28</sup>Department of Molecular Neuroscience, Weizmann Institute of Science, Rehovot, Israel. <sup>29</sup>The Epigenetics Institute, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA, USA. <sup>30</sup>The Center for Stem Cell Biology, Developmental Biology Program, Sloan Kettering Institute for Cancer Research, New York, NY, USA. <sup>31</sup>Department of Clinical Application, Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan. <sup>32</sup>Neural Stem Cell Institute, Albany, NY, USA. <sup>33</sup>Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy. <sup>34</sup>Human Technopole, Viale Rita Levi Montalcini, Milan, Italy. <sup>35</sup>Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland. <sup>36</sup>Child Study Center, Yale University, New Haven, CT, USA. <sup>37</sup>Department of Neuroscience, Yale University, New Haven, CT, USA. <sup>38</sup>Yale Kavli Institute for Neuroscience, New Haven, CT, USA. <sup>39</sup>VIB-KU Leuven Center for Brain and Disease Research, Leuven, Belgium. <sup>40</sup>Department of Neurosciences, Leuven Brain Institute, Leuven, Belgium. <sup>41</sup>Department of Neurology, Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. <sup>42</sup>Harvard Stem Cell Institute, Harvard University, Cambridge, MA, USA. ✉e-mail: [spasca@stanford.edu](mailto:spasca@stanford.edu)



**Fig. 1 | Differentiation of hPS cells into unguided or guided organoids, assembloids and their transplantation into animals.** Boxes indicate assays and other variables that should be taken in consideration when designing

experiments. MRI, magnetic resonance imaging; QC, quality control. Adapted with permission from Mesa Schumacher.

This consensus arises from our experience in modelling the development, evolution and diseases of the human nervous system. Although we encourage methods development, we emphasize that experiments with these models should be thoughtfully designed to answer a specific scientific question as knowledge of the variables can also mitigate the cost of experiments. In other words, considering the variety of experimental models available, including 2D models, with their advantages and caveats, it is critical to match the experimental system to the question. Moreover, we underscore the need for transparency in sharing protocols and experimental details, including how quantifications were performed and the requirements for depositing raw data in public databases.

Quality control measures can vary widely based on experimental, translational or clinical goals. However, the variables that we describe here should offer sufficient guidance for tailoring experimental designs to a wide array of objectives and needs. We hope that this summary, derived from our collective experience, will serve as guidance to those in or entering this field, especially trainees, and ultimately will lead to more optimally designed studies. Finally, we anticipate that this framework for the experimental process may have implications for regulatory agencies such as the US Food and Drug Administration, the UK Medicines and Healthcare products Regulatory Agency, the European Medicines Agency and the Pharmaceuticals and Medical Devices Agency of Japan that are relevant to developing therapies in an academic or industry setting, and that many of these guidelines will be relevant to both 2D cellular models and to 3D models of other organs.

### Quality of hPS cells as a foundational step

One of the most important considerations is the quality of the human pluripotent stem cell (hPS cell; encompassing human embryonic stem (hES) cells and human induced pluripotent stem (iPS) cells) lines used to derive neural and other cell lineages. These cells are prone to spontaneous differentiation and sensitive to cell culture conditions and growth factors to maintain pluripotency, and each iPS cell line can carry hundreds to thousands of single nucleotide mutations, copy

number variants and genomic or epigenomic changes. Some genomic variants can be acquired in culture even after relatively short periods (5–10 passages), whereas others reflect somatic mutations that are inherited from founder cells<sup>3–6</sup>. Even pluripotent stem cell lines that are considered isogenic may not always be strictly isogenic and reprogramming itself can be mutagenic<sup>4</sup>. Therefore, it is crucial to confirm genome integrity status at the onset of obtaining the cells and to then regularly verify and report this, as well as to authenticate pluripotent stem cell lines, as suggested by the International Society for Stem Cell Research (ISSCR) guidelines on stem cell research and modelling<sup>7</sup>. It is important to note that classic karyotyping methods or low-pass whole-genome sequencing (WGS) may lack the necessary resolution, and array-based DNA technologies such as array-comparative genomic hybridization (array CGH) and single nucleotide polymorphism (SNP) arrays are recommended. There is no consensus on a minimal number of de novo genetic events in a pluripotent stem cell line that warrants exclusion, and finding a genetic variant does not necessarily compromise the results, as long as they are described in detail in the manuscript and unless they affect genes that are known to interfere with cellular function directly<sup>7</sup>. We also caution that erosion of X-chromosome inactivation and abnormal imprinting can affect the reliability of disease models<sup>8</sup>, particularly in female-derived cells, where it can lead to abnormally high levels of gene expression from the normally inactive X chromosome<sup>9</sup>. Crucially, experiments should be accompanied by detailed information on the cell lines used (including the sex, age and ancestry background of the donor), the culture conditions, the range of passage numbers used, the genome integrity and *Mycoplasma* contamination status following, for instance, the checklist recommendations by the ISSCR<sup>7</sup>. Moreover, pluripotent stem cell lines that are compared should be cultured under the same conditions. The use of control and patient cell lines that are not matched demographically or obtained from different sources is not recommended. For example, deriving all control lines in one laboratory and all patient lines in another can introduce a confounding variable. To minimize these issues, gene editing should ideally be performed in the same laboratory and using the same methods.

With the advent of CRISPR engineering, there has been an increase in the use of genetically engineered iPS cell lines in addition to the use of patient-derived cell lines. Although this approach provides greater control over genetic background, it is important to note that CRISPR–Cas9-modified iPS cell lines require thorough re-characterization, particularly in terms of pluripotency and genome integrity, if subcloning of iPS cell lines is used after editing. The consensus is that multiple genetic backgrounds (individuals) should be subjected to editing to generalize findings. Experimental design could include iPS cell lines that are unmodified at the targeted locus that have been through the process of editing, the parental line, and multiple edited clones to increase the likelihood that phenotypic effects are attributable to the desired on-target editing and not to artefacts of cloning or editing. Another valuable control is to use revertant lines, in which the mutations are corrected in the same cell line. Additionally, orthogonal validation with patient-derived lines or other strategies, such as small hairpin RNAs (shRNAs), CRISPR activation (CRISPRa), CRISPR interference (CRISPRi) or anti-sense oligonucleotides, is recommended. When assessing the effect of environmental factors or studying copy number variations that cannot be engineered into control iPS cell lines, large cohorts of patient-derived lines with demographically matched controls remain the gold standard. Cell villages, chimerooids and mosaic organoids are advancing the ability to achieve the large sample sizes needed for these studies.

Production of iPS cells requires appropriate donor consent for making and using these lines. Consent restrictions, differences in the international regulatory landscape and institutional policies governing the sharing of human cell lines have sometimes prevented the broader use of a set of lines across laboratories and, overall, present a significant challenge in comparing experiments. The KOLF2.1J line (of European ancestry), has been proposed as an all-around well-performing iPS cell line for collaborative studies<sup>10</sup>, but more than one line is needed. Several repositories currently provide iPS cell lines (California Institute for Regenerative Medicine, the US National Institutes of Health (NIH), Allen Institute, the Simons Foundation Autism Research Initiative, Coriell and The Jackson Laboratory). Moving forwards, creating a collection of broadly consented, well-characterized, genetically and ancestrally diverse iPS cell lines that are named in a standardized manner could greatly facilitate protocol benchmarking and accelerate the validation of disease phenotypes across various research settings.

## Generating and characterizing neural cells

Human neural development spans hundreds of days, extending decades into the postnatal period. Timing of in vitro neural differentiation is largely conserved in most culture conditions and consequently, modelling neurogenesis, astrogenesis, oligodendrogenesis, synapse or circuit formation with human cells will involve long, laborious experiments. Given the complexity and duration of these experiments, it is crucial to design and power them with great care<sup>11</sup>. Owing to variability across cell lines, particularly when establishing new protocols and confirming phenotypes, we recommend utilizing multiple iPS or hES cell lines and experimental differentiation batches to enable robust statistical analysis (for example, using two different XX and two different XY donor iPS cell lines, ideally in several independent differentiation experiments). It is critical to set *a priori* criteria for inclusion of lines and avoid removing cell lines *post hoc*. Differentiations will inevitably need to be staggered, and incorporating quality control steps at key stages can enhance reliability and save resources. For instance, during the first month of cortical organoid differentiation, a panel of region-specific markers covering domains of the nervous system can be readily checked by quantitative PCR (qPCR) or similar inexpensive methods (for example, *FOXP2* for forebrain, *EMX1* for dorsal forebrain, *NKX2-1* or *DLX2* for ventral forebrain, *TCF7L2* for diencephalon, *EN1* for midbrain or anterior hindbrain including cerebellum, *HOX* family genes for hindbrain and spinal cord, and *TTR* for choroid plexus). Morphological measures, such as diameter

## Box 1

# Recommendations for characterization and validation of new differentiation protocols

- Morphological characterization, including monitoring and documenting the size and shape of organoids over time, as well as establishing cell health of the 3D culture (for example, using dyes for cell viability)
- Validation across multiple, genetically distinct iPS or hES cell lines
- Dynamic assessment of expression patterns of region-specific markers through gene expression (qPCR) or spatial assessment of key markers by immunocytochemistry, ideally compared with primary brain tissue samples or published relevant datasets
- Single-cell or nucleus RNA profiling of individual or pooled organoids at several stages, followed by computational mapping onto human reference atlases (such as Allen Brain Atlas, Human Cell Atlas or UCSC). After reproducibility is established with single organoid profiling (such as by single-cell RNA sequencing or qPCR), organoids can be pooled.
- Assessment of reliability across lines, experimental batches and individual organoids
- Functional characterization through imaging and electrophysiological recordings

and shape, can sometimes indicate issues of cell identity or viability with some caveats depending on the method. If grown in suspension, organoids need to be separated to avoid fusion, whereas if embedded in extracellular matrices, it can be difficult to run certain assays for which access to cells is needed. Cell death and organoid morphology should be monitored and documented over time and across batches. These characteristics can vary depending on the differentiation method and the culture conditions, and this variability may interfere with phenotyping. It will be helpful to report the degree of cell death in the core of the organoid and to describe how this was accounted for in experiments or mitigated to reduce cavitation, especially in long-term cultures. Immunocytochemistry can also be used, but whereas some cytoarchitectural features can reliably be observed (for example, ventricular-like zones), more mature neuroanatomical features are not generally present in current multi-cellular preparations. For instance, distinguishing subventricular-like zones and outer subventricular-like zones in cortical organoids or defining multiple cortical layers can often be challenging. Functional quality control experiments should be considered, including characterization using calcium imaging, patch clamp or extracellular recording, or neurotransmitter release.

Developing a new differentiation protocol necessitates extensive characterization and validation (Box 1). Implementing established protocols should, at a minimum, confirm cell identity in the neural cultures derived from the pluripotent stem cell lines used and assess assay variability for the phenotype of interest.

Patch clamping or extracellular recordings can be complemented by calcium or voltage imaging in intact organoids or assembloids. Immunocytochemical characterization of organoids has often been limited in the literature, with only selected examples or parts of organoids being presented. Moving forwards, this characterization should include multiple organoid sections, multiple organoids, and multiple experimental batches with appropriate normalization procedures that consider unbiased sampling of cell distribution, quantification of the total number of cells and, when possible, comparison with primary tissue samples. For the development of assembloids, chimerooids and mosaic organoids,

it is essential to carefully detail the conditions for integration, including the optimal stages of differentiation for fusion and mixing, respectively, and assays. For instance, when it comes to generating neuro-immune assembloids in which neural organoids are integrated with immune cells, the timing of integration and the compatibility of the cell culture medium must be carefully considered. For chimerooids and mosaics organoids, examination of clonal dynamics (pluripotent stem or differentiated cells), patterning of neural progenitors used for mixing and the ideal timing is required. Claims of circuit formation in assembloids should include evidence of connectivity by imaging of projections and, ideally, analyses of functional connectivity, such as retrograde or anterograde tracing, synapse formation and electrophysiology.

There should be transparency in reporting experimental details. Methods sections should include information on cell culture medium and composition, cell culture dishes, cell density, oxygen levels, concentration and timing of growth factors added, passage number of starting cells and passage method, the efficiency of formation of a particular structure and how the structure is defined, the lots of various reagents (including extracellular matrices), and the presence of serum or other undefined components. Finally, it is important to deposit protocols in public repositories such as protocols.io, provide cell-type annotations and be available to address questions and share reagents to facilitate method reproducibility across laboratories.

We also consider that it is time to broadly implement established best practices for single-cell analysis across modalities<sup>12</sup>, including detailed reporting of metadata and the uploading of raw data (if permitted by donor consent) and partially processed data (for example, count matrix) on public repositories such as NIH Gene Expression Omnibus, database of Genotypes and Phenotypes (dbGaP), Neuroscience Multi-Omic (NeMO) archive, CellxGene, National Institute of Mental Health Data Archive (NDA), University of California and Santa Cruz (UCSC) browser, even when publishing in journals that do not require this. This will be increasingly important as integration across datasets and comparison to primary tissue molecular atlases become standard procedures. Benchmarking to developmental cell atlases, when available, is necessary to substantiate claims about deriving specialized cell types, rather than solely relying on the expression of a few selected markers. Similarly, claims regarding physiological properties should be supported by evidence beyond the mere presence of a specific cell type in an organoid. Although dissection of parts of organoids for profiling has been commonly performed, we recommend profiling intact organoids to obtain a comprehensive overview of cell composition. If only subregions are used or organoids are pooled, this should be transparently reported. Developing quality scores of similarities and user-friendly tools, such as VoxHunt<sup>13</sup>, is important to expedite the derivation and validation of new cell types.

Considerable effort has been dedicated to developing neural differentiation protocols. However, a major challenge arises when implementing these protocols in other laboratories, as they are often subject to modifications or optimizations. Although improvements are always welcome, even minor adjustments to a protocol, such as altering the concentration or timing of growth factor application, or the schedule for medium changes, may require revalidation of reliability and the identity and functionality of resulting 3D cultures. This practice also poses challenges for reproducibility efforts, as meta-analyses or comparisons across published studies become exceedingly difficult.

### Probing and manipulating neural cells in 3D

One of the next major challenges in leveraging the potential of human multi-cellular models is likely to be the lag in developing and optimizing tools for capturing the functional properties of these complex 3D cultures. So far, most phenotyping efforts have focused on assessing cell diversity and molecular signatures through single-cell transcriptomics or immunocytochemistry and the morphology of organoids. However,

unless there are severe brain structural defects in patients (for example, microcephaly), the morphology of organoids is rarely a reliable measure for phenotyping. A marked reduction in the size of organoids, especially those carrying mutations from individuals with only minor or no changes in brain or cortical volume, is often more indicative of technical issues related to genetic engineering, differentiation processes or variability between different cell lines. Indeed, characteristics of iPS cell lines, their genetic background and the stage of differentiation (timing), contribute to a large fraction of the variability in these models<sup>14,15</sup>. Therefore, it may be necessary to increase the number of lines used in experiments. Even when sample sizes are limited, precise reporting of technical details, including specific timepoints, the number of cell lines, batches and replicates for each experiment, is crucial for reproducibility and reliability. Special attention should be paid to the research question, the scalability of the assay, and the corresponding sample size. For instance, addressing polygenic inheritance is likely to require large numbers of lines and readouts that can be adapted at scale. To substantiate claims about genetic background effects in disease models, a large number of independent lines is likely to be required, often necessitating orthogonal assays for validation or the examination of multiple pedigrees. In these cases, it is also useful to report data on multiple lines from a single individual. Furthermore, the timing for assessing alterations is a crucial factor and will vary depending on the brain region model, differentiation method or disorder being studied. For the spinal cord, which develops faster, experiments can extend up to 40–50 days in vitro. However, for corticogenesis, several months or more are necessary to observe the generation of deep and superficial layer neurons (neurogenesis), unless strategies to accelerate or stop neurogenesis are used. For astrogenesis, assays at early stages in cortical organoids are often confounded by progenitors with astrocytic features. For example, radial glia rather than astrocytes—the latter often require 4–5 months in vitro to develop, unless strategies to accelerate or stop gliogenesis are used<sup>16</sup>. Notably, very early stages of development of guided neural organoids (that is, the first few weeks of in vitro differentiation) can sometimes exhibit more variability than later stage organoids, probably owing to the time needed for various iPS cell lines and states to reach key stages. Related to timing, assessing the maturation or ‘ageing’ status of cell-based organoids and assembloids is crucial for establishing the relevance of disease models, especially for studying neurodegenerative diseases. The same principles hold for neurotoxicology, where the physiological relevance and the developmental stage of the model used should be described and assessed for effective integration with epidemiological data.

Functional assays in organoids and assembloids provide insights into neural development, maturation and disease modelling, with fluorescence imaging, neurotransmitter release and electrophysiological methods being central to these efforts. Calcium imaging with genetically encoded indicators, such as gCaMPx, is relatively straightforward to implement and can now be used to monitor activity across large populations of neurons in intact neural organoids or across multi-part assembloids. Combining this imaging readout with reporter labelling facilitates the identification of cell-type-specific effects. However, challenges remain in finding, validating and delivering cell-type-specific reporters for human neural cells, and these reporters often become silenced over time. However, in contrast to electrophysiological methods, calcium imaging has a slower temporal resolution, is susceptible to photobleaching and phototoxicity, and cannot be equated with spiking or synaptic activity, especially in developing human neural networks.

Patch clamp electrophysiology, sometimes coupled with optogenetic and chemogenetic manipulations, although having low throughput and requiring advanced expertise, remains the gold standard for measuring the activity of individual neurons due to its resolution and precision in capturing the electrical state and membrane properties of a cell. Because this approach has low throughput, it is important to perform power analysis, considering cell-type diversity and maturity, to accurately establish the needed sample size. Meanwhile, extracellular

recordings provide a less invasive alternative that captures the spiking activity of neuronal populations and can be coupled with optogenetic stimulation or neurotransmitter-uncaging strategies. These methods are typically used acutely or require the organoids or assembloids to be sectioned, which can disrupt their 3D structure and potentially affect physiological properties. Multielectrode arrays can be limited by lower signal resolution and can introduce artefacts, especially if 3D organoids require long-term plating on a flat surface. There is need for recording platforms that can seamlessly integrate with and record chronically (ideally for hundreds of days) from intact neural organoids and assembloids without interfering with their 3D self-organization and differentiation, unlike methods that require 2D culture plating for recording. Moreover, human primary tissue recordings are important as a benchmark for neural recording studies involving organoids and assembloids.

Finally, we emphasize that the complexity of an assay should not preclude the rigour or the sample size that is required to obtain robust results. A range of cell lines, biological and technical replicates, and a combination of functional assays are often necessary to ensure that findings are robust.

### Transplanting human neural cells into animals

Organoid and assembloid models of the nervous system are constrained by their lack of functional vascularization, incomplete morphological and electrophysiological maturity of neurons and glial cells, and lack of meaningful sensory and other physiological inputs that shape circuit activity during development. Transplantation of human neural cells into host organisms can mitigate some limitations by enabling vascularization and integration into living host circuits. Transplantation methods hold potential for developing cell therapies and potentially offering a preclinical platform for testing drugs directly on human cells.

Several aspects of experimental design must be considered when developing transplantation models, whether grafting whole organoids or dissociated neural cells derived from 2D or 3D cultures. These experiments are complex and demand surgical expertise. The developmental stage of the transplanted cells and the host at the time of engraftment—whether prenatal, early postnatal or adult—could influence outcomes. Quality control measures are important, given the extended duration of these experiments, sometimes exceeding nine months. Ensuring cell quality prior to transplantation, in terms of viability, purity and cell identity, is vital for enhancing reproducibility. Magnetic resonance imaging offers a non-invasive and efficient method to observe graft integration and growth, although at low resolution, and can facilitate the refinement of new transplantation procedures. When possible, two-photon imaging can be useful and informative, given its resolution and ability to monitor circuit function, but it remains technically challenging<sup>17</sup>. For graft characterization, we advise verifying survival, proliferation, maturation and migration within the central nervous system of the host, alongside monitoring changes in cell identity and graft composition over time and assessing possible cell fusion—a process in which individual human and host cells become one by merging their plasma membranes. Identifying human cells may necessitate staining with human-specific antibodies (for example, anti-HNA (human neutrophil antigen), STEM121, STEM123, anti-NCAM and anti-H-NUC (also known as CDC27)) or orthogonal approaches to distinguish host cells from transplanted cells. Finally, functional integration into host circuits requires validation, including anatomical reconstruction of projections, labelling of efferent and afferent connections, potentially with viral vectors, and electrophysiological characterization, in vivo imaging and optogenetic manipulation, to ensure correct identification of the transplanted cells versus the host cells, and, where relevant, assessment of potential effects on host behaviour.

It is crucial to recognize that xenograft transplantation into the rodent nervous system typically occurs against an immunocompromised background or in immunocompetent animals receiving immunosuppressive

treatments, constraining the neuro-immune interactions that can be tested. Developmental timelines are species-specific. Therefore, engraftment is intrinsically heterochronic and rodents are not always the ideal species for transplantation; other laboratory animals may have to be considered, although the use of primates as a host is ethically challenging.

The transplantation of human cells into the nervous system of animals introduces various ethical, moral, societal and legal concerns that are the focus of ongoing discussions<sup>18</sup>. Here we underscore the importance of ethical aspects, including obtaining human donor consent for the use of human cells in vivo in animals, ensuring animal welfare and being mindful of public perceptions regarding the purpose of these experiments and their promise to develop therapies. As the field progresses and transplantation techniques improve, it is essential to identify potential ethical tipping points, including emergent features, and ensure that these considerations are integrated into the research process and animal welfare<sup>19</sup>.

### Modelling development, evolution and disease

At the core of developing useful, predictive models of biological processes is rigorous benchmarking. To ensure the reliability and relevance of experimental outcomes, it is critical to seek to obtain some ‘ground truth’ information based on normal human brain development and disease models. This is often challenging for neuropsychiatric disorders, but can sometimes be achieved by comparison with human postmortem or surgically resected tissue and clinical or neuroimaging data, using ex vivo fetal tissue preparations, or relevant animal models (including non-human primates), which remain essential tools for neuroscience. Many disease phenotypes have been described in neural cultures derived from iPS cells over the past 15 years. Some of these involve major changes in cell-type composition, electrophysiological function or cell morphology that are in apparent contrast to the clinical presentation in patients. Therefore, it is possible that such phenotypes may be artefactual, a partial manifestation of the in vivo phenotype, or a manifestation of a phenotype that would be compensated in vivo or at a neural circuit level. This is unsurprising, as in vitro differentiation is often noisy and can amplify phenotypes, and these defects should be interpreted in the context of the disease observed in humans. Therefore, asserting major claims regarding the pathophysiology of complex and poorly understood psychiatric disorders without corroborating findings through independent validation in alternative experimental systems or with clinical data can be misleading. It is important to consider that these are only models of aspects of disease and should be interpreted as such. This is especially important for neuropsychiatric disorders (neurodevelopmental and neurodegenerative), as most such conditions are defined behaviourally.

Ensuring robust statistical and experimental design, including the use of randomization, blinding and a priori power calculations, constitutes another fundamental principle. It is crucial to define samples: individual cells, individual organoids versus bulk organoids or assembloids, differentiation batches, transplanted animals, or iPS cell lines of the same or different donors. When multiple data points are collected from the same experimental batch without maintaining independence, there is a risk of inflating false-positive results. Graphs should show all data collected and points should be coloured (or shaped) by the line or batch used; figure legends should include details of what the points represent. A supplementary table could indicate lines and replicates used in each experiment. It is also important to distinguish between technical (assessing variability introduced by the experimental measurement or procedure) and biological (accounting for variability between different biological entities or samples) replicates. Additionally, transparency regarding which organoids or assembloids were excluded from analysis and the rationale behind these decisions<sup>2</sup>, including the definition of technical failure, is paramount. Such transparency can promote the

integrity and reproducibility of research, especially with complex cellular models of the human nervous system.

When modelling evolutionary or developmental features using organoids and assembloids, several experimental considerations and limitations should be considered. This includes the developmental stage of the organoids and assembloids, especially as comparison across species is challenging and claims of ‘uniquely human’ evolutionary features may be related to heterochrony, a phenomenon by which similar developmental processes unfold over different timescales in different species, with neoteny potentially having a role in retaining early developmental features. Variation in pluripotency states across species adds another layer of complexity, potentially affecting the timing of differentiation in vitro and the ability to make accurate phenotypic comparisons. Cell culture conditions under which these models are maintained, particularly the use of high-glucose media and hyperoxia, may alter developmental and evolutionarily relevant processes, and could influence experimental outcomes. Importantly, claims about modelling extinct hominid brain development or function are particularly challenging, as we have limited information on the cognitive and social functioning of these species. These considerations underscore the need for an experimental design that is suited to these unique challenges and the acknowledgment of inherent limitations when using models to study complex biological processes.

Finally, modelling disease using human stem cell-based organoids and assembloids necessitates a deep understanding of the inherent complexities and limitations of these systems. Bridging the gap between cellular models and clinical presentations remains a formidable challenge, and we urge caution in interpreting results and linking cellular phenotypes to complex cognitive and behavioural processes. A primary consideration is the genetic heterogeneity among iPS cell donors, which is significantly larger than the genetic variance seen in inbred animal models. This variability underscores the critical importance of choosing appropriate controls for these experiments, including obtaining and reporting detailed donor information and, ultimately, registering iPS cell lines to ensure that other researchers can easily access them. For iPS cell studies, matching clinical features of the cohort, such as sex and ancestry, or using genetically related controls may be an approach to reduce heterogeneity. Unlike in clinical trials, where sample size estimation and power analyses are standard prerequisites, such practices are not mandatory for in vitro experiments, even when utilizing patient-derived cells. However, the implementation of power analysis to determine the necessary sample size for detecting an expected effect size is essential. In cases where specific power analysis is not feasible or relevant, it is important to use the largest possible sample size. The utilization of multiple cell lines from the same individual has been shown to reach a power plateau<sup>11</sup>, indicating that incorporating cells from multiple individuals is more likely to increase robustness. Ultimately, unlocking the translational potential of this field will necessitate the adoption of enhanced statistical methodologies; for instance, employing linear and generalized mixed effect models<sup>20</sup> that account for data dependence can provide a more accurate analysis compared with traditional methods such as *t*-tests and ANOVA, which assume independence among observations. It is advisable to consult statisticians to ensure the appropriate application of these methods.

## Final remarks

We are at an exciting time in neuroscience, fuelled by the potential of human neural cellular models. These human stem cell-based models, combined with xenograft and other experimental approaches, are poised to accelerate our understanding of human development, evolution and disorders of the nervous system. In conjunction with the prior consensus on nomenclature<sup>1</sup>, the framework for the experimental process that we propose here will hopefully accelerate the application of organoid, assembloid and other multi-cellular models, bringing us

closer to realizing their translational potential. This framework and further international collaborative initiatives across laboratories promise to enhance reliability and reproducibility and improve communication within and across disciplines, informing trainees and those entering the field, as well as publishing, funding and regulatory agencies.

1. Pasca, S. P. et al. A nomenclature consensus for nervous system organoids and assembloids. *Nature* **609**, 907–910 (2022).  
**This consensus article provided a framework for the nomenclature of human multicellular models of nervous system development and disease.**
2. The RIVER working group. Reporting in vitro experiments responsibly—the RIVER recommendations. Preprint at *MetaArxiv* <https://doi.org/10.31222/osf.io/x6aut> (2023).
3. Lezmi, E., Jung, J. & Benvenisty, N. High prevalence of acquired cancer-related mutations in 146 human pluripotent stem cell lines and their differentiated derivatives. *Nat. Biotechnol.* **42**, 1667–1671 (2024).
4. Panda, A. et al. Clonally selected lines after CRISPR–Cas editing are not isogenic. *CRISPR J.* **6**, 176–182 (2023).
5. Abyzov, A. et al. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* **492**, 438–442 (2012).
6. International Stem Cell, I. et al. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat. Biotechnol.* **29**, 1132–1144 (2011).
7. Ludwig, T. E. et al. ISSCR standards for the use of human stem cells in basic research. *Stem Cell Rep.* **18**, 1744–1752 (2023).  
**This report details the ISSCR recommendations and reporting criteria for enhancing the reproducibility and rigour of stem cell research.**
8. Mekhoubad, S. et al. Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10**, 595–609 (2012).
9. Bar, S., Seaton, L. R., Weissbein, U., Eldar-Geva, T. & Benvenisty, N. Global characterization of X chromosome inactivation in human pluripotent stem cells. *Cell Rep.* **27**, 20–29.e23 (2019).
10. Pantazis, C. B. et al. A reference human induced pluripotent stem cell line for large-scale collaborative studies. *Cell Stem Cell* **29**, 1685–1702.e1622 (2022).
11. Brunner, J. W. et al. Power and optimal study design in iPSC-based brain disease modelling. *Mol. Psychiatry* **28**, 1545–1556 (2023).
12. Heumos, L. et al. Best practices for single-cell analysis across modalities. *Nat. Rev. Genet.* **24**, 550–572 (2023).
13. Fleck, J. S. et al. Resolving organoid brain region identities by mapping single-cell genomic data to reference atlases. *Cell Stem Cell* **28**, 1177–1180 (2021).
14. Gordon, A. et al. Long-term maturation of human cortical organoids matches key early postnatal transitions. *Nat. Neurosci.* **24**, 331–342 (2021).
15. Jourdon, A. et al. Modeling idiopathic autism in forebrain organoids reveals an imbalance of excitatory cortical neuron subtypes during early neurogenesis. *Nat. Neurosci.* **26**, 1505–1515 (2023).
16. Wang, M. et al. Morphological diversification and functional maturation of human astrocytes in glia-enriched cortical organoid transplanted in mouse brain. *Nat. Biotechnol.* **43**, 52–62 (2025).
17. Schafer, S. T. et al. An in vivo neuroimmune organoid model to study human microglia phenotypes. *Cell* **186**, 2111–2126.e2120 (2023).
18. Farahany, N. A. et al. The ethics of experimenting with human brain tissue. *Nature* **556**, 429–432 (2018).
19. National Academies of Sciences Engineering and Medicine. *The Emerging Field of Human Neural Organoids, Transplants, and Chimeras: Science, Ethics, and Governance* (National Academies, 2021).  
**This report provides an overview of ethical and governance issues associated with human neural organoids, assembloids and their transplantation put forwards by a committee of scientists, ethicists and legal scholars convened by the National Academies of Sciences Engineering and Medicine.**
20. Yu, Z. et al. Beyond *t* test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron* **110**, 21–35 (2022).

**Acknowledgements** The authors thank members of their laboratories who provided feedback on earlier versions of this manuscript, including A. Jourdon, V. Mariano, T. L. Li, N. Caporale, E. Villa and M. Sutcliffe.

**Author contributions** All authors discussed the content of the manuscript and reviewed and edited the entire manuscript. S.P.P. led the writing of the manuscript following extensive discussions and with input from all authors.

**Competing interests** The authors declare no competing interests.

## Additional information

**Correspondence and requests for materials** should be addressed to Sergiu P. Paşca.

**Peer review information** *Nature* thanks Silvia Velasco and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.