The Journal of Molecular Diagnostics, Vol. 28, No. 3, March 2026

ELSEVIER

the Journal of Molecular Diagnostics
jmdjournal.org

# Piloting an Interpretive External Quality Assurance Model for Genomic Testing for Childhood Syndromes and Intellectual Disability



Ben Lundie,[*†] Sze Yee Chai,[‡] Alicia B. Byrne,[§] Dimitar Azmanov,[¶] John Christodoulou,[‖**] Matilda A. Haas,[‖††] Karin S. Kassahn,[‡‡§§] Sebastian Lunke,[‖¶¶‖‖] Ami Stott,[††] Bryony A. Thompson,[***] Tony Badrick,[‡] and Bruce Bennetts[†††]

*From Pathology Queensland,* Brisbane, Queensland, Australia; the Institute for Medical Bioscience,† University of Queensland, St Lucia, Queensland, Australia; the Royal College of Pathologists of Australasia Quality Assurance Programs,‡ Sydney, New South Wales, Australia; the Program in Medical and Population Genetics,§ Broad Institute of MIT and Harvard, Cambridge, Massachusetts; the Department of Diagnostic Genomics,¶ PathWest Laboratory Medicine, Perth, Western Australia, Australia; the Murdoch Children's Research Institute,‖ Parkville, Victoria, Australia; the Departments of Paediatrics** and Pathology,‖‖ University of Melbourne, Melbourne, Victoria, Australia; Australian Genomics,†† Parkville, Victoria, Australia; Genetics and Molecular Pathology,‡‡ SA Pathology, Adelaide, South Australia, Australia; the Adelaide Medical School,§§ The University of Adelaide, Adelaide, South Australia, Australia; Victorian Clinical Genetics Services,¶¶ Parkville, Victoria, Australia; the Department of Pathology,*** Royal Melbourne Hospital, Parkville, Victoria, Australia; and the Western Sydney Genetics Program,††† The Children's Hospital at Westmead, Sydney, New South Wales, Australia*

Accepted for publication November 19, 2025.

Address correspondence to Ben Lundie, B.App.Sc., M.Sc., Pathology Queensland, Level 4, Block 7, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia.
E-mail: ben.lundie@health.qld.gov.au.

Few external quality assurance programs adequately address the complexity of largescale human genome or exome analysis. To bridge this gap, Australian Genomics and the Royal College of Pathologists of Australasia Quality Assurance Programs (QAP) developed a pilot interpretive module focused on genomic testing for childhood syndromes and intellectual disabilities. The program assessed laboratories' proficiency in interpreting complex genomic data for pediatric disorders. Six clinically accredited laboratories analyzed standardized genomic, phenotypic, and referral data. Reports were evaluated using a rubric adapted from the European Molecular Genetics Quality Network model, covering genotyping, variant classification, interpretation, and report content. Feedback included comparative performance results and individual recommendations. All laboratories correctly identified and classified target variants, but variation was observed in report structure, inclusion of genetic counseling advice, and application of the American College of Medical Genetics and Genomics/Association for Molecular Pathology classification framework. Participants noted that data-sharing limitations and differences in local reporting practices contributed to scoring inconsistencies. The pilot demonstrated the feasibility of a disease-specific interpretive QAP for complex pediatric genomic testing. Future rounds will address logistical challenges, refine scoring criteria, and strengthen standardization, supporting broader implementation. This initiative lays the groundwork for integrating specialized QAP modules into routine practice to improve diagnostic accuracy and consistency across laboratories. *(J Mol Diagn 2026, 28: 227—237; https://doi.org/10.1016/j.jmoldx.2025.11.006)*

The advent of next-generation sequencing (NGS), including genome and exome sequencing, has greatly improved the diagnostic yield for genetic disorders.[1] NGS enables the detection of a wide range of genetic alterations, including single nucleotide variants, copy number variations, structural variants, gene fusions, methylation patterns, and repeat

Supported by the Quality Use of Pathology Program (QUPP) funding grant 4-GY32AAH. Research conducted at the Murdoch Children's Research Institute (MCRI) was supported by the Victorian Government's Operational Infrastructure Support Program. The Chair in Genomic Medicine (J.C.) is supported by The Royal Children's Hospital Foundation. Australian Genomics is funded by National Health and Medical Research Council grants 1113531, 2000001, and 2035846, and the Medical Research Future Fund, administered by the Murdoch Children's Research Institute. A.B.B. is supported by National Institutes of Health grant U24 HG006834.

Portions of this work were presented at the 2024 American College of Medical Genetics and Genomics annual Clinical Genetics Meeting, March 12—16, 2024, Toronto, ON, Canada.


https://doi.org/10.1016/j.jmoldx.2025.11.006

expansions.[2–4] As NGS evolves, it may replace many traditional diagnostic methods, making proficiency in NGS-based assays increasingly critical.

External quality assurance programs (EQAP), or proficiency testing, are vital for laboratory accreditation. EQAP ensures the accuracy, reliability, and consistency of laboratory testing by independently assessing performance, identifying areas for improvement, and ensuring adherence to best practices.[5] Participation in EQAP enhances technical proficiency, fosters confidence in diagnostic results, and reduces diagnostic errors, ultimately improving patient outcomes.[6–8]

Standards set out by various agencies, including the Australian National Pathology Accreditation Advisory Council (NPAAC), ISO15189:2022, and the United States Clinical Laboratory Improvement Amendments (CLIA) mandate the participation in EQAP to maintain accreditation.[9–11] These EQAP activities must be conducted across the full scope of a laboratory's service including pre-analytical, analytical, and post-analytical activities. Despite these mandates, few existing quality assurance frameworks specifically address complex genomic testing activities, such as variant prioritization and interpretation in a pediatric exome sequencing or genome sequencing testing context. Laboratories use various methods for sample enrichment, quantification, sequencing platforms, and bioinformatics pipelines, all of which can contribute to variability in test outputs. Although clinical laboratory validation ensures that laboratories can detect clinically significant variants, the ability to detect only a few high-impact variants may not be sufficient to assess overall analytical performance. Instead, a comprehensive evaluation should encompass the detection and nondetection of variants throughout the entire genome.

Recognizing this, the European Molecular Genetics Quality Network (EMQN) introduced the DNA Sequencing − NGS (Germline SNVs and indels) (vGermline) module to benchmark analytical performance in NGS-based assays using a consensus approach. This module enables laboratories to demonstrate analytical proficiency across various NGS assays covering diverse genetic alterations. Many disease-specific EQAP modules are available through EMQN, United Kingdom National External Quality Assessment Service (UK NEQAS), College of American Pathologists (CAP), and Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP); however, with the exception of the CAP NGS undiagnosed disorders-exome, none currently assess the challenges posed by complex, multisystemic conditions such as syndromic neurodevelopmental disorders. Furthermore, most existing programs do not fully address the growing trend towards laboratory generalization in molecular diagnostics, where laboratories increasingly offer a broad array of tests rather than specializing in specific genetic disorders.[12]

In 2020, the Australian Medicare Benefits Schedule (MBS) items 73358 and 73359 were introduced to subsidize exome sequencing or genome sequencing to diagnose monogenic disorders in children, so-called childhood syndromes. In this context, RCPAQAP, in collaboration with diagnostic laboratories, developed a pilot of a data-only, interpretive EQAP designed to assess the clinical performance of laboratories engaged in genomic testing for complex disorders, currently not well covered by existing EQAP programs. Specifically, this module is designed to be general in nature with laboratories being provided clinical details that may include multisystem complex disorders, requiring them to select appropriate target genes, if they adopt a targeted approach, and interpret variants across a wide number of genes based on phenotype/genotype correlation, rather than current disease and system-specific modules.

One of the key tenets of the pilot study was that it should reflect a model that was sustainable and was not reliant on the volunteering of experts, especially in the scoring. An important part was to compare scoring by experienced molecular geneticists and competent non-molecular staff. This program aims to bridge the gap in current EQAP offerings (Figure 1) and to ensure that laboratories are equipped to meet the diagnostic challenges posed by emerging technologies and increasingly complex clinical cases.

## Materials and Methods

### Ethics Approval

Research ethics approval was obtained from the Royal Children's Hospital (Melbourne) Research Ethics Committee (HREC/81777/RCHM-2022).

### Development of Scoring Criteria

Scoring criteria (Supplemental Table S1) were adapted from the EMQN QA system with support from EMQN. A sub-working group tailored the criteria to a pilot-specific dry interpretive module, addressing genotyping, variant classification and clinical interpretation, and report content. Each section was worth 2.0 marks using a deductive scoring system. A detailed guide based on the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) classification standards[13] supplemented the criteria, ensuring consistency for nonexperts. Two expert working group members (B.A.T. and J.C.) (senior genomics scientist and clinical geneticist/genetic pathologist) determined the ACMG/AMP criteria and classifications for scoring.

### Case Sourcing

EQA samples were sourced from the Australian Genomics (AG) Genomic Data Repository (GDR). The GDR collates genomic and phenotypic data from participants enrolled in

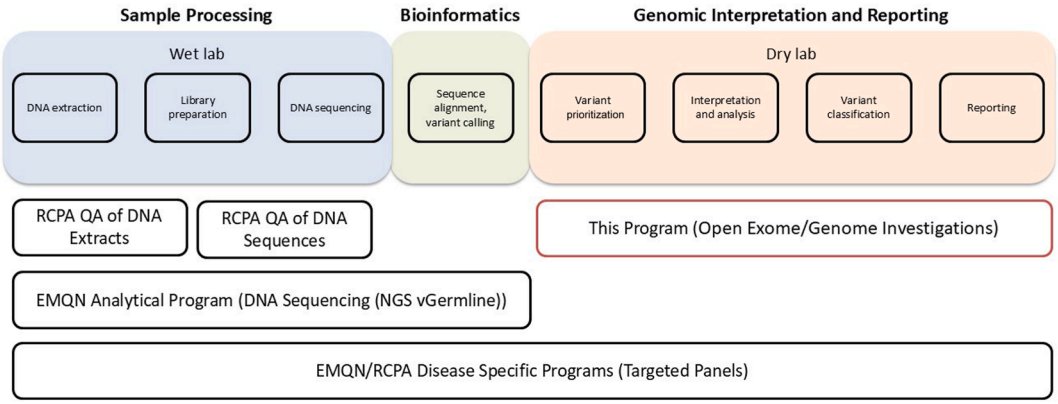


**Figure 1** External quality assurance programs (EQAP) mapped against laboratory processes showing the gap addressed by this program. No EQAP currently exist to comprehensively assess the scope and complexity associated with analysis of exome or genome data. EMQN, European Molecular Genetics Quality Network; NGS, next generation sequencing; QA, quality assurance; RCPA, Royal College of Pathologists of Australasia.

AG Flagship studies across rare disease, cancer, and carrier screening. Participants provided consent for their data to be used in secondary research, and access was granted following approval by the AG Data Access Committee.

To simulate real-world testing conditions, the project plan considered two options for data provision: diagnostic laboratories or the AG GDR. The GDR was selected as the most appropriate source due to its repository of ethically approved, pre-consented cases with associated genomic and phenotypic data. The working group elected to provide a trio dataset, while allowing laboratories to analyze the proband as a singleton if that reflected their standard practice. Case filtering was undertaken by the AG data team according to the eligibility criteria for Medicare Benefits Schedule items 73358/9, and suitable trio cases were identified for use in this study (Figure 2).

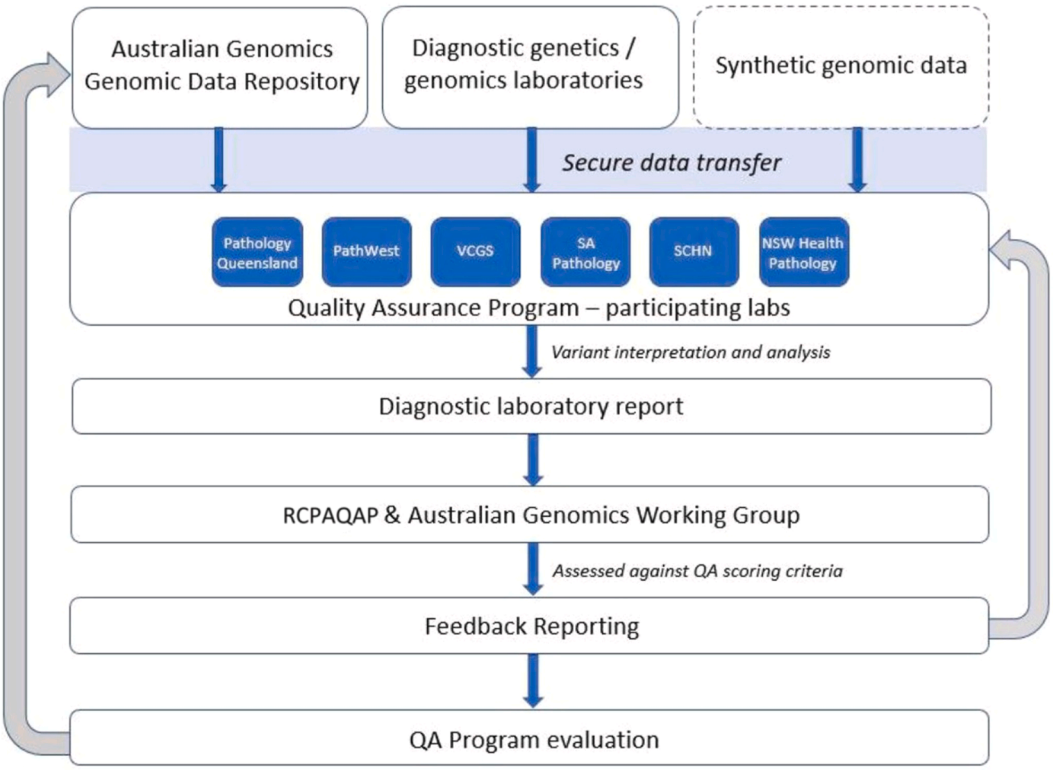


**Figure 2** Process overview for the quality assurance pilot, noting potential sources of data with different data access/request and secure transfer processes (**blue arrows**), including synthetic data (**dotted box**). Laboratories participating in the pilot received a feedback report benchmarking accuracy against the scoring criteria. Ongoing evaluation mechanisms will inform QA program (QAP) development (**gray arrows**). NSW Health Pathology, New South Wales Health Pathology; QA, quality assurance; SA Pathology, South Australia Pathology; SCHN, Sydney Children's Hospital Network; VCGS, Victorian Clinical Genetics Services.

### Data Sharing, Analysis, and Interpretation

Genomic and phenotypic data were provided to six National Association of Testing Authorities (NATA)-accredited Australian laboratories. Genomic data (FASTQ, BAM, VCF) were securely shared for processing with existing pipelines, and referral data were supplied via password-protected PDFs. Laboratories adhered to privacy and security standards, analyzing the data to produce diagnostic reports using their established practices.

### Assessment and Scoring

Diagnostic reports were reviewed by working group members using the scoring criteria and guide. Reports were assessed in an open, identifiable process by both experts and nonexperts, who also provided detailed feedback. Expert assessors included heads of diagnostic genomics services, genetic pathologists, and senior genomics scientists routinely performing variant curation. Nonexpert assessors included research program coordinators and a RCPAQAP senior scientist. Members did not score reports from their own laboratories.

### Feedback

Scores and feedback were compiled into reports for each lab, following the RCPAQAP format. Reports included variant classification details, performance summaries across genotyping, clinical interpretation, report content, and comparative scores (Figure 3).

### Evaluation

Participating laboratories completed a survey delivered in REDCap[14,15] to evaluate pilot acceptability, identify improvement areas, and share their experience (Supplemental Appendix S1). Survey results will inform future QA pilot rounds and publications.

### Data Availability

Data access requests from Australian Genomics are accepted via an online application form that will require approval from the Australian Genomics Data Access Committee. Approved applicants will be required to abide by the Australian Genomics data policies and agreements (see *https://www.australiangenomics.org.au/tools-and-resources/accessing-australian-genomics-data*, last accessed April 22, 2024).

## Results

### Operationalizing the Pilot

The working group was established in May 2021, with initial efforts focused on research ethics submission and data access planning. Following guidance from the Royal Children's Hospital (Melbourne) HREC, research ethics approval (HREC/81777/RCHM-2022) was obtained in August 2022 to meet requirements for academic publication, manage additional findings from data re-analysis, and enable project evaluation through laboratory surveys.

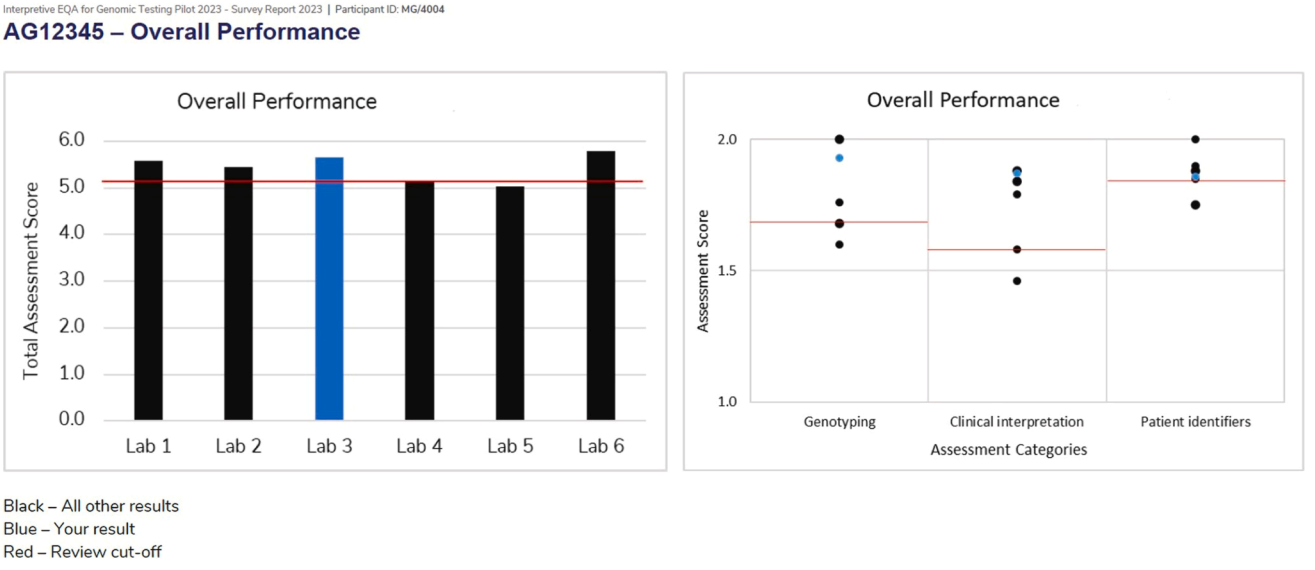


**Figure 3** Example overall performance for laboratory 3 from the individual laboratory feedback report showing comparison of laboratory performance against other participants.

## Case Selection

The final case selected was an 11−month-old female who had undergone investigations suggestive of leukodystrophy. Reported phenotypic features included poor feeding from birth, regression of motor skills at 8 months, irritability, hypotonia, loss of rolling, and reduced head control, with no seizures, hearing loss, or ophthalmological abnormalities. Additional clinical findings included a small anterior fontanelle and short up-slanted palpebral fissures.

Previous testing in this patient had identified compound heterozygous variants in *EIF2B5*, comprising a paternally inherited missense variant in exon 6 [NM_003907.3: c.806G>A (p.Arg269Gln), ClinVar Variant ID 942204, *https://www.ncbi.nlm.nih.gov/clinvar*, last accessed October 4, 2025] and a maternally inherited nonsense variant in exon 8 [NM_003907.3:c.1264C>T (p.Arg422*), ClinVar Variant ID 1806183].

## Participating Laboratories, Data Delivery, and Report Return

Six diagnostic laboratories participated in the pilot: PathWest, South Australia Pathology, Pathology Queensland, New South Wales Health Pathology, Victorian Clinical Genetics Services, and Sydney Children's Hospital Network. Three laboratories required site-specific governance in addition to data-sharing agreements , whereas the others required only data-sharing agreements. Finalizing governance processes for all sites took several months. Data were shared progressively from May to July 2023, with the final diagnostic report returned in October 2023. Average turnaround time was 9 weeks (range: 4 to 18 weeks). Minimal instructions were provided, allowing labs to follow standard diagnostic practices as they were involved in pilot planning.

## Pilot Results

Most laboratories (5/6) performed open investigations, examining variants across all known Mendelian disease genes; 2/6 laboratories complemented these open investigations with targeted analysis using virtual gene panels and/or phenotype-driven analyses: PanelApp[16] Australia (*https://panelapp-aus.org*, last accessed September 16, 2025; intellectual disability syndromic and non-syndromic v0.5393, paediatric leukodystrophy v0.296, regression v0.530, Mendeliome v1.1164, leukodystrophy superpanel v.0.398); and PanelApp UK (Genomics England PanelApp, *https://panelapp.genomicsengland.co.uk*, last accessed September 16, 2025; childhood onset leukodystrophy v.14.2, inherited white matter disorders v.1.175). One laboratory performed a targeted analysis of a selected virtual gene panel only (PanelApp Australia, pediatric leukodystrophy v0.244).

**Table 1** Average Nonexpert, Expert, and Overall Scores for Participating Laboratories

| Lab | Nonexpert | Expert | Combined |
|---|---|---|---|
| 1 | 5.7 | 5.4 | 5.5 |
| 2 | 5.9 | 5.4 | 5.6 |
| 3 | 5.3 | 5.5 | 5.4 |
| 4 | 5.0 | 5.1 | 5.1 |
| 5 | 5.5 | 6.0 | 5.8 |
| 6 | 5.0 | 5.0 | 5.0 |
| Average | 5.4 | 5.4 | 5.4 |

All laboratories accurately identified, classified, and interpreted the target variants. Working group members, including experts and nonexperts, assessed the diagnostic reports using the established scoring criteria and variant classification guide. Of the maximum possible score of 6.0, the pilot achieved an average score of 5.4 (range: 5.0 to 5.8) (Table 1). Five of the six laboratories performed trio analysis, whereas one laboratory restricted the analysis to a singleton due to a misinterpretation of the scheme instructions. This laboratory reported the correct target variants; however, being a singleton analysis, phasing of variants could not be established. In addition to the expected variants, the laboratory performing singleton analysis reported two additional pathogenic variants. One laboratory erroneously switched the parental origin of the target variants for which points were deducted in the genotyping score.

## Assessment

The assessment of diagnostic reports highlighted inconsistencies in scoring and reporting approaches (Table 1). Variations arose from the deductive scoring system, especially criteria with multiple deduction values (eg, 0.2, 0.5, or 1.0 marks for insufficient or incorrect evidence for variant classification). Nonexpert scorers struggled when ACMG/AMP criteria differed from the variant classification guide, or when laboratories did not use ACMG/AMP codes. Determining appropriate evidence levels was particularly challenging without these codes.

Some working group members deducted marks inconsistently or opted to provide comments instead. The presence of two variants further complicated scoring, with some members deducting marks for each variant separately. Differences in diagnostic reports, such as layout, sample type recording, genetic counseling recommendations, approaches to variant prioritization and curation, and assay limitations, also influenced scoring. Relevant scoring criteria have been updated to address these issues.

## Diagnostic Report Layout

The reporting style varied across laboratories, with report length varying from 3 to 5 pages (average 3.8 pages). Some

**Table 2** Genetic Counseling Recommendations Included by Participating Laboratories

| Lab | Genetic counseling recommendation |
|---|---|
| 1 | "The parents of this patient are at 1:4 risk of a recurrence of this condition in any subsequent pregnancy. Referral to a clinical genetics service is recommended for clinical review and professional genetic and reproductive counselling. Prenatal diagnosis is available for these variants." |
| 2 | "Genetic counselling is recommended. Genetic test results may have significant medical implications for both the patient and genetic relatives. Referral of this patient's parents to clinical genetics is recommended to discuss the reproductive implications of these results. Correlation with the patient's clinical phenotype, other investigations and family history is recommended." |
| 3 | "Genetic counselling is recommended." |
| 4 | "This result may have important implications for the extended family, and genetic counselling should be considered." |
| 5 | "Genetic counselling is recommended. Diagnostic/carrier testing in at-risk family members and prenatal testing, where appropriate, is available through this laboratory." |
| 6 | "Genetic counselling is recommended." |

laboratories included variant classification information and associated clinical recommendations/indications upfront. By contrast, others distributed this information throughout the report or towards the end after presenting variant classification and interpretation sequentially. Additionally, some presented the critical information in box/bolded, whereas others included it in the report's body. As such, information was not always located in the same sections of the report.

## Reporting Sample Type and Identification

The working group noted great variability in reporting the specimen type. Reported specimen types included "DNA extracted from blood," "DNA," "Data only" "EDTA whole blood," and "Blood." Some laboratories did not include the provided patient identifier on their diagnostic report.

## Genetic Counseling Recommendations

The level of detail provided regarding genetic counseling recommendations was diverse across laboratories (Table 2). The relevant scoring criterion guidance for reviewer was: "counselling and/or follow-up is relevant but not recommended" with a deduction of either 0 or 0.5 marks.

## Variant Classification and Clinical Interpretation

Most laboratories ($n$ = 5/6) used the ACMG/AMP standards for variant classification. One laboratory used its own modified ACMG/AMP classification system, which was not outlined in the report or publicly available. Of the five laboratories that did use the ACMG/AMP standards, only four included ACMG/AMP criteria codes in the diagnostic report, and the remaining laboratory was able to provide this information on request (although the pilot did not request it). Although the ACMG/AMP criteria and strengths were applied differently, this did not significantly impact the overall classification of the variants for this case. Variant classification was the same across all laboratories except one where one of the two variants were classified as likely pathogenic where all other laboratories called the variant pathogenic (Table 3). Additionally, some reports did not note the parent of origin for each variant.

## Assay Limitations and Variant Prioritization Strategies

There was considerable variability in the documentation of assay details and limitations between the diagnostic reports. One assessor deducted points from laboratories for not providing the scope of the assay because it was not clear from the report whether exome or panel analysis was done.

In other reports, issues such as test sensitivity not being provided or sequencing details appearing to be standard text were noted. It was challenging for nonexperts to assess this, and they were less likely to deduct points or provide comments about these criteria for all reports.

## Evaluation Survey Results

Governance and data-sharing agreements were challenging ($n$ = 5/6), though two comments noted this is unlikely to be an issue for non-research QA programs. Downloading the data package was straightforward for all laboratories, though two labs noted missing metadata and pedigree information. Providing data in accessible formats was raised as a potential future issue. Analysis and reporting took 1 to 5 hours for most laboratories ($n$ = 5/6) and 6 to 10 hours for one laboratory, involving four staff members per site: bioinformaticians ($n$ = 5/6), genetic pathologists ($n$ = 5/6), medical scientists ($n$ = 6), and clinical geneticists ($n$ = 2/6).

Half the laboratories ($n$ = 3/6) found scoring and feedback transparent, while others were neutral ($n$ = 2/6) or disagreed ($n$ = 1/6). Suggestions included clearer instructions to avoid unnecessary point deductions, more detailed feedback across all criteria, faster turnaround times, and better accommodation of diverse approaches to variant classification.

**Table 3** ACMG/AMP Variant Classification Criteria Applied by Laboratories to Classify the Variants

| Expected ACMG/AMP criteria, strength, and classification | Lab 1 | Lab 2 | Lab 3 | Lab 4 | Lab 5 (ACMG/AMP criteria not reported) | Lab 6 (ACMG/AMP framework not used) |
|---|---|---|---|---|---|---|
| Variant 1 *EIF2B5* - NM_003907.3:c.1264C>T (p.Arg422*) | | | | | | |
| PVS1 | PVS1 | PVS1 | PVS1 | PVS1 | — | — |
| *PS4* | *PS4_supporting* | — | — | — | — | — |
| PM2_supporting | PM2_supporting | PM2 | PM2 | PM2_supporting | — | — |
| PM3_strong | PM3 | PM3_strong | PM3 | PM3 | — | — |
| PM5 | — | — | — | — | — | — |
| PP4 | — | — | PP4 | — | — | — |
| Pathogenic | Pathogenic | Pathogenic | Pathogenic | Pathogenic | Pathogenic | Pathogenic |
| Variant 2 *EIF2B5* - NM_003907.3:c.806G>A (p.Arg269Gln) | | | | | | |
| *PS4* | *PS4* | — | — | — | — | — |
| *PM1* | — | — | — | *PM1* | — | — |
| PM2_supporting | PM2_supporting | PM2 | PM2 | PM2_supporting | — | — |
| PM3_very strong | PM3 | PM3_very strong | PM3_strong | PM3_strong | — | — |
| *PM5* | *PM5* | *PM5_supporting* | *PM5* | — | — | — |
| PP3_moderate | PP3_moderate | PP3 | PP3 | PP3_moderate | — | — |
| PP4 | — | — | PP4 | — | — | — |
| Pathogenic | Pathogenic | Pathogenic | Pathogenic | Likely pathogenic | Pathogenic | Pathogenic |

Variants can be accessed via ClinVar (*https://www.ncbi.nlm.nih.gov/clinvar*, last accessed October 4, 2025). Italicized rows were not part of the expected classification.

Most laboratories ($n$ = 5/6) expressed interest in ongoing involvement if the QA program became permanent. One laboratory declined but noted the program's value with modifications. Most laboratories ($n$ = 4/6) supported expansion to other phenotypes/organ systems, suggesting areas such as renal, ocular, immunology, neurology, or prenatal genomics.

One laboratory indicated that ongoing involvement might replace participation in other schemes, whereas another noted that low overhead costs could make yearly participation feasible, despite their usual 3-year cycle for QA programs. Suggestions for improvement included better case ethnic representation, avoiding variants in Shariant,[17] considering synthetic data, enhancing data labelling, and clearer instructions for data-only analysis. Laboratories emphasized the program's value and the importance of continued improvement.

## Discussion

EQAP plays a crucial role in evaluating genomic testing laboratories, ensuring quality assurance for participants. Unlike most quality assurance programs, which focus on targeted gene panels with specific disease applications, this program assesses laboratories' ability to prioritize and interpret disease-causing variants in open genomic investigations based on clinical indications, unrestricted by specific gene panels. Such investigations are common for childhood syndromes, which this QAP pilot addresses. With the consolidation of molecular techniques into comprehensive genomic technologies, it is increasingly important to evaluate laboratories' capability to interpret diverse disease states using a single, versatile platform.

Establishing ethics approvals, governance, and data-sharing agreements posed challenges during the pilot but are unlikely to affect formal EQAP implementation. Standardized data-sharing practices will minimize such hurdles, which are unique to the pilot phase. Formal EQAPs typically avoid site-specific governance applications or complex agreements, relying instead on standard EQA program terms.

Encouragingly, all laboratories successfully prioritized and reported the target variants based on the clinical referral information, despite employing diverse analytical approaches. These included a range of in-house and commercial software tools for variant annotation and prioritization from trio exome sequencing data, as well as virtual gene panels and phenotype-guided strategies. One laboratory that conducted singleton analysis identified two additional pathogenic variants associated with autosomal recessive leukodystrophy. However, laboratories performing trio analysis discarded these variants after parental segregation analysis revealed they were in cis. Although trio analysis typically results in only a modest increase in diagnostic yield, its key advantage lies in identifying *de novo* events and phasing variants, as demonstrated in this case.[18]

The pilot revealed variability in how laboratories reported specimen types, critical for both EQA and routine diagnostics, especially in data-only analyses or re-analyses. Some reports omitted key identifiers such as patient IDs,

compromising clarity and traceability. Standardized report elements, such as specimen collection date, type, methodology, and testing limitations, are essential to ensure uniform, interpretable, and comparable results across laboratories.[19,20] Variability in reporting specimen types and assay methods likely reflects differences in laboratory systems, which can limit how this information is recorded.

Although professional societies have provided recommendations on the content and structure of genomic reports,[20,21] these remain evolving frameworks rather than universally harmonized standards. In Australia, the National Pathology Accreditation Advisory Council (NPAAC) sets mandatory requirements for reporting multigene panel sequencing and human genetic variation.[22,23] These include the use of established nomenclature, classification of variants according to clinical significance with evidence-based guidelines, clear statements of reportable variants and assay scope, and disclosure of limitations. Reports must also specify the reference genome used, unambiguously identify the gene(s) tested, and provide recommendations for follow-up where relevant. Although these requirements establish a baseline for consistency and clarity, they do not yet provide the comprehensive detail or harmonization needed to ensure uniformity in clinical genomic reporting across laboratories, reflecting the broader international landscape where guidance is still maturing.

To reduce inconsistencies, EQAP instructions should specify required report inclusions, including standardized formats for describing specimen types. For processes not performed by the participating laboratory, allowances should be made for omitting nonapplicable details, such as analytical performance, since these cannot be evaluated. Alternatively, EQAP providers could recommend standard methods and limitations as if all processes were conducted in-house. Standardizing these practices will ensure all relevant information is captured, supporting consistent and accurate assessments of laboratory performance.

Despite broad concordance in the final classification of variants, a key challenge lies in standardizing the evaluation of variant classifications and clinical interpretations, particularly given the evolving nature of guidelines for variant classification. Although many laboratories follow the ACMG/AMP standards,[13] the 2015 framework has been supplemented with various updates, disease-specific adaptations, and alternative classification systems such as Sherloc and the Association for Clinical Genomic Science (ACGS) guidelines.[24,25] This diversity in practices adds complexity to scoring, as laboratories may use different evidence thresholds and variant classification codes, or omit classification codes entirely.

One laboratory applied PS4 in a recessive context, which was not included in the recommended scoring criteria. This was based on the use of the ACGS 2020 framework,[25] which permitted PS4 at supporting strength when a variant had been observed in an unrelated affected individual ("proband counting") and did not explicitly exclude recessive disorders. The revised ACGS 2024 guidelines[26] now clarify that PM3 should be used to count biallelic cases in recessive conditions, reserving PS4 for case-control analyses of more common variants.

PM2 was applied at supporting strength by some laboratories in line with updated ACMG recommendations,[27] but retained at moderate strength by those following ACGS guidance. Similarly, PM3 was variably applied from moderate to very strong, reflecting both policy differences around proband counting and whether additional evidence was accrued once a pathogenic threshold had already been reached.

Phenotype-driven evidence (PP4) was applied by only one laboratory, despite being part of the recommended scoring, based on a highly specific leukodystrophy phenotype and comprehensive panel coverage. This illustrates how laboratories differ in their policies when weighing phenotype specificity against genotype. *In silico* evidence (PP3) was also capped at supporting strength in several laboratories, despite sequence variant interpretation guidance[28] permitting moderate strength, reflecting ongoing caution around the use of predictive algorithms.

For variant 1, PM5 was recommended on the basis of other pathogenic nonsense variants at the same residue, yet no laboratory applied this criterion. Although omission did not incur a deduction, it highlights continuing uncertainty about whether PM5 is appropriate outside its original scope as a missense-level criterion. Although the rationale for PM5 is that alternative substitutions at the same residue are more likely deleterious, its use for truncating variants is not universal. Some ClinGen Expert Panels [eg, *CDH1* (ClinGen CDH1 Expert Panel Specifications to the ACMG/AMP Variant Interpretation Guidelines Version 3.1; *https://clinicalgenome.org/docs/clingen-cdh1-expert-panel-specifications-to-the-acmg-amp-variant-interpretation-guidelines-version-3.1*, last accessed October 4, 2025) allow a supporting version of PM5 under tightly defined conditions, such as predicted nonsense-mediated decay or upstream position relative to known pathogenic truncations. However, because truncating variants are already captured by PVS1, applying PM5 in this context risks double-counting unless carefully constrained.

These findings show how different guideline versions, local policy decisions, and divergent interpretations of ambiguous criteria contribute to inter-laboratory variability. They also reinforce the importance of harmonization in the adoption of evolving frameworks and the role of EQA in driving convergence.

Discordance between the use of specific ACMG/AMP criteria is not unique to this pilot; a recent study showed concordance rates of 27% to 72% (average 46%) between laboratories classifying a set of 10 variants.[29] To address these differences, the scoring criteria could be modified to include general requirements for evidence types rather than relying strictly on specific ACMG/AMP classification codes. This approach would allow for a more adaptable

assessment that accommodates alternative frameworks while maintaining rigorous standards.

A point of contention in the scoring criteria was whether to deduct points for not detailing assay methods and limitations when a pathogenic variant was identified. While the criteria stipulated deductions only when no pathogenic variant was detected, it can still be valuable for laboratories to specify how variants were prioritized and what strategies or filters were employed during data analysis. These strategies may differ significantly, such as using virtual gene panels that restrict broad genomic investigations to specific gene subsets. Including these details in reports provides essential context, even if the expected pathogenic variants are found.

Scoring consistency between expert and nonexpert assessors was generally high in this pilot; however, challenges emerged in achieving uniformity across all evaluators. This variability underscored the need for a more detailed and structured marking rubric to guide assessors on when and how to deduct points. Clear, explicit instructions would not only help nonexpert assessors apply the criteria consistently but also promote standardized scoring across the board. This is particularly crucial given the diverse reporting practices observed across laboratories.

The working group considered the inclusion of competent nonexpert assessors as a key element for scaling the program, addressing the anticipated shortage of expert resources in diagnostic genomics, both nationally and internationally (Australasian Professional Genetic Workforce Census Reports; *https://hgsa.org.au/Web/Web/HP-Resources/Australasian-Professional-Genetic-Workforce-Census-Reports.aspx*, last accessed October 4, 2025).[30] Leveraging nonexpert assessors expands the pool of available evaluators, while establishing an escalation pathway to expert assessors ensures high standards are maintained in critical cases, such as discordant scores or laboratory appeals, without overburdening scarce expert resources.[31]

A key challenge for comprehensive end-to-end quality assurance programs is the reliance on sourcing biological specimens, especially given the rarity of certain diseases and the need for large quantities of DNA. As the number of laboratories enrolling in such programs grows, so will the demand for biological material, which may present logistical and ethical challenges, such as obtaining consent and dealing with the potential for previously undetected findings. Data-only programs, which can be derived from biological specimens but distributed digitally, offer a scalable alternative. By reducing reliance on physical samples, these programs can alleviate some of the burden on laboratories and lower costs of participation while maintaining high standards of assessment.

The vGermline module by EMQN assesses analytical performance for NGS-based assays, currently focusing on single nucleotide variants and indels $<50$ bp, with a pilot module for copy number variation detection underway. Future expansions could include techniques such as repeat expansion, structural variant detection, and mtDNA sequencing, covering the full scope of NGS capabilities. Although valuable for systematically assessing proficiency across multiple genes, these technical programs lack interpretive aspects and should complement, not replace, disease-specific programs, as each provides unique insights into laboratory performance.[32]

Further, a nonsynchronous EQAP model where analytical and disease-specific modules are applied separately could allow laboratories that perform only part of the NGS workflow to participate. For example, some laboratories may outsource the sequencing phase to external service providers due to limitations in capacity or resources, focusing primarily on interpreting the data returned. In such cases, laboratories conducting the sequencing can be evaluated through an analytical or technical module, while laboratories tasked with interpreting the data could participate in a separate data-only interpretation module. This separation ensures that laboratories are assessed based on their particular role within the testing pipeline, creating a more tailored and relevant assessment of their capabilities.

Traditional methods of quality control, which rely on patient samples or cell lines, also face several challenges. These include the risk of culturing artefacts in cell lines and complex consent requirements. The use of synthetic quality control materials, such as those generated through recombinant DNA techniques or oligonucleotide synthesis, presents a cost-effective solution, particularly for multiplex or panel-based assays. However, synthetic materials may not always fully replicate the characteristics of actual patient samples, potentially limiting their use in specific EQAP contexts.[33]

The next phase of this program will focus on addressing one of the key limitations identified during the pilot; the reliance on previously reported variants, which may limit the assessment of a laboratory's ability to identify novel variants. To overcome this, a second pilot round is currently underway using *in silico* spiked-in variants, where a novel variant is computationally inserted into a composite sample derived from the 1000 Genomes Project.[34] This approach will allow for a genuinely blinded assessment of variant prioritization and classification and provide an opportunity to evaluate laboratories' ability to interpret previously unseen variants. Using *in silico* data derived from biological samples may overcome some of the previously mentioned challenges. While the use of *in silico* data in EQAP is not new, with the exception of the CAP NGS undiagnosed disorders-exome module, previous efforts have focused primarily on analytical performance rather than clinical interpretation.[35] Following the next phase, it is intended that this module will be offered through the RCPAQAP as an annual program both within and outside of Australia.

## Conclusion

This study highlights the need for flexible, scalable EQA programs to accommodate the growing diversity of

genomic testing methods and disease states. Data-only interpretive modules offer a promising avenue for achieving this, allowing for both cost reduction and improved accessibility. However, ensuring the robustness of these programs will require ongoing refinement of assessment frameworks, incorporation of synthetic data, and careful attention to the standardization of laboratory practices and reporting.

## Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the authors used ChatGPT 4.0 (OpenAI, San Francisco, CA) in order to suggest alternative wording and reduce word count. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Author Contributions

All the authors conceptualized the study and designed the methodology; M.A.H., A.S., S.Y.C., and B.B. administered the project; J.C., A.B.B., and B.A.T. curated the data; K.S.K., D.A., B.A.T., A.B.B., S.Y.C., M.A.H., and A.S. analyzed the data; S.Y.C. and B.A.T. validated the data; J.C., T.B., and B.B. supervised the project; B.L. wrote the original manuscript; and K.S.K., B.B., S.Y.C., S.L., J.C., B.A.T., T.B., M.A.H., A.S., and D.A. reviewed and edited the manuscript.

## Disclosure Statement

T.B. and S.Y.C. are employed by RCPAQAP.

## Supplemental Data

Supplemental material for this article can be found at *https://doi.org/10.1016/j.jmoldx.2025.11.006*.

## References


1. Wigby KM, Brockman D, Costain G, Hale C, Taylor SL, Belmont J, Bick D, Dimmock D, Fernbach S, Greally J, Jobanputra V, Kulkarni S, Spiteri E, Taft RJ: Evidence review and considerations for use of first line genome sequencing to diagnose rare genetic disorders. NPJ Genom Med 2024, 9:15
2. Wojcik MH, Lemire G, Berger E, Zaki MS, Wissmann M, Win W, et al: Genome sequencing for diagnosing rare diseases. N Engl J Med 2024, 390:1985−1997
3. Record CJ, Reilly MM: Lessons and pitfalls of whole genome sequencing. Pract Neurol 2024, 24:263−274
4. Kolmogorov M, Billingsley KJ, Mastoras M, Meredith M, Monlong J, Lorig-Roach R, Asri M, Alvarez Jerez P, Malik L, Dewan R, Reed X, Genner RM, Daida K, Behera S, Shafin K, Pesout T, Prabakaran J, Carnevali P, Yang J, Rhie A, Scholz SW, Traynor BJ, Miga KH, Jain M, Timp W, Phillippy AM, Chaisson M, Sedlazeck FJ, Blauwendraat C, Paten B: Scalable Nanopore sequencing of human genomes provides a comprehensive view of haplotype-resolved variation and methylation. Nat Methods 2023, 20:1483−1492
5. Sciacovelli LSS, Zardo L, Plebani M: The role of the external quality assessment. Biochem Med (Zagreb) 2010, 20:160−164
6. de Stefano MC, Floridia G, Censi F, Tosto F, Salvatore M, Civolani A, Crescenzi B, Giardino D, Lenzini E, Lisi E, Lonardo F, Mancini M, Novelli A, Piombo G, Stioui S, Taruscio D: The Italian National External Quality Assessment Program in Cytogenetics: 4 years of activity (2013-2016) following the introduction of poor performance criteria. Ann Ist Super Sanita 2018, 54:109−116
7. Jang MA, Yoon YA, Song J, Kim JH, Min WK, Lee JS, Lee YW, Lee YK: Effect of accreditation on accuracy of diagnostic tests in medical laboratories. Ann Lab Med 2017, 37:213−222
8. Kalman LV, Lubin IM, Barker S, du Sart D, Elles R, Grody WW, Pazzagli M, Richards S, Schrijver I, Zehnbauer B: Current landscape and new paradigms of proficiency testing and external quality assessment for molecular genetics. Arch Pathol Lab Med 2013, 137:983−988
9. Centers for Medicare & Medicaid Services: Clinical Laboratory Improvement Amendments (CLIA). Baltimore, MD, Centers for Medicare & Medicaid Services, 1988
10. International Organization for Standardization: Medical Laboratories — Requirements for Quality and Competence. Geneva, Switzerland, International Organization for Standardization, 2022
11. National Pathology Accreditation Advisory Council: Requirements for Medical Pathology Services. Canberra, Australia, Australian Government Department of Health, 2018
12. Schobers G, Derks R, den Ouden A, Swinkels H, van Reeuwijk J, Bosgoed E, Lugtenberg D, Sun SM, Corominas Galbany J, Weiss M, Blok MJ, Olde Keizer RACM, Hofste T, Hellebrekers D, de Leeuw N, Stegmann A, Kamsteeg E-J, Paulussen ADC, Ligtenberg MJL, Bradley XZ, Peden J, Gutierrez A, Pullen A, Payne T, Gilissen C, van den Wijngaard A, Brunner HG, Nelen M, Yntema HG, Vissers LELM: Genome sequencing as a generic diagnostic strategy for rare disease. Genome Med 2024, 16:32
13. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL: 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 2015, 17:405−424
14. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN; REDCap Consortium: The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019, 95:103208
15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG: Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009, 42:377−381
16. Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, Leong IUS, Smith KR, Gerasimenko O, Haraldsdottir E, Thomas E, Scott RH, Baple E, Tucci A, Brittain H, de Burca A, Ibañez K, Kasperaviciute D, Smedley D, Caulfield M, Rendon A, McDonagh EM: PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. Nat Genet 2019, 51:1560−1565
17. Tudini E, Andrews J, Lawrence DM, King-Smith SL, Baker N, Baxter L, et al: Shariant platform: enabling evidence sharing across Australian clinical genetic-testing laboratories to support variant interpretation. Am J Hum Genet 2022, 109:1960−1973
18. Tan TY, Lunke S, Chong B, Phelan D, Fanjul-Fernandez M, Marum JE, Kumar VS, Stark Z, Yeung A, Brown NJ, Stutterd C, Delatycki MB, Sadedin S, Martyn M, Goranitis I, Thorne N, Gaff CL, White SM: A head-to-head evaluation of the diagnostic efficacy and costs of trio versus singleton exome sequencing analysis. Eur J Hum Genet 2019, 27:1791−1799
19. Payne DA, Baluchova K, Russomando G, Ahmad-Nejad P, Mamotte C, Rousseau F, van Schaik RHN, Marriott K, Maekawa M,

Chan KCA: Toward harmonization of clinical molecular diagnostic reports: findings of an international survey. Clin Chem Lab Med 2018, 57:78−88
20. Nisselle A, Janinski M, Martyn M, McClaren B, Kaunein N; Reporting Item Standards for Education and its Evaluation in Genomics Expert Group, Barlow-Stewart K, et al: Ensuring best practice in genomics education and evaluation: reporting item standards for education and its evaluation in genomics (RISE2 Genomics). Genet Med 2021, 23:1356−1365
21. Rehder C, Bean LJH, Bick D, Chao E, Chung W, Das S, O'Daniel J, Rehm H, Shashi V, Vincent LM: Next-generation sequencing for constitutional variants in the clinical laboratory, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). Genet Med 2021, 23:1399−1415
22. National Pathology Accreditation Advisory Council: Requirements for Human Genetic Variation. Canberra, Australia, National Pathology Accreditation Advisory Council, Commonwealth of Australia, 2013
23. National Pathology Accreditation Advisory Council: Requirements for Massively Parallel Sequencing (MPS). Canberra, Australia, National Pathology Accreditation Advisory Council, Commonwealth of Australia, 2017
24. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, Kobayashi Y, Patil N, Thusberg J, Westbrook M, Invitae Clinical Genomics Group, Topper S: Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. Genet Med 2017, 19: 1105−1117
25. Ellard S, Baple EL, Callaway A, Berry I, Forrester N, Turnbull C, Owens M, Eccles DM, Abbs S, Scott R, Deans ZC, Lester T, Campbell J, Newman WG, Ramsden S, McMullan DJ: ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020. London, UK, Association for Clinical Genomic Science, 2020
26. Durkie M, Cassidy E-J, Berry I, Owens M, Turnbull C, Scott RH, Taylor RW, Deans ZC, Ellard S, Baple EL, McMullan DJ: ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2024. London, UK, Association for Clinical Genomic Science, 2024
27. ClinGen Sequence Variant Interpretation Working Group: ClinGen Sequence Variant Interpretation Recommendation for PM2 − Version 1.0. Bethesda, MD, Clinical Genome Resource (ClinGen), 2020
28. Pejaver V, Byrne AB, Feng BJ, Pagel KA, Mooney SD, Karchin R, O'Donnell-Luria A, Harrison SM, Tavtigian SV, Greenblatt MS, Biesecker LG, Radivojac P, Brenner SE; ClinGen Sequence Variant Interpretation Working Group: Calibration of computational tools for missense variant pathogenicity classification and ClinGen recommendations for PP3/BP4 criteria. Am J Hum Genet 2022, 109: 2163−2177
29. Houge G, Bratland E, Aukrust I, Tveten K, Žukauskaitė G, Sansovic I, et al: Comparison of the ABC and ACMG systems for variant classification. Eur J Hum Genet 2024, 32:858−863
30. Maiese DR, Lyon M, Reddi HV, Blitzer MG, Bodurtha JN, Muenke M: The 2019 medical genetics workforce: a focus on laboratory geneticists. Genet Med 2023, 25:100834
31. Richards CS, Grody WW: Alternative approaches to proficiency testing in molecular genetics. Clin Chem 2003, 49:717−718
32. Schrijver I, Aziz N, Jennings LJ, Richards CS, Voelkerding KV, Weck KE: Methods-based proficiency testing in molecular genetic pathology. J Mol Diagn 2014, 16:283−287
33. Chen B, O' Connell CD, Boone DJ, Amos JA, Beck JC, Chan MM, et al: Developing a sustainable process to provide quality control materials for genetic testing. Genet Med 2005, 7:534−549
34. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al: A global reference for human genetic variation. Nature 2015, 526:68−74
35. Duncavage EJ, Abel HJ, Merker JD, Bodner JB, Zhao Q, Voelkerding KV, Pfeifer JD: A model study of in silico proficiency testing for clinical next-generation sequencing. Arch Pathol Lab Med 2016, 140:1085−1091