

Scaling national and international improvement in virtual gene panel curation via a collaborative approach to discordance resolution

Zornitza Stark,^{1,2,3,*} Rebecca E. Foulger,⁴ Eleanor Williams,⁴ Bryony A. Thompson,^{3,5} Chirag Patel,⁶ Sebastian Lunke,^{2,3} Catherine Snow,⁴ Ivone U.S. Leong,⁴ Arina Puzriakova,⁴ Louise C. Daugherty,⁴ Sarah Leigh,⁴ Christopher Boustred,⁴ Olivia Niblock,⁴ Antonio Rueda-Martin,⁴ Oleg Gerasimenko,⁴ Kevin Savage,⁴ William Bellamy,⁴ Victor San Kho Lin,⁷ Roman Valls,⁷ Lavinia Gordon,⁷ Helen K. Brittain,⁴ Ellen R.A. Thomas,^{4,8} Ana Lisa Taylor Tavares,⁴ Meriel McEntagart,^{4,9} Susan M. White,^{2,3} Tiong Y. Tan,^{2,3} Alison Yeung,^{2,3} Lilian Downie,^{3,10} Ivan Macciocca,² Elena Savva,² Crystle Lee,² Ain Roesley,² Paul De Fazio,² Jane Deller,¹¹ Zandra C. Deans,¹¹ Sue L. Hill,¹¹ Mark J. Caulfield,⁴ Kathryn N. North,^{1,3,10} Richard H. Scott,⁴ Augusto Rendon,⁴ Oliver Hofmann,⁷ and Ellen M. McDonagh^{4,12}

Summary

Clinical validity assessments of gene-disease associations underpin analysis and reporting in diagnostic genomics, and yet wide variability exists in practice, particularly in use of these assessments for virtual gene panel design and maintenance. Harmonization efforts are hampered by the lack of agreed terminology, agreed gene curation standards, and platforms that can be used to identify and resolve discrepancies at scale. We undertook a systematic comparison of the content of 80 virtual gene panels used in two healthcare systems by multiple diagnostic providers in the United Kingdom and Australia. The process was enabled by a shared curation platform, PanelApp, and resulted in the identification and review of 2,144 discordant gene ratings, demonstrating the utility of sharing structured gene-disease validity assessments and collaborative discordance resolution in establishing national and international consensus.

Introduction

Robust gene-disease validity assessments are the foundation of accurate variant interpretation in diagnostic genomics. Reporting variants in genes lacking clinical validity poses risks to patient care through the potential for misinterpretation and erroneous diagnosis. Conversely, the rapid growth in new gene-disease relationships resulting from the deployment of genomic sequencing in research and clinical care presents the risk of missed diagnoses if new knowledge is not continuously integrated into diagnostic systems. Gene knowledge databases commonly used in diagnostic systems have traditionally been curated for other primary purposes such as cataloguing Mendelian gene-disease associations (Orphanet, OMIM) or facilitating diagnosis in specific phenotypic groups (Gene2Phenotype).¹ The development of the ClinGen framework significantly advanced evidence-based approaches to gene-disease curation by proposing a scheme for objectively scoring genetic and experimental data.² The application of this framework by expert groups has resulted in the curation of over 900 gene-disease pairs to date,^{3–8} including the systematic identification of

genes with disputed or refuted relationships to disease.⁹ However, with over 3,000 genes reported in association with Mendelian disease,¹⁰ and many of these linked to multiple phenotypes, there is a significant unmet need for evidence-based curated gene resources that can be applied in diagnostic practice today.

Adding further complexity, a common strategy used to increase diagnostic efficiency is the collation of gene-disease validity assessments into “virtual” gene panels for specific testing indications such as intellectual disability. This approach targets genomic analysis to variants that are most likely to be clinically relevant while reducing secondary findings.¹¹ However, it necessitates appropriate panel design, maintenance, and application to mitigate the risk of missing diagnoses, particularly for patients with complex, poorly defined, still evolving, or blended phenotypes. The wide variability in panel content between diagnostic providers is well recognized and raises significant concerns about impact on diagnostic and patient outcomes.^{12–15} Although guidance on panel design and maintenance has been recently provided by the American College of Medical Genetics and Genomics (ACMG),¹⁶ there are few,

¹Australian Genomics Health Alliance, Melbourne, VIC 3052, Australia; ²Victorian Clinical Genetics Services, Murdoch Children’s Research Institute, Melbourne, VIC 3052, Australia; ³University of Melbourne, Melbourne, VIC 3010, Australia; ⁴Genomics England, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK; ⁵Royal Melbourne Hospital, Melbourne, VIC 3050, Australia; ⁶Genetic Health Queensland, Royal Brisbane and Women’s Hospital, Brisbane, QLD 4006, Australia; ⁷Centre for Cancer Research, University of Melbourne, Victorian Comprehensive Cancer Centre, Melbourne, VIC 3000, Australia; ⁸Guy’s and St Thomas’s NHS Trust, London SE1 9RS, UK; ⁹St George’s University Hospitals NHS Trust, London SW17 0QT, UK; ¹⁰Murdoch Children’s Research Institute, Melbourne, VIC 3052, Australia; ¹¹National Health Service England and National Health Service Improvement, London SE1 6LH, UK; ¹²Open Targets and European Molecular Biology Laboratory - European Bioinformatics Institute, Wellcome Genome Campus, Hinxton CB10 1SD, UK

*Correspondence: zornitza.stark@vcgs.org.au

<https://doi.org/10.1016/j.ajhg.2021.06.020>

© 2021 American Society of Human Genetics.



if any, efforts to systematically harmonize these processes across healthcare systems,¹⁷ which results in variability in practice and duplication of effort.

Here, we describe our experience of using PanelApp, an open-access gene and panel curation platform, across two healthcare systems and multiple diagnostic providers with the aim of harmonizing gene panel content through systematic discordance resolution.

PanelApp background

PanelApp is an open-access platform developed to support virtual gene panel curation and incorporate crowd-sourced reviews from multiple experts worldwide.¹⁸ The application programming interface (API) can be integrated into sequencing analysis pipelines or variant curation platforms to assist in variant prioritization, and all panel content is downloadable. Gene-disease validity assessments are captured through structured reviews, including disease association, mode of inheritance, and supporting publications. The platform also enables curation of recurrent copy number variant sites (CNVs) and known short tandem repeat sites (STRs). PanelApp curators evaluate the reviews and current evidence to designate an overall rating to the gene-disease assessment. Three thresholds based on a traffic light system determine which genes have sufficient evidence to be analyzed for a clinical indication. The “green” rating denotes genes used in diagnostic reporting and requires case-level evidence from three unrelated families or two unrelated families with convincing functional data. The green PanelApp rating was developed from G2P and ClinGen guidelines specifically to be relevant for a wide range of rare diseases. “Amber” and “red” ratings reflect gene-disease associations with moderate or low levels of evidence, respectively. The PanelApp infrastructure is an open-source Python application under the Apache 2 license and is freely available with deployment support for commercial cloud vendors and a local environment for development and testing. The cloud version utilizes managed services wherever possible, reducing maintenance requirements and simplifying integration with existing computational infrastructure.

The platform was designed by Genomics England for the 174 rare disease and cancer susceptibility panels used in the 100,000 Genomes Project.¹⁹ Building upon this, the National Health Service (England) (NHSE) Genomic Medicine Service (GMS) is poised to deliver genomic testing through a network of genomic laboratory hubs (GLHs) for clinical indications defined in the National Genomic Test Directory. The Genomics England PanelApp instance has transitioned to support the GMS by harnessing clinical and laboratory knowledge from GLHs to enable expert curation of 173 panels for diagnostic use. Clinical validity assessments are currently available for 5,859 genes, 63 CNVs, and 21 STRs (Figure 1A).

PanelApp Australia was deployed in 2019 by the Australian Genomics Health Alliance, a national collaborative

research partnership piloting approaches and developing evidence to inform national integration of genomics into mainstream care.²⁰ It currently hosts 255 virtual gene panels developed by Australian genomics rare disease and cancer cohort studies, as well as panels created by diagnostic laboratories and clinical services. Clinical validity assessments are available for 5,105 genes, 62 CNVs, and 30 STRs. The Australian instance has been used for consolidation of virtual panels between Australian diagnostic services with the aim of national harmonization.

Maintaining separate PanelApp instances allows separate governance in relation to panel repertoire, which is influenced by local factors such as genomic test funding for specific clinical indications. It also supports autonomous decision-making and facilitates local accreditation and integration into diagnostic pipelines. For example, gene panel versions used for diagnosis in the NHSE are overseen by the NHSE Genomics Clinical Reference Group and reviewed by test evaluation working groups, while individual Australian diagnostic laboratories retain oversight of the panels they use. On the other hand, the use of the same underlying software enables technical co-development and pooling of curated data and expertise to drive quality improvements.

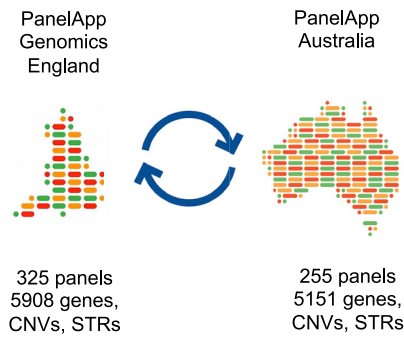
Panel comparison and discordance resolution

Comparing genetic variant pathogenicity classifications by different laboratories to identify discordance is a powerful tool for improving the quality of variant curation. It has been facilitated at scale by platforms such as ClinVar, where variant assessments are deposited by multiple providers via a structured format.^{21,22} The publication of widely adopted standardized criteria for assessing variant pathogenicity,²³ and quality improvement schemes measuring concordance, such as those initiated by the Clinical Sequencing Evidence-Generating Research (CSER) Consortium,^{24,25} the European Molecular Genetics Quality Network (EMQN), and Genomics Quality Assessment (GenQA), which is part of the UK National External Quality Assessment Service (NEQAS) Consortium, have improved consistency between diagnostic providers. Similar efforts in gene curation are hampered by the lack of agreed terminology, agreed curation standards, and platforms that can be used to identify and resolve discrepancies at scale. The use of a common platform including a shared data model and gene curation scheme enabled us to undertake rapid systematic comparison between gene panels used in the UK and Australia with the aim of harmonizing content and identifying areas for process improvement.

Comparing approaches to panel design

First, we performed mapping to identify comparable gene panels (panel pairs) between the two PanelApp instances,

A OVERVIEW



B PANEL COMPARISON: MITOCHONDRIAL DISEASE PANEL

i	Green TACO1 2 reviews	Compare	TACO1 Green 0 reviews
ii	Green TANGO2 2 reviews		
iii	Amber TARS2 2 reviews	Compare	TARS2 Green 0 reviews
i	Green TAZ 2 reviews	Compare	TAZ Green 0 reviews
iv	Red TIMM44 1 review		
ii			TIMM50 Green 0 reviews
i	Green TIMM8A 2 reviews	Compare	TIMM8A Green 0 reviews

C PANEL HARMONIZATION: INTELLECTUAL DISABILITY PANEL DIAGNOSTIC-GRADE GENES

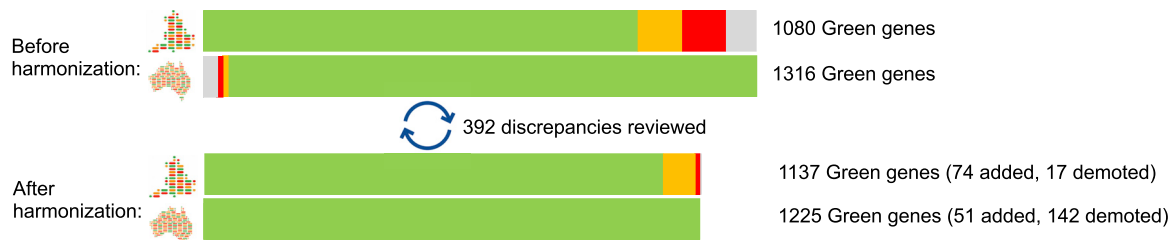


Figure 1. Virtual gene panel comparison and harmonization via PanelApp

(A) The PanelApp instances deployed in the UK and Australia use a common data model and curation guidelines but maintain separate governance and panel repertoires.

(B) The “compare panels” feature in PanelApp can be used to visualize discrepancies between gene panels, as displayed here for a section of the mitochondrial disease panels. The comparison tool highlighted gene rating agreement (i) as well as discrepancies requiring review, including diagnostic-grade genes included in one panel but not the other (ii) and genes with different ratings (iii). Red genes on one panel but missing on the other (iv) were not reviewed because of low level of evidence for inclusion.

(C) Improved alignment between the two intellectual disability panels following exchange of reviews. The green bars represent number of genes rated “green” on each panel. The gray, amber, and red bars represent genes that were rated “green” on one panel but were either absent (gray) or discordantly rated on the other.

which presented several immediate challenges. Gene panel names do not utilize disease ontologies but are named on the basis of common clinical indications for testing. Although the names are usually a good guide to content, panel comparability is also determined by the rules applied in panel design and by clinical context. For example, the two “photosensitivity disorders” panels had non-overlapping content: one contained genes primarily related to porphyrias and the other contained genes primarily related to DNA repair disorders. The “hearing loss” and “deafness” panels only partially overlapped in content because of the inclusion of syndromic disorders in one panel but not the other, and a similar issue was identified between the two “genetic epilepsy” panels: metabolic disorders were included in one panel but not the other. We also encountered significant differences in how panels are subdivided within disease categories. For example, PanelApp Australia contains 17 panels covering primary immunological disorders, whereas Genomics England PanelApp has a single panel. To enable comparison, we used the PanelApp “superpanels” functionality, which groups panels together, such as within a disease category (i.e., immunological disorders).

The mapping exercise highlighted common tensions in panel design in relation to limiting or broadening panel scope¹⁴ on the basis of factors such as age of presentation (e.g., adult-onset ataxia) or absence of syndromic features (e.g., non-syndromic hearing loss). Smaller panels can widen access to genomic testing by enabling sub-specialists to order genomic tests within their area of expertise. From a laboratory perspective, smaller panels increase clinical specificity while limiting the number of variants of uncertain significance. However, this approach carries diagnostic risks if limited panels are applied to investigate patients with complex, evolving, and blended phenotypes. Further, smaller panels can add complexity to test requisition and can act as a barrier for mainstream clinicians. The two PanelApp instances increasingly support overarching panels or superpanels within and across disease categories (“kidneyome,” “pediatric disorders,” “Mendeliome”) to facilitate comprehensive analysis for complex patients and aid the transition of genomic testing toward a first-tier mainstream test used early in the diagnostic trajectory. These shifts need to be coupled with education, training, and decision support tools to promote best practice as part of test requisition, analysis, and reporting. Mapping

Box 1. Key recommendations for the development, use, and maintenance of virtual gene panels

- Use a standardized gene curation framework during panel development and set clear evidence thresholds for diagnostic use
- Use standard ontologies and record case-level and functional evidence for each gene-disease association and proposed mode of inheritance
- Provide clear panel descriptions to guide clinicians and laboratories in panel application, explaining the clinical indication and listing any significant clinical or technical exclusions
- Encourage application of multiple panels (co-application) and the use of larger panels within and across disease categories to mitigate the risks of restricted analysis
- Make your gene-disease and virtual gene panel data openly available and promote collaboration and harmonization nationally and internationally to reduce duplication of effort and improve diagnostic outcomes
- Focus expert gene curation effort on resolving discrepancies and on continuous surveillance and assessment of new evidence for gene-disease associations to enable rapid gains in evidence-based practice
- Use version-controlled panels and provide easy access to previous versions to ensure a clear understanding of the analysis performed and guide future re-analysis
- Develop specific training and competencies in gene and virtual panel curation across the genomic workforce and enable wide participation in the curation effort

of panel clinical indications to disease ontologies such as the Monarch Disease Ontology (MONDO)²⁶ will assist in both harmonization and automation efforts. Most importantly, the rules governing panel design and use need to be articulated clearly to guide clinicians and diagnostic laboratories, particularly because open resources like PanelApp have a wide range of international users (Box 1, key recommendations).

Comparing and harmonizing panel content

We identified 80 comparable panel pairs and, allowing for differences arising from panel scope definition, reviewed their content to identify discrepancies arising from differences in gene-disease validity assessment and relevance to panel indication. We uploaded panels from the Genomics England instance to the Australian instance in order for the panels to be compared side-by-side with the “compare panels” functionality in PanelApp; visualizing panels side-by-side readily identifies discrepant gene pairs. This includes genes that are present on one panel but absent from the other and genes with discrepant ratings, for example, “green-red” (Figure 1B). Genes rated red in one instance but absent from the other were not assessed. All discrepant gene pairs were reviewed initially by the Australian curation team, assessing the evidence provided by the Genomics England curation team. Where the Australian team disagreed, a review with the additional evidence was provided on the Genomics England PanelApp instance. These reviews were assessed independently by the Genomics England PanelApp curation team, resulting in further discrepancy resolution.

Of the 80 panel pairs compared, only three relatively small panels covering highly specific clinical indications

were completely concordant (hematuria, osteopetrosis, and chondrodysplasia punctata). A total of 2,144 discrepant gene ratings and inclusions were identified in the remaining 77 panel pairs. The evidence for gene-disease association and relevance to panel clinical indication was assessed for each discrepancy, and structured reviews were exchanged between the two curation teams via the PanelApp platforms for the promotion of discordance resolution. Of the reviews exchanged, 1,008 reviews (47%) indicated that there was sufficient evidence for diagnostic use (green rating) and 1,139 reviews (53%) indicated moderate or limited evidence for gene-disease association or panel inclusion (425 amber and 714 red). Common sources of discordance are presented in Table S1, and all reviews exchanged and curator assessments are publicly available on the PanelApp instances.

As an example, the largest panel, intellectual disability (ID), had the largest number of discrepancies, 595, including 392 affecting diagnostic-grade (green) genes. Of these, 39 were due to a difference in panel design and were not reviewed further: the Genomics England PanelApp panel contains genes associated with moderate to profound ID, and complex disorders are included if ID is a presenting feature. Individuals presenting predominantly with other clinical features can access testing through alternative pathways and their associated virtual panels (e.g., inborn errors of metabolism) because genomic testing is funded through the healthcare system for a range of indications. By contrast, funding of genomic testing in Australia is more variable; federal funding is currently only available for syndromic and non-syndromic ID, necessitating more inclusive panel design. Following re-evaluation of the evidence for each of the remaining 353 discordances involving diagnostic-grade (green) genes, 298 (84%) were resolved (Figure 1C). Fifty-five (16%)

remained discordant; all were rated green in PanelApp Australia and amber/red in Genomics England PanelApp. The most common reasons for remaining discrepant ratings were variability in the level of ID reported (22), insufficient cases to determine whether ID is a key presenting feature (17), and differential weighing of functional evidence where only two unrelated families are reported in the literature (9).

Building the evidence base for gene-disease associations

The pace of rare disease gene discovery continues unabated,²⁷ presenting a constant challenge for curated databases of gene-disease associations to maintain currency with flow-on effects on diagnostic outcomes both for analysis of new data and reanalysis of existing data.^{28,29} The ACMG recommends regular monitoring and revision of diagnostic panels and at least a 6-monthly review cycle incorporating consultation with disease experts, literature surveys, and monitoring of curated databases.¹⁶ We combined the curator resources across the two PanelApp instances to conduct a monthly literature review of 20 genetics and other subspecialty journals (Table S2) to augment crowdsourced expert reviews and monitoring of other curated databases. Over the first 6 months, this process resulted in 219 new gene curations. This included 157 novel gene-disease associations and 62 changes to existing gene curations incorporating reports of novel allelic disorders, phenotype expansions, new modes of inheritance, or additional case reports or functional data altering evidence ratings.

A total of 107 new gene-disease associations with enough evidence for diagnostic use were identified between June and November 2020. Of these, 41 (38%) are yet to be included in OMIM, highlighting the importance and value of a proactive approach to the identification and curation of new evidence for diagnostic benefit. We found frequent, regular searches for new evidence easier to integrate with existing workflows, while the pooling of curator resources on an open platform reduced individual curation burden. However, the optimal frequency of panel updates needs to take governance and other laboratory requirements into consideration; some diagnostic providers may elect to integrate new information into diagnostic pipelines at less frequent intervals, for example quarterly or 6-monthly, by utilizing panel version control features.

Gene curation and panel harmonization: Future directions

The harmonization effort described here has several, chiefly technical, limitations highlighting the need for ongoing platform development.³⁰ Our approach relied on manual transfer of reviews, which is laborious; auto-

rating key notifications between PanelApp instances, for example when genes are promoted to or demoted from a diagnostic rating (green), would increase efficiency and benefit the wider community. The use of literature mining tools to identify reports of gene-disease associations would make the new gene identification effort more comprehensive while allowing gene curators to focus on critical appraisal. Most of the work undertaken to date relates to genes implicated in rare Mendelian disorders, and further co-development should address the broader applications of diagnostic and research genomics, including somatic cancer testing and complex polygenic disorders.

In conclusion, maintaining curated gene-disease association databases for diagnostic use is a complex task and benefits from pooling expertise across multiple professional groups, including diagnostic scientists, bioinformaticians, clinicians, and researchers. The necessity for broad input has to be coupled with expertise in gene curation, and we need to develop resources, provide training, and establish competencies across the genomic workforce. Sharing interpretation knowledge and diagnostic experience across multiple providers, including internationally, is key to accelerating the implementation of genomics in healthcare.³¹

We have demonstrated the utility of a collaborative approach to identifying and resolving discrepancies in gene and virtual gene panel curation facilitated by the use of a common data platform. The PanelApp software is open source and available for other healthcare systems to implement and participate in data sharing. Harmonization of gene-disease validity assessments between gene curation initiatives will be facilitated by the development of agreed terminology, federated platforms for sharing knowledge via standardized formats, and a focused curation effort to resolve discrepancies. We anticipate that the establishment of the Gene Curation Coalition, of which both Genomics England PanelApp and PanelApp Australia are members, will serve as a focal point for international harmonization efforts. The large number of discrepancies resolved here highlights the scale of the challenge but also provides a blueprint for extending this effort in the future.

Supplemental information

Supplemental information can be found online at <https://doi.org/10.1016/j.ajhg.2021.06.020>.

Acknowledgments

We thank all PanelApp reviewers and those who have contributed feedback or gene lists to help in the development of PanelApp; individual panels show the names and affiliations of contributors. The Australian Genomics Health Alliance is supported by the Australian National Health and Medical Research Council (GNT1113531). M.J.C. was funded by the National Institute for Health Research (NIHR) as part of the portfolio of translational research of the NIHR Biomedical Research Center at Barts and

The London School of Medicine and Dentistry. He is supported as an NIHR senior investigator, and this work was funded by the MRC eMedLab award. This research was made possible through access to the data and findings generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health). The 100,000 Genomes Project is funded by the NIHR and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project uses data provided by patients and collected by the NHS England as part of their care and support. We thank all participants in the 100,000 Genomes Project.

Declaration of interests

Kathryn North is a member of the Editorial Board of *The American Journal of Human Genetics*. The other authors declare no competing interests.

Web resources

ClinGen, <https://clinicalgenome.org>
European Molecular Genetics Quality Network (EMQN), <https://www.emqn.org>
Gene2Phenotype, <https://www.ebi.ac.uk/gene2phenotype>
Gene Curation Coalition, <https://thegencc.org/>
Genomics England PanelApp, <https://panelapp.genomicsengland.co.uk/>
Genomics Quality Assessment (GenQA), <https://www.genqa.org/>
NHS England National Genomic Test Directory, <https://www.england.nhs.uk/publication/national-genomic-test-directories/omim>, <https://omim.org/>
PanelApp Australia, <https://panelapp.agma.umccr.org/>
PanelApp GitLab, <https://gitlab.com/genomicsengland/panelapp/panelapp>
United Kingdom National External Quality Assessment Service (UK NEQAS), <https://ukneqas.org.uk>

References

1. Thormann, A., Halachev, M., McLaren, W., Moore, D.J., Svinti, V., Campbell, A., Kerr, S.M., Tischkowitz, M., Hunt, S.E., Dunlop, M.G., et al. (2019). Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. *Nat. Commun.* *10*, 2373.
2. Strande, N.T., Riggs, E.R., Buchanan, A.H., Ceyhan-Birsoy, O., DiStefano, M., Dwight, S.S., Goldstein, J., Ghosh, R., Seifert, B.A., Sneddon, T.P., et al. (2017). Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource. *Am. J. Hum. Genet.* *100*, 895–906.
3. Adler, A., Novelli, V., Amin, A.S., Abiusi, E., Care, M., Nannenberg, E.A., Feilotter, H., Amenta, S., Mazza, D., Bikker, H., et al. (2020). An International, Multicentered, Evidence-Based Reappraisal of Genes Reported to Cause Congenital Long QT Syndrome. *Circulation* *141*, 418–428.
4. McGlaughon, J.L., Pasquali, M., Wallace, K., Ross, J., Senol-Cosar, O., Shen, W., Weaver, M.A., Feigenbaum, A., Lyon, E., Enns, G.M., et al. (2019). Assessing the strength of evidence for genes implicated in fatty acid oxidation disorders using the ClinGen clinical validity framework. *Mol. Genet. Metab.* *128*, 122–128.
5. DiStefano, M.T., Hemphill, S.E., Oza, A.M., Siegert, R.K., Grant, A.R., Hughes, M.Y., Cushman, B.J., Azaiez, H., Booth, K.T., Chapin, A., et al. (2019). ClinGen expert clinical validity curation of 164 hearing loss gene-disease pairs. *Genet. Med.* *21*, 2239–2247.
6. Ingles, J., Goldstein, J., Thaxton, C., Caleshu, C., Corty, E.W., Crowley, S.B., Dougherty, K., Harrison, S.M., McGlaughon, J., Milko, L.V., et al. (2019). Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. *Circ. Genom. Precis. Med.* *12*, e002460.
7. Grant, A.R., Cushman, B.J., Cavé, H., Dillon, M.W., Gelb, B.D., Gripp, K.W., Lee, J.A., Mason-Suares, H., Rauén, K.A., Tartaglia, M., et al. (2018). Assessing the gene-disease association of 19 genes with the RASopathies using the ClinGen gene curation framework. *Hum. Mutat.* *39*, 1485–1493.
8. Renard, M., Francis, C., Ghosh, R., Scott, A.F., Witmer, P.D., Adès, L.C., Andelfinger, G.U., Arnaud, P., Boileau, C., Callewaert, B.L., et al. (2018). Clinical Validity of Genes for Heritable Thoracic Aortic Aneurysm and Dissection. *J. Am. Coll. Cardiol.* *72*, 605–615.
9. Hosseini, S.M., Kim, R., Udupa, S., Costain, G., Jobling, R., Liston, E., Jamal, S.M., Szybowska, M., Morel, C.F., Bowdin, S., et al. (2018). Reappraisal of Reported Genes for Sudden Arrhythmic Death: Evidence-Based Evaluation of Gene Validity for Brugada Syndrome. *Circulation* *138*, 1195–1205.
10. Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am. J. Hum. Genet.* *97*, 199–215.
11. Matthijs, G., Souche, E., Alders, M., Corveleyn, A., Eck, S., Feenstra, I., Race, V., Sistermans, E., Sturm, M., Weiss, M., et al. (2016). Guidelines for diagnostic next-generation sequencing. *Eur. J. Hum. Genet.* *24*, 1515.
12. Chambers, C., Jansen, L.A., and Dhamija, R. (2016). Review of Commercially Available Epilepsy Genetic Panels. *J. Genet. Couns.* *25*, 213–217.
13. Pেকেles, H., Accogli, A., Boudrahem-Addour, N., Russell, L., Parente, F., and Srour, M. (2019). Diagnostic Yield of Intellectual Disability Gene Panels. *Pediatr. Neurol.* *92*, 32–36.
14. Pellacani, S., Dosi, C., Valvo, G., Moro, F., Mero, S., Sicca, F., and Santorelli, F.M. (2020). Customized multigene panels in epilepsy: the best things come in small packages. *Neurogenetics* *21*, 1–18.
15. Erger, F., Schaaf, C.P., and Netzer, C. (2019). Which genes to assess in the NGS diagnostics of intellectual disability? The case for a consensus database-driven and expert-curated approach. *Mol. Cell. Probes* *45*, 84–88.
16. Bean, L.J.H., Funke, B., Carlston, C.M., Gannon, J.L., Kantarci, S., Krock, B.L., Zhang, S., Bayrak-Toydemir, P.; and ACMG Laboratory Quality Assurance Committee (2020). Diagnostic gene sequencing panels: from design to report-a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.* *22*, 453–461.
17. Krahn, M., Biancalana, V., Cerino, M., Perrin, A., Michel-Calemard, L., Nectoux, J., Leturcq, F., Bouchet-Séraphin, C., Acquaviva-Bourdain, C., Campana-Salort, E., et al. (2019). A National French consensus on gene lists for the diagnosis of myopathies using next-generation sequencing. *Eur. J. Hum. Genet.* *27*, 349–352.

18. Martin, A.R., Williams, E., Foulger, R.E., Leigh, S., Daugherty, L.C., Niblock, O., Leong, I.U.S., Smith, K.R., Gerasimenko, O., Haraldsdottir, E., et al. (2019). PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat. Genet.* *51*, 1560–1565.
19. Turnbull, C., Scott, R.H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F.B., Halai, D., Baple, E., Craig, C., Hamblin, A., et al. (2018). The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. *BMJ* *361*, k1687.
20. Stark, Z., Boughtwood, T., Phillips, P., Christodoulou, J., Hansen, D.P., Braithwaite, J., Newson, A.J., Gaff, C.L., Sinclair, A.H., and North, K.N. (2019). Australian Genomics: A Federated Model for Integrating Genomics into Healthcare. *Am. J. Hum. Genet.* *105*, 7–14.
21. Harrison, S.M., Dolinsky, J.S., Chen, W., Collins, C.D., Das, S., Deignan, J.L., Garber, K.B., Garcia, J., Jarinova, O., Knight Johnson, A.E., et al. (2018). Scaling resolution of variant classification differences in ClinVar between 41 clinical laboratories through an outlier approach. *Hum. Mutat.* *39*, 1641–1649.
22. Harrison, S.M., Dolinsky, J.S., Knight Johnson, A.E., Pesaran, T., Azzariti, D.R., Bale, S., Chao, E.C., Das, S., Vincent, L., and Rehm, H.L. (2017). Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet. Med.* *19*, 1096–1104.
23. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* *17*, 405–424.
24. Amendola, L.M., Muenzen, K., Biesecker, L.G., Bowling, K.M., Cooper, G.M., Dorschner, M.O., Driscoll, C., Foreman, A.K.M., Golden-Grant, K., Grealley, J.M., et al. (2020). Variant Classification Concordance using the ACMG-AMP Variant Interpretation Guidelines across Nine Genomic Implementation Research Studies. *Am. J. Hum. Genet.* *107*, 932–941.
25. Amendola, L.M., Jarvik, G.P., Leo, M.C., McLaughlin, H.M., Akkari, Y., Amaral, M.D., Berg, J.S., Biswas, S., Bowling, K.M., Conlin, L.K., et al. (2016). Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. *Am. J. Hum. Genet.* *99*, 247.
26. Mungall, C.J., McMurry, J.A., Köhler, S., Balhoff, J.P., Borromeo, C., Brush, M., Carbon, S., Conlin, T., Dunn, N., Engelstad, M., et al. (2017). The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Res.* *45* (D1), D712–D722.
27. Bamshad, M.J., Nickerson, D.A., and Chong, J.X. (2019). Mendelian Gene Discovery: Fast and Furious with No End in Sight. *Am. J. Hum. Genet.* *105*, 448–455.
28. Wright, C.F., McRae, J.F., Clayton, S., Gallone, G., Aitken, S., FitzGerald, T.W., Jones, P., Prigmore, E., Rajan, D., Lord, J., et al. (2018). Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genet. Med.* *20*, 1216–1223.
29. Stark, Z., Schofield, D., Martyn, M., Rynehart, L., Shrestha, R., Alam, K., Lunke, S., Tan, T.Y., Gaff, C.L., and White, S.M. (2019). Does genomic sequencing early in the diagnostic trajectory make a difference? A follow-up study of clinical outcomes and cost-effectiveness. *Genet. Med.* *21*, 173–180.
30. Knowles, R.M., Mateen, B.A., and Yehudi, Y. (2021). We need to talk about the lack of investment in digital research infrastructure. *Nat. Comput. Sci.* Published online March 25, 2021. <https://doi.org/10.1038/s43588-021-00048-5>.
31. Stark, Z., Dolman, L., Manolio, T.A., Ozenberger, B., Hill, S.L., Caulfield, M.J., Levy, Y., Glazer, D., Wilson, J., Lawler, M., et al. (2019). Integrating Genomics into Healthcare: A Global Responsibility. *Am. J. Hum. Genet.* *104*, 13–20.

Supplemental information

Scaling national and international improvement in virtual gene panel curation via a collaborative approach to discordance resolution

Zornitza Stark, Rebecca E. Foulger, Eleanor Williams, Bryony A. Thompson, Chirag Patel, Sebastian Lunke, Catherine Snow, Ivone U.S. Leong, Arina Puzriakova, Louise C. Daugherty, Sarah Leigh, Christopher Boustred, Olivia Niblock, Antonio Rueda-Martin, Oleg Gerasimenko, Kevin Savage, William Bellamy, Victor San Kho Lin, Roman Valls, Lavinia Gordon, Helen K. Brittain, Ellen R.A. Thomas, Ana Lisa Taylor Tavares, Meriel McEntagart, Susan M. White, Tiong Y. Tan, Alison Yeung, Lilian Downie, Ivan Macciocca, Elena Savva, Crystle Lee, Ain Roesley, Paul De Fazio, Jane Deller, Zandra C. Deans, Sue L. Hill, Mark J. Caulfield, Kathryn N. North, Richard H. Scott, Augusto Rendon, Oliver Hofmann, and Ellen M. McDonagh

Table S1. Common sources of discordant gene inclusion or rating on diagnostic panels and further recommendations for practice.

Gene-disease validity assessments	
New evidence since original assessment	<p><i>HARS2</i> rating upgraded to diagnostic on Hearing Loss panel due to publication of 6 unrelated families in 2019.^{1; 2}</p> <p>Mode of inheritance for <i>SCN8A</i> changed from mono-allelic to both mono-allelic and bi-allelic due to reports of families with bi-allelic variants and more severe phenotype in 2019.³</p> <p>Recommendation: new evidence for gene-disease relationships should be sought proactively on a regular basis.</p>
Reappraisal of historical gene-disease associations	<p>Association between <i>HOXA11</i> and radioulnar synostosis with amegakaryocytic thrombocytopenia downgraded to Amber as two families reported in 2000, and none since.</p> <p>Association between <i>KIRREL3</i> and intellectual disability (ID) downgraded to Red as the genetic variants originally reported are present at a high frequency in population datasets.</p> <p>Association between <i>TUBA8</i> and cortical malformations downgraded to Red as publication reporting alternative diagnosis in original families identified.⁴</p> <p>Recommendation: evidence for gene-disease associations should be critically appraised, especially those predating genome-wide approaches to gene discovery and the existence of large reference population datasets. Academic publishers should encourage an evidence-based approach to asserting gene-disease associations, particularly in subject reviews.</p>
Gene-disease associations reported in isolated populations	<p><i>PET100</i> promoted to Green on ID panel: initially founder variant in eight Lebanese individuals reported, but one further family in different population identified, as well as functional data.</p> <p>Recommendation: caution should be exercised when gene-disease associations rest on evidence solely derived in isolated populations as the reported variant could be in linkage disequilibrium with the true causative variant which affects a different gene.</p>
Use of unpublished evidence from diagnostic	<p>Association between <i>FAM20C</i> and craniosynostosis rated as Green due to personal communication with research group reporting unpublished cases.</p>

laboratories and research groups	<p>Recommendation: unpublished evidence should be used judiciously when provenance can be established, and ideally the primary data supporting variant classification reviewed. Decisions based upon unpublished or private data should be explained in as much detail as is possible (the PanelApp software enables comments to be made which are publicly available).</p>
Differential weighting of functional evidence	<p><i>HOMER2</i> rating unresolved: two families reported with mono-allelic variants and deafness. Bi-allelic mouse model has deafness, but mice with mono-allelic variants have normal hearing.</p> <p>Recommendation: detailed gene curation schemes such as the ClinGen framework are particularly well suited to objectively resolving edge cases.</p>
Quality of evidence	<p><i>GRIA1</i> rating unresolved: multiple reports of <i>de novo</i> variants in large intellectual disability cohort studies, but paucity of phenotypic, segregation, or functional data.</p> <p>Recommendation: higher thresholds should be applied where novel gene-disease associations are reported without robust supporting evidence. Publishers should encourage the inclusion of case-level data where possible.</p>
Panel inclusion decisions	
Inclusion due to phenotypic overlap	<p><i>ALMS1</i> included in ciliopathy panels, despite Alstrom syndrome not being considered a ciliopathy.</p> <p>Recommendation: careful consideration should be given to genes that need to be included in panels due to phenotypic overlap and to providing guidance that promotes panel co-application or more extended analysis.</p>
Inclusion of gene family members	<p>Inclusion of <i>PEX</i> genes in disorders of cortical malformation unresolved: only some have published evidence documenting cortical malformations.</p> <p>Recommendation: a consistent approach should be adopted to this scenario and clear guidance and explanation provided to clinical and diagnostic users (e.g. co-application of other panels such as Peroxisomal Disorders encouraged if all genes are not included).</p>
Tractability of variant types by different Next	<p>Inclusion of <i>PRDM13</i> remains unresolved: the only variant types reported are non-coding variants upstream of the gene which</p>

<p>Generation Sequencing (NGS) assays and sample types</p>	<p>may not be tractable by all assays (e.g., exome sequencing).</p> <p>Recommendation: a consistent approach should be adopted to this scenario and clear guidance provided to clinical and diagnostic users (e.g., the gene is included in the panel but analysis and reporting are dependent on technical factors).</p>
<p>Proportion of individuals with the disorder who have a particular phenotypic manifestation</p>	<p>Inclusion of Fanconi anemia (FA) genes on the ID panel remains unresolved: ~20% of individuals with FA have ID.</p> <p>Recommendation: clinical guidance is required for this decision, and consideration of the scope of the panel, established clinical pathways for testing as well as risk of incidental findings.</p>

References:

1. Demain, L.A.M., Gerkes, E.H., Smith, R.J.H., Molina-Ramirez, L.P., O'Keefe, R.T., and Newman, W.G. (2020). A recurrent missense variant in HARS2 results in variable sensorineural hearing loss in three unrelated families. *Journal of human genetics* 65, 305-311.
2. Karstensen, H.G., Rendtorff, N.D., Hindbaek, L.S., Colombo, R., Stein, A., Birkebaek, N.H., Hartmann-Petersen, R., Lindorff-Larsen, K., Hojland, A.T., Petersen, M.B., et al. (2020). Novel HARS2 missense variants identified in individuals with sensorineural hearing impairment and Perrault syndrome. *European journal of medical genetics* 63, 103733.
3. Wengert, E.R., Tronhjem, C.E., Wagnon, J.L., Johannesen, K.M., Petit, H., Krey, I., Saga, A.U., Panchal, P.S., Strohm, S.M., Lange, J., et al. (2019). Biallelic inherited SCN8A variants, a rare cause of SCN8A-related developmental and epileptic encephalopathy. *Epilepsia* 60, 2277-2285.
4. Diggle, C.P., Martinez-Garay, I., Molnar, Z., Brinkworth, M.H., White, E., Fowler, E., Hughes, R., Hayward, B.E., Carr, I.M., Watson, C.M., et al. (2017). A tubulin alpha 8 mouse knockout model indicates a likely role in spermatogenesis but not in brain development. *PLoS one* 12, e0174264.

Table S2. Journals reviewed as part of the PanelApp rare disease updates

	Journal name
1	American Journal of Human Genetics
2	American Journal of Medical Genetics
3	Annals of Clinical and Translational Neurology
4	Annals of Neurology
5	Brain
6	Cell
7	Clinical Genetics
8	European Journal of Human Genetics
9	Genetics in Medicine
10	Human Molecular Genetics
11	Human Mutation
12	Journal of Clinical Investigation
13	Journal of Medical Genetics
14	Nature
15	Nature Communications
16	Nature Genetics
17	Neurology
18	Neurology Genetics
19	PLOS Genetics
20	Science