Genetics in Medicine (2024) 26, 101116

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Determining priority indicators of utility for genomic testing in rare disease: A Delphi study

Zoe Fehlberg[1,2], Ilias Goranitis[1,2], Andrew J. Mallett[1,3,4,5], Zornitza Stark[1,2,6], Stephanie Best[1,2,7,8,*]

[1]*Australian Genomics, Melbourne, VIC, Australia;* [2]*University of Melbourne, Melbourne, VIC, Australia;* [3]*College of Medicine and Dentistry, James Cook University, Douglas, QLD, Australia;* [4]*Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD, Australia;* [5]*Department of Renal Medicine, Townsville University Hospital, Douglas, QLD, Australia;* [6]*Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Melbourne, VIC, Australia;* [7]*Department of Health Services Research, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia;* [8]*Victorian Comprehensive Cancer Centre Alliance, Melbourne, VIC, Australia*

ARTICLE INFO

Article history:
Received 25 October 2023
Received in revised form
1 March 2024
Accepted 4 March 2024
Available online 6 March 2024

Keywords:
Delphi methodology
Genomic testing
Rare disease
Utility
Value assessment

ABSTRACT

Purpose: Determining the value of genomic tests in rare disease necessitates a broader conceptualization of genomic utility beyond diagnostic yield. Despite widespread discussion, consensus toward which aspects of value to consider is lacking. This study aimed to use expert opinion to identify and refine priority indicators of utility in rare disease genomic testing.

Methods: We used 2 survey rounds following Delphi methodology to obtain consensus on indicators of utility among experts involved in policy, clinical, research, and consumer advocacy leadership in Australia. We analyzed quantitative and qualitative data to identify, define, and determine priority indicators.

Results: Twenty-five experts completed round 1 and 18 completed both rounds. Twenty indicators reached consensus as a priority in value assessment, including those relating to prognostic information, timeliness of results, practical and health care outcomes, clinical accreditation, and diagnostic yield. Whereas indicators pertaining to discovery research, disutility, and factors secondary to primary reason for testing were considered less of a priority and were removed.

Conclusion: This study obtained expert consensus on different utility indicators that are considered a priority in determining the value of genomic testing in rare disease in Australia. Indicators may inform a standardized approach to evidence generation and assessment to guide future research, decision making, and implementation efforts.

The Article Publishing Charge (APC) for this article was paid by Andrew J. Mallett.

Zornitza Stark and Stephanie Best are joint senior authors.

*Correspondence and requests for materials should be addressed to Stephanie Best, The University of Melbourne, Grattan Street, Parkville, Victoria, 3010, Australia. *Email address:* stephanie.best@unimelb.edu.au

doi: https://doi.org/10.1016/j.gim.2024.101116

Introduction

For some time, there has been an appreciation that a broader conceptualization of genomic test utility exists.[1,2] The American College of Medical Genetics and Genomics (ACMG) describes utility beyond outcomes pertaining to death and serious disease or disability, to include effects on clinical management, prognostic implications, benefits of the information for patients and their family, and the cost impact on health care systems.[3] In rare disease, (defined as a disease with a prevalence of <1 in 200,000 in the United States or <1 in 2000 in the European Union), genomic testing is primarily used as a diagnostic tool, with emerging applications in screening, such as newborn and reproductive *carrier* screening. Typically characterized by early onset, rare diseases greatly impact affected individuals and families and require substantial and sustained health care. Further, owing to the heterogeneity of rare diseases and diverse health and non-health impacts, pinpointing a singular test outcome to define utility is challenging. Despite this, diagnostic yield is the most frequently reported indicator of clinical value.[4] Given the acceleration of genomic testing in rare disease, it is essential that the broader aspects of utility are considered and reported consistently to optimize the implementation and realization of the benefits of genomic testing for patients, families, and health systems.[5-8]

Research demonstrating the shift toward comprehensive valuations of genomic utility in rare disease[9-11] and inclusion in economic assessment[12-14] has enabled cost-benefit analyses that incorporates aspects of diagnostic, clinical, and non-clinical value across patient populations. Increasingly, evidence generated through such approaches has been used to inform health care priorities and facilitate funding approval decisions.[14,15] Although advantageous in quantifying the value of genomics for specific rare disease conditions and types of genomic medicine, such approaches are limited in providing a generic framework for assessing the value of genomic medicine. Likewise, evidence-based valuation processes such as health technology assessment (HTA) agencies, which are used by health care payers to guide policy and reimbursement decisions[16] have highlighted the need to work within a standardized approach so that genomic medicine is not disadvantaged by funding application tools.[17] Currently limiting a standardized approach to generating evidence of overall utility is the lack of agreement toward which indicators of utility should be captured and how to summate the disparate measures.[18]

Other efforts to standardize the reporting of utility have led to the development and validation of tools to separately measure clinical and personal utility. The Clinician-reported Genetic testing Utility InDEx C-Guide acts as a checklist for clinicians to report perceived clinical utility, or the changes in clinical management and improvements in health outcomes, from genomic sequencing.[19-21] The Personal Utility scale measures the personal utility, or the value of a genomic result beyond a change in management perceived by patient and families.[22] Although both act as valuable data collection tools, the problem remains as to how to consider utility through a singular lens[23] and how to include stakeholders at various levels, such as health payers and researchers.

Conceptual frameworks have been proposed as an approach to cohesively determine overall utility. Smith et al conceived a patient-centered model and identified clinical and non-clinical domains of patient perceived utility.[24] Although this approach integrates clinical and personal impacts, further dimensions of utility can be considered, including the extent of impact, the relationship and perspective of utility, and the time horizon.[25] Through such multidimensional frameworks, it may be possible to form a cumulative point of utility that better reflects all the possible aspects of genomic testing utility in rare disease. To effectively utilize such frameworks, further refinement of domains and indicator measures is needed. Capturing a holistic consensus on what matters is essential given the role clinicians and researchers play in evidence generation and the impact on consumers. In this study, we aimed to use expert opinion from a wide range of key stakeholders to identify and define priority indicators of utility for rare disease genomic testing in Australia.

Materials and Methods

Study design

We used the Delphi technique, which is a multi-stage process to obtain group insight toward a current or future challenge and examine levels of consensus among participating experts.[26,27] The process comprises of iterative "rounds," whereby participants are asked to respond to aggregated results from previous rounds until a sufficient level of consensus is reached. This study involved 2 survey rounds of prioritization and revision conducted between January 2023 and April 2023. The study was approved by Royal Children's Hospital Melbourne, Human Research Ethics Committee (HREC/89720/RCHM-2022). Participation was voluntary and implied consent indicated by survey commencement after the opportunity to view participant information material.

Participants and recruitment

A diverse range of experts in the field of genomic testing for pediatric and adult-onset rare disease were purposively selected through the extensive national *Australian Genomics*[5] network consisting of more than 100 organizations, including diagnostic laboratories, clinical genetics services, and research and academic institutions. Experts were identified for their leadership (ie, are a director, chair, lead, or head of department) within a health-service, state or federal government health department, HTA, research, or their

executive management of research. The patient and/or caregivers of patients' voices were represented through inviting leaders of consumer advocacy groups, many of whom have lived experience of rare disease and are able to represent the interests of the rare disease community more broadly while providing strategic level insight. We sought individuals with the appropriate role, expertise, or experience with funding decisions for genomic testing in rare diseases, including individuals with professional backgrounds in rare disease clinical genetics, laboratory genetics, research, advocacy, and policy. Potential participants were invited via email, with 2 follow-up reminders as needed. If unable to participate, individuals were encouraged to recommend a suitable replacement. No additional participants were sought beyond the first round.

Delphi development

A modified Delphi method was used to develop round 1, which comprised 37 predetermined utility indicators housed within 9 domains drawn from a review of the literature and the study team's experience. We thought of domains as high-level categories used to group relative indicators and indicators as a practical item that can be assessed to gauge utility. The search strategy reflected our aim to build upon existing frameworks while incorporating current dialog around the broader conceptualization toward genomic testing utility and measurement of indicators. To achieve this, secondary articles were sought except for previously conducted Delphi studies because we considered them a rich source of information. One member of the study team (Z.F.) applied the relevant article type filters (eg, reviews, systematic reviews, editorials, and commentary) with the appearance of the terms for genomic testing (eg, genetic testing, exome sequencing, and genetic screening) and utility (eg, utility and value) to articles indexed in PubMed (MEDLINE 2012-October 2022). Empirical studies reporting utility outcomes and articles related to genomic testing in cancer or common diseases were excluded. The search was supplemented by reviewing the reference lists of found articles and keyword and key authors searches in Google Scholar and prominent genetics journals. Examples of search strategies are shown in Supplemental Methods 1. Article characteristics, utility domains, and indicators were abstracted for 27 articles (eg, Grosse,[2,] Hayeemz et al,[19] and Smith et al[23]) with the entire list presented in Supplemental Data 1. Members of the study team with extensive knowledge of genomic testing in Australia (I.G., A.J.M., Z.S., and S.B.) revised the list and incorporated additional indicators derived from professional experience. Indicators were then categorized within domains of utility and descriptions toward their measurement generated. The list was discussed among the study team until agreement was reached (Supplemental Data 2).

After analyzing the round 1 data, the entire study team reviewed the revised list of indicators through iterative discussion to develop round 2. This included the recategorization of indicators within domains, merging and splitting indicators, and incorporating suggested refinement. Modifications to the survey are presented in the Results and Supplemental Material 5.

Delphi rounds

Two Delphi rounds were completed online using surveys hosted in the Research Electronic Data Capture (REDCap) database platform. During development, the survey instruments were pre-tested for content and usability by 2 external staff members with expertise in REDCap and survey design. The complete instruments can be found in Supplemental Methods 2. We stopped after 2 rounds because a sufficient level of consensus was met for most indicators or were trending downward in their agreement fraction and feedback on modifications to definitions was minimal.

Round 1

Consisted of 3 parts and was available to complete over 1 month. Part 1 included a series of demographics questions. Part 2 was organized into domain tables that housed indicators. For each indicator, it was mandatory for participants to accept or modify the name and description; if selected, the participant was prompted to suggest their modification. Next, participants were asked to select whether they would consider the measure to be a "top priority," "high priority," "low priority," "not a priority," or "unsure" in the measurement of rare-disease genomic testing utility, including diagnostic testing and screening contexts. Participants could comment or justify their decision and propose additional indicators. In part 3, participants were asked if they had any domain level modifications.

Round 2

Consisted of 2 parts and was available to complete over 2 weeks. Part 1 followed the same format as round 1, except participants were only asked to rate the priority of indicators that did not reach consensus in the previous round or were reworked or newly included. Part 2 presented the list of indicators included in round 2 and asked participants to select their top 5 across all the domains.

Analysis

Each round generated quantitative and qualitative data that were downloaded from REDCap and analyzed. Descriptive statistics were performed on all quantitative variables. Consistent with previous approaches,[28-30] items prioritized as a "top" or "high" priority by ≥75% of participants were included, if between 51% and 75% they were considered undecided and included in round 2 for further review, and items with ≤50% were excluded. A stricter inclusion criterion was applied to round 2 with indicators removed if the

agreement fraction was <75%. Proposed indicators were added if suggested by 2 or more participants. Incorporating suggested modifications to indicator names and definitions was a process of editing for accuracy and comprehensibility, which did not require multiple instances for the change to be made. General comments and justifications were analyzed using conventional content analysis, looking for patterns both within and across indicators and resulted in merging indicators, rehousing them within other domains, or understanding participant decision making. All changes were discussed at regular study team meetings before being incorporated. Statistics were computed in Stata SE version 17 and qualitative data managed using spreadsheets.

Results

Characteristics of participants

Of the 45 invited experts, 25 completed round 1 (response rate of 55.55%), and 18 (response rate of 72%) completed both rounds. Two individuals declined and suggested an alternative, and the rest did not respond. Participants were from a range of leadership areas shown in Table 1 alongside demographic characteristics. The lowest initial response rate was individuals with leadership roles within state or federal government health departments (1/7).

Table 1 Participant characteristics

Characteristic, *n* (%)	Round 1 (*n* = 25)	Round 2 (*n* = 18)
Leadership area		
Clinic/health service	7 (28)	2 (11)
State or Federal Health Department	1 (4)	1 (5.5)
Health technology assessment	2 (8)	1 (5.5)
Patient/consumer advocacy	3 (12)	3 (17)
Research	8 (32)	7 (39)
Executive management/advisory	4 (16)	4 (22)
Location in Australia		
Australian Capital Territory	1 (4)	0 (0)
New South Wales	8 (32)	7 (39)
Queensland	1 (4)	1 (5.5)
South Australia	1 (4)	0 (0)
Tasmania	1 (4)	1 (5.5)
Victoria	11 (44)	8 (44.5)
Western Australia	2 (8)	1 (5.5)
Age (years)		
18-30	0 (0)	0 (0)
31-40	3 (12)	2 (11)
41-50	8 (32)	7 (39)
51-60	10 (40)	6 (33)
61-70+	4 (16)	3 (17)
Gender		
Male	9 (36)	5 (28)
Female	16 (64)	13 (72)
Experience (years)		
0-5	0 (0)	0 (0)
6-10	4 (16)	4 (22)
11-20	8 (32)	5 (28)
20+	13 (52)	9 (50)

Delphi results

The results from the 2 survey rounds are presented in Table 2, and Table 3 contains the final list of priority indicators and descriptions.

Priority indicators

After 2 rounds, 20 out of the 40 indicators reached consensus as a priority in the measurement of genomic testing utility in rare disease (Table 2). High agreement was met for "Prognostic Information" (96%), "Timely Result" and "Practical Outcomes" (95%), "Clinical Accreditation," "Health Care Process Outcomes," and "Health Outcomes" (92%). Interestingly, high-priority agreement did not necessarily translate to "Top 5" selection available in round 2. "Diagnostic Yield (Symptomatic and Screening)" (61%) was the most frequently selected top 5 followed by "Clinical Accreditation," and "Medication and Medical Device Management (Alters Outcomes)" (55.5%). "Health Outcomes" (50%) and "Cost Impact on Health Care" (44%) also featured highly. Top 5 justifications (n = 8) included remarks such as "*I rated outcomes that improve QoL (Quality of Life) for the person diagnosed as highest priority, ie, patient centered. There were many similar ones that I could have put in top 5, difficult to pick!*" (Round 2, P26).

Upon analysis of qualitative data, 4 indicators that reached consensus in round 1 were merged because they were considered to overlap: "Prognostic Information for Families" merged with "Practical Outcomes," "Decreasing Patient's Need for High Acuity, Urgent, or Emergency Care" merged with "Health Care Process Outcomes," "Affective Utility" merged with "Health Outcomes," and "Cost-effectiveness" merged with "Cost Impact on Health Budget." Although "Timeliness of Diagnosis" reached consensus in round 1 (80%), it was reworked into "Timely Result" with substantial changes to the description. The changes were made to better capture the impact of the result on clinical decision making and therefore required reassessment in round 2 in which it then reached higher consensus (95%).

Overall, 12 indicators did not reach consensus as a priority and were excluded. Four in round 1 (Request Further Testing, Actionable Secondary Findings, Creation of Large Data Sets to Enable Commercial Research, and Societal Acceptability) and 8 following round 2 (Actionable Inconclusive Result, Disutility, Improving Diagnostic Outcomes for Future Patients, Creation of Large Data Sets to Enable a Broad Range of Health-Related Research, Impact on

Table 2 Final list of indicators and descriptions with agreement percentage and number of appearances in top 5 selections

Indicator	Description	Agreement %	Top 5, *n*
Domain: Test Performance			
Clinical Accreditation (Analytical Validity and Clinical Validity)	The evidence is available to show how well the test (a) detects the genotype of interest accurately and reliably and (b) predicts the clinical disorder or phenotype associated with the genotype.	92	10
Diagnostic Yield (Symptomatic and Screening)	% of the tested population who either (a) receive a genomic diagnosis when testing individuals with an existing clinical disorder or (b) receive a genomic diagnosis in a screening context, *carrier* testing, or targeted testing context and, from those, the proportion who develop a disease phenotype.	84	11
Domain: Clinical			
Medication and Medical Device Management (Alters Outcomes)	% of the tested population for whom the result changes medication/or device utilization that are expected to alter health outcomes. For example, starting or stopping medications, access to funded precision treatments, or registered but unfunded treatments.	88	10
Medication and Medical Device Management (Symptoms of the Condition)	% of the tested population for whom the result changes medication/or device utilization for symptoms of their condition or disease progression.	88	1
Non-Medication Management	% of the tested population for whom the result changes access to symptomatic treatments/therapies with allied health (eg, occupational therapy, physiotherapy, and speech pathology).	84	3
Health Care Process Outcomes	% of the tested population for whom the result changes the process of medical care. For example, additional presymptomatic monitoring, referrals for management/support and surveillance (eg, genetic counselor, additional specialist[s]), transitions in care (eg, long-term supported and palliative care), hospitalization events, length of stay in hospital/ intensive care, surgery or organ transplantation decisions, or need for high-acuity, urgent/emergency care.	92	2
Health Outcomes	% of the tested population for whom the result affects clinical outcomes directed by a change in management. For example, risk management strategies to reduce risk of disease, patient family reported quality of life, prevention of disability, or prolonging of life.	92	9
Timely Result	% of the tested population who receive the result within a clinically appropriate timeframe to inform decision making, including changes in management.	80	2
Avoid Further Investigation	% of the tested population for whom the result negates the need for further simple, complex, invasive (eg, muscle biopsy) testing, and/or monitoring.	84	2
Access to Clinical Trial(s)	% of the tested population for whom the result enables entry into clinical trial(s).	76	2
Diagnostic Purpose	% of the tested population who receive a genomic diagnosis that either (a) cannot be reached by other investigations or (b) confirm, amend, subclassify, or refute existing phenotypic diagnoses or (c) rules out diagnoses.	80	3

(continued)

Table 2 Continued

Indicator	Description	Agreement %	Top 5, *n*
Domain: Prognostic			
Prognostic Information	% of the tested population who receive a change of prognosis (general, clarified, and precise) without any change in treatment. For example, information on natural history, related conditions, severity and likelihood of symptoms, age of onset, future health risks, and risk of death.	96	1
Domain: Individual and Family			
Establish Accurate Recurrence Risk	% of the tested population for whom the result establishes accurate recurrence risk for themself and biological family members.	84	2
Access to Cascade Testing	% of the tested population with a genomic diagnosis for whom the result enables access to cascade testing in the biological family to identify other at-risk individuals.	88	1
Knowledge Outcomes	% of the tested population for whom the result impacts individuals' knowledge about oneself or child and feelings toward the use of gained knowledge.	89	1
Practical Outcomes	% of the tested population for whom the result affects practical outcomes for individuals and or families. For example, future autonomy, planning and contribution (financial, employment, social, and education) access to support (eg, mental health support, support groups, and day-to-day living supports), and improved communication with family or child.	95	2
Reproductive Planning	% of the tested population for whom the result affects reproductive planning (eg, PND, IVF, PGT-M, and CVS/amniocentesis) and number of children.	88	4
Domain: Economic			
Cost impact on health care	Measuring financial effects on health budgets of delivering or not delivering the test. For example, impacts on length of hospital stay, diagnostic test volumes and costs, value of information analysis, and workforce and productivity outcomes.	80	8
Cost impact on the individual or family	Measuring direct and indirect financial effects of having or not having the test and impact of the test result on individuals and families. For example, the cost of testing, cost of medical care, and workforce participation.	76	0
Domain: Societal			
Accessibility	Measuring whether access to testing and associated medical care or counseling services are likely to be equitable across socioeconomic, geographical, and cultural groups.	89	4

Clinical Trial Outcomes, Cost Impact on the Broader Economy, Societal Disutility, and Evidentiary Confidence). There was general concordance in the comments as to why indicators were rated as a "low" or "not a priority," for example, the comments or justifications as to why "Actionable Secondary Findings" was considered a low priority included "*utility should be limited to the condition being investigated*" (Round 1, P4) or "*not the purpose of the test, so I wouldn't measure the value of the test by it*" (Round 1, P9) and "*useful side-effects of the test but not essential*" (Round 1, P32). However, there was some discordance for indicators within the "Discovery (research)" domain, for example, "Creation of Large Data Sets to Enable Commercial Research" was rated low (47%), albeit 1 participant selected it as a top priority reasoning "*This is the only pathway to affordable treatment"* (Round 1, P24).

Table 3 Results following 2 Delphi rounds and exemplar justifications

Indicators	Round 1, % (n = 25)	Round 2, % (n = 18)	Top 5, n (n = 18)	Outcome	Exemplar Justification
Domain: Test Performance					
Clinical Accreditation (Analytical Validity and Clinical Validity)	92	-	10	Included	*I feel this is less important in my world, where there is lots of research and really high-level domain-specific expertise. But in the broader setting we really need to be sure the results of genetic testing are correct and not causing harm.* (Round 1, Top Priority, P40)
Diagnostic Yield (Symptomatic and Screening)	84	-	11	Included	*But depends, sometimes an exclusion is required clinically ie, excluding a genetic diagnosis with a prognostic management or surveillance impact.* (Round 1, High Priority, P9)
Domain: Clinical					
Medication and Medical Device Management (Alters Outcomes)	88	-	10	Included	*Super important but perhaps rarer examples of genetics having impact in this way*. (Round 1, High Priority, P40)
Medication and Medical Device Management (Symptoms of the Condition)	88	-	1	Included	
Non-Medication Management	-	84	3	Included	*For many this will be the only pathway.* (Round 2, High Priority, P24)
Health Care Process Outcomes	92	-	2	Included	*I really would like a "medium priority" option. Selected low because sometimes this just isn't relevant.* (Round 1, Low Priority, P31)
Decreasing Patient's Need for High Acuity, Urgent or Emergency Care	88	-	-	Merged[a]	*I think this is high priority but very specific to certain (maybe rarer) situations, ie, not common but very impactful if you can affect change here.* (Round 1, High Priority, P40)
Transitions in Care	60	-	-	Merged[a]	*Of value but less so than others.* (Round 1, Low Priority, P36)
Diagnostic Purpose	80	-	3	Included	*I see this one as a measure of value of the test - is the information gathered of value to the patient and family.* (Round 1, Top Priority, P24)
Actionable Inconclusive Result	56	39	0	Excluded	*Very few would use a non-diagnostic (inconclusive) result in such a manner.* (Round 2, Not a Priority, P2)
Actionable Secondary Findings	36	-	-	Excluded	*Should not be considered when requesting diagnostic testing. Utility should be limited to the condition being investigated.* (Round 1, Not a Priority, P4)
Avoid Further Investigation	84	-	2	Included	*Nobody likes a diagnostic odyssey!* (Round 1, Top Priority, P40)
Request Further Testing	48	-	-	Excluded	*Not an issue.* (Round 1, Not a Priority, P30)
Access to Clinical Trial(s)	76	-	2	Included	*The priority depends on the disease setting, stage, and access to existing therapies.* (Round 1, High Priority, P18)
Health Outcomes	92	-	9	Included	

(continued)

Table 3 Continued

Indicators	Round 1, % (n = 25)	Round 2, % (n = 18)	Top 5, n (n = 18)	Outcome	Exemplar Justification
Timely Result	80	95	2	Included	*Top priority for those having screening or diagnostic testing for reproductive purposes or where results are used to determine treatment or management*. (Round 2, Top Priority, P32)
Domain: Prognostic					
Prognostic Information	96	-	1	Included	*This information is of huge value over time and for future patients and families. It's also very valuable for health system planning*. (Round 1, Top Priority, P24)
Prognostic Information for Family	88	-	-	Merged[a]	*The societal impacts of testing are not currently considered in HTA assessments*. (Round 1, Low Priority, P18)
Domain: Individual & Family					
Practical Outcomes	72	95	2	Included	*Again, this can be the only area of life that individuals and families feel they have some control over*. (Round 2, High Priority, P24)
Knowledge Outcomes	56	89	1	Included	*Important for personal utility*. (Round 2, High Priority, P37)
Establish Accurate Recurrence Risk	84	-	2	Included	*Level of priority depends on the life-stage of the patient and their relevant family members*. (Round 1, Unsure, P35)
Access to Cascade Testing	88	-	1	Included	*This will allow families to make informed decisions impacting their future*. (Round 1, Top Priority, P24)
Reproductive Planning	88	-	4	Included	*While this is important to capture, it's unclear whether a test that has a high utility of patients accessing reproductive options means it has high utility as this can also be impacted by patient values and access*. (Round 1, Unsure, P35)
Reproductive Outcomes	76	-	-	Merged[a]	
Affective Utility	76	-	-	Merged[a]	*A bit unpredictable - other issues too - worry over family members who are affected*. (Round 1, Low Priority, P30)
Access To Social or Community Support	68	-	-	Merged[a]	*This provides a measure of the role of different support mechanisms across the life course - it can inform planning, service and support development and practice, coordination, reduce duplication etc*. (Round 1, Top Priority, P24)
Disutility	60	50	0	Excluded	*Not an issue for my patients - when they already know that they have the disease*. (Round 1, Not a Priority, P30)

(continued)

Table 3 Continued

Indicators	Round 1, % (n = 25)	Round 2, % (n = 18)	Top 5, n (n = 18)	Outcome	Exemplar Justification
Domain: Discovery (research)					
Improving Diagnostic Outcomes for Future Patients	72	67	1	Excluded	*Important to use data that is generated, to learn from it, publish etc. Not so much a priority for an individual.* (Round 2, Low Priority, P11)
Creation of Large Data Sets to Enable a Broad Range of Health-Related Research	64	50	2	Excluded	*This is the only pathway to change*. (Round 1, Top Priority, P24)
Creation of Large Datasets to Enable Commercial Research	47	-	-	Excluded	*This is the only pathway to affordable treatment*. (Round 1, Top Priority, P24)
Impact on Clinical Trial Outcomes	60	62	0	Excluded	*This is how we will learn more effectively and efficiently and have more impact*. (Round 1, Top Priority, P24)
Domain: Economic					
Cost-Effectiveness	84	-	0	Merged[a]	
Cost Impact on the Broader Economy	52	50	0	Excluded	*This is a metric that is generally important to governments. Failing to capture this in genomics could be disadvantageous*. (Round 2, High Priority, P36)
Cost Impact on Healthcare	80	-	8	Included	*Struggling to see what this has to do with genomic testing, so I suspect I am missing something*. (Round 1, Unsure, P31)
Cost Impact on the Individual or Family	76	-	0	Included	
Economic Impact on Families	68	-	0	Merged[a]	*Important to understand but is not captured in HTA for reimbursement currently*. (Round 1, Not a Priority, P18)
Domain: Societal					
Societal Acceptability	48	-	0	Excluded	*For the great majority of genomic testing this is established already*. (Round 1, Low Priority, P31)
Societal Disutility	52	50	0	Excluded	*These change within subgroups of a society and over time therefore lower priority*. (Round 2, Low Priority, P32)
Accessibility	72	89	4	Included	*Gaps need to be identified that can be targeted for improvements*. (Round 2, High Priority, P37)
Evidentiary confidence					
Evidentiary Confidence	-	72	-	Excluded	*This is an important consideration if seeking government funding to support testing*. (Round 2, High Priority, P37)

Notes. A dash (-) indicates indicator was not present in survey.
[a]Further information on merged indicators can be found in Supplemental Material 5.

Following reassessment in round 2, "Creation of Large Data Sets to Enable a Broad Range of Health-Related Research" was removed (50%); however, it was selected twice as a "top 5" priority.

Of the 10 undecided indicators reassessed in round 2, 4 reached consensus for inclusion (Knowledge Outcomes, Practical Outcomes, and Accessibility), and the other 7 were excluded. "Non-Medication Management" was the only additional indicator suggested and was included in round 2, prompting 1 participant to comment "*Really like the inclusion on non-medication management*" (Round 2, P24) and reached consensus for inclusion (84%). A team decision to add "Evidentiary Confidence" was made to elicit feedback toward the level of certainty decision makers may require for the utility gains claimed, and although it scored 72%, it did not reach the threshold for final inclusion.

Although emphasizing their importance, some indicators were considered by participants to be separate from utility. For example, participants raised "Clinical Accreditation" as a precursor requirement, eg, "*The description is spot on, but I see this is as a characteristic/requirement of the test. This feature should be a given, but it isn't describing utility*" (Round 1, P36). Both "Personal Disutility" and "Societal Disutility" were reasoned to represent a decrement in utility and therefore unable to indicate utility gains, eg, "*I'm not convinced that "negative effects" of a test should be included in a model to measure utility*" (Round 2, P35). Some perceived the "Accessibility" of a genomic test an important associated indicator or something that affects utility, as 1 person put it "*equity is an important goal but should not determine whether an important test is available*" (Round 2, P11). None of the Discovery (research) indicators reached priority agreement with lines of reasoning that they are secondary gains, eg, "*I've responded unsure of priority because while there is evidence that patients cite contributing to research as a form of utility from test results, I'm not sure it should be a priority for someone undergoing clinical genomic testing*" (Round 1, P35).

Throughout, comments acknowledged the influence of a participant's area of expertise. As 1 person put it "*My research focus might be making me rate these (Discovery indicators) higher than others…*" (Round 1, P40) and at other times 1 person stated, "*I don't think I have the expert knowledge about health economics as to whether this is a priority*" (Round 1, P35). Another participant reflected on the complexity of determining priority indicators, in that indicators may be utilized for different purposes as they wrote "*this question (picking a top 5) has more than one answer depending on the type of analysis required*" (Round 2, P18). Also, recurring in the comments section was that indicators are interrelated, eg, "*the diagnostic outcome depends on the test purpose(s) and setting*" (Round 1, P18) or '*…hard to pick 5 many are inter-linked—ie, if a result is not timely, it will not help with management in a meaningful way*" (Round 2, P11).

Refinement of domains, indicators, and descriptions

All but 2 participants provided qualitative feedback in round 1 (resulting in 258 comments), and everyone contributed to round 2 (resulting in 93 comments). Directed by round 1 feedback, changes to the domains and indicators, including descriptions, were undertaken. Below we discuss examples with a complete comparison provided in Supplemental Data 3. Analysis of comments prompted the "Psychosocial" domain to be renamed "Individual and Family" and 4 of the initial proposed indicators (Access to Social or Community Support, Affective Utility, Cognitive Utility, and Behavioral Utility) were reworked into discrete indicators (Health Outcomes, Knowledge Outcomes, and Practical Outcomes). The "Reproductive" and "Investigative" domains were removed and indicators rehoused within the "Individual and Family" and "Clinical" domains. Suggested modifications to the indicator descriptions were largely about providing examples, ensuring that a range of contexts were captured and creating overall consistency in the phrasing. Round 2 resulted in fewer comments and wording modifications, with 21 of the 28 indicators having only 1 minor, for example, replacing "molecular diagnosis" with "genomic diagnosis," or no further modifications suggested.

Discussion

Providing evidence of the value of genomic testing for patients and families, the wider community, and health systems remains a challenge to implementing genomics in health care.[31,32] Our study identified 20 utility indicators within 5 domains considered by a diverse body of experts as priorities for assessing genomic testing in rare disease. Spanning diagnostic to societal, the final 5 domains used to categorize indicators align with previous conceptual models,[19,24,25] what is important to key stakeholders,[33] and evidence-based assessment requirements.[16] By using a process of iterative expert refinement, we intend for the indicators to be easily understood and operationalized among key stakeholders. Our findings move toward standardized terminology, comparable evidence generation, and consistent reporting of genomic testing. These features will assist research and evaluation and thereby support effective implementation and realization of the benefits of genomic testing in health care.

Many of the findings maintains a patient-centered approach to determining utility and likely reflects the sample who were predominantly clinical decision makers and researchers with clinical professional backgrounds. For example, the indicators that reached the highest consensus were "Prognostic Information," "Timely Result," "Practical Outcomes," "Clinical Accreditation," "Health Care Process Outcomes," and "Health Outcomes." Interestingly,

"Diagnostic Yield," although the most reported outcome in literature[34] did not have 1 of the highest agreement fractions, it was, however, the most frequently selected top 5. Future research could work to ascertain the level of diagnostic yield that is valued by stakeholders and the trade-off between diagnostic yield in the presence of other utility gains. Diagnostic yield, in the setting of our study remained centered to attaining a confirmed genetic diagnosis for a patient's/family's phenotype and did not include uncertain, partial, or secondary findings as part of the definition. Further, proposed indicators relating to these outcomes did not reach consensus, although we acknowledge that some may infer or interpret them to indirectly contribute to diagnostic yield.

In line with previous research, our participants supported the growing and welcomed inclusion of a number of patient/family measures[9,35] and other contextual factors[17,36] within overall utility. Three of the indicators in the individual-family domain "Practical Outcomes," "Knowledge Outcomes," and "Reproductive Planning" reassuringly aligned with The Personal Utility scale,[22] enabling the integration of a validated approach to measurement. "Accessibility," was the only societal level indicator that reached adequate consensus and is considered an important consideration across multiple settings in which genomic testing is implemented.[17,37] "Accessibility" was also the only indicator that was not identified in the empirical literature search that reached consensus. Although "Societal Acceptability" was removed and may be established in some settings, as the application of genomic testing expands, eliciting societal preferences remains an essential consideration in evidence generation and health care decision making.[9,38]

Similarly, many of the indicators that are indicative of the complex picture of genomic medicine did not make final inclusion. For example, we sought to build on conceptual work,[25] by gaining perspective on utility outcomes relating to discovery research. However, no indicator in this domain reached priority consensus, albeit "Improving Diagnostic Outcomes for Future Patients" and "Creation of Data sets to Enable a Broad Range of Health-Related Research" were selected as a top 5 and justified as a pathway to population health and scientific advancements.[39] Policy decisions on the large-scale implementation of genomics spans not just health care but includes innovation and research areas of government and can be seen as an investment in improving the care of future patients, for example, through gene discovery, methods development, and the development of precision treatments.[5] Likewise, no direct outcomes relating to disutility reached priority consensus and are frequently omitted in the reporting of studies.[34] As the settings in which genomic testing is applied continually expands, the ability to conceptualize both the perceived harms and benefits of genomic testing remains an important consideration, as seen in participants giving priority to understanding the "Cost Impact on Families and Individuals." These findings necessitate the need for future iterations to be agile to meet changing environments and the contexts in which genomic testing is utilized.

Two economic indicators "Cost Impact to Health Care" and "Cost Impact on Families and Individuals" were considered priorities for determining utility. Each of the indicator's descriptions included "measuring financial effects of delivering or not delivering the test" and reflects the high cost of genomic testing and implementation within a resource constrained public health system. Also, at the forefront of participants' minds when justifying how they rated the economic indicators were the implications on equity of access when the cost of testing falls on the patients and families. The third economic indicator "Cost-impacts on the Broader Economy" was removed, again possibly reflecting the sample. As emphasized by a participant and previous research,[40] failure to capture the funding needs for the wider workforce and infrastructure required when introducing a new complex intervention arguably impedes implementation. Many national strategies utilize hybrid/effectiveness implementation study designs[41] to promote the assessment of such outcomes and accelerate implementation.[31]

The notion of interrelated utility, whereby the value of one indicator is dependent on the outcome of another, was raised by participants and remains a challenge for evidence-based evaluation assessments, such as HTA. For example, the clinical benefits of a test depends on effective access to appropriate interventions[42] or the health and cost benefit of the test increases with the timing of result[12] or if cascade testing in the family is accessed as a result of testing in the proband.[17] This outcome, confirms the need for future applications of a holistic and comprehensive approach to utility.

Limitations

Although the initial response rate was not high, our participants were highly engaged and actively contributed to both rounds. Delphi studies require considerable time and effort on the behalf of participants, and we invited experts at senior leadership levels, which may reflect the overall low participation from stakeholders in state and federal health departments and relatively high drop-out from health services. Around 80% of our participants were located in the 2 most populated states in Australia, which likely reflects the overrepresentation. Further research could investigate the degree to which the indicators identified or removed in this study are priorities for stakeholders at these levels and other geographic locations. Although we did not directly involve individual consumers in the study, we did have high participation from consumer representatives and built on research that includes end-users. Although our selection of consensus threshold was based on the literature, we acknowledge the limitations of using a blunt threshold and throughout report qualitative data to assist with justification and provide a balanced argument. An example of this limitation is the outcome of the "Evidentiary Confidence"

measure, which was removed after 1 round with a borderline agreement threshold. Understanding decision-makers' tolerance for prospective utility or economic gains within levels of uncertainty is likely an important future research direction. Finally, our participants principally work within a public health care system, and we acknowledge the influence of health system structures on the assessment of utility. Further work could look at how the indicators function within other health care systems.

Evidence of utility is essential for funders of health care to make policy, resource allocation, and service planning decisions. Using Delphi methodology, we have obtained expert consensus on a broad range of utility indicators that are considered priorities in determining the value of genomic testing in rare disease. Findings may be used to further develop a standardized approach to measuring overall genomic utility and assist with a whole-of-systems approach to implementing genomics into health care.

Data Availability

The data sets generated during and/or analyzed during the current study can be made available from the corresponding author upon request.

Acknowledgments

The authors would like to thank all the experts who took part in Delphi study for their time and contribution, Tiffany Boughtwood, John Christodoulou, Lilian Downie, Kristina Elvidge, Monica Ferrie, Clara Gaff, Jodie Ingles, Kristi Jones, Edwin Kirk, Ivan Macciocca, Kristen Nowak, Nicholas Pachter, Judy Savige, Christopher Semsarian, Clare Stuart, Colman Taylor, Erin Turbitt, Ingrid Winship, and the remaining participants who have elected to remain anonymous.

Funding

This project is funded through the Australian Government's National Health and Medical Research Council (NHMRC) as part of the "Australian Genomics Grant Program Proposal 2021-23" (Grant GNT2000001).

Author Information

Conceptualization: Z.F., I.G., A.J.M., Z.S., S.B.; Methodology: Z.F., I.G., A.J.M., Z.S., S.B.; Investigation: Z.F., I.G., A.J.M., Z.S., S.B.; Formal Analysis: Z.F., I.G., A.J.M., Z.S., S.B.; Writing-original draft: Z.F., S.B.; Writing-review and editing: Z.F., I.G., A.J.M., Z.S., S.B.

Ethics Declaration

The study was approved by Royal Children's Hospital Melbourne, Human Research Ethics Committee (HREC/89720/RCHM-2022). Participation was voluntary and implied consent indicted by survey commencement after viewing participant information materials.

Conflict of Interest

The authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2024.101116) contains supplemental material, which is available to authorized users.

References

1. Foster MW, Mulvihill JJ, Sharp RR. Evaluating the utility of personal genomic information. *Genet Med*. 2009;11(8):570-574. http://doi.org/10.1097/GIM.0b013e3181a2743e
2. Grosse SD, Rasmussen SA. Exome sequencing: value is in the eye of the beholder. *Genet Med*. 2020;22(2):280-282. http://doi.org/10.1038/s41436-019-0674-z
3. ACMG Board of Directors. Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2015;17(6):505-507. http://doi.org/10.1038/gim.2015.41
4. Smith HS, Swint JM, Lalani SR, et al. Clinical application of genome and exome sequencing as a diagnostic tool for pediatric patients: a scoping review of the literature. *Genet Med*. 2019;21(1):3-16. http://doi.org/10.1038/s41436-018-0024-6
5. Stark Z, Boughtwood T, Haas M, et al. Australian Genomics: outcomes of a 5-year national program to accelerate the integration of genomics in healthcare. *Am J Hum Genet*. 2023;110(3):419-426. http://doi.org/10.1016/j.ajhg.2023.01.018
6. Stranneheim H, Lagerstedt-Robinson K, Magnusson M, et al. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med*. 2021;13(1):40. http://doi.org/10.1186/s13073-021-00855-5
7. 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR et al. 100,000 Genomes pilot on rare-disease diagnosis in health care – preliminary report. *N Engl J Med*. 2021;385(20):1868-1880. http://doi.org/10.1056/NEJMoa2035790
8. Mordaunt DA, Dalziel K, Goranitis I, Stark Z. Uptake of funded genomic testing for syndromic and non-syndromic intellectual disability in Australia. *Eur J Hum Genet*. 2023;31(9):977-979. http://doi.org/10.1038/s41431-023-01417-6
9. Goranitis I, Best S, Christodoulou J, Stark Z, Boughtwood T. The personal utility and uptake of genomic sequencing in pediatric and adult conditions: eliciting societal preferences with three discrete choice experiments. *Genet Med*. 2020;22(8):1311-1319. http://doi.org/10.1038/s41436-020-0809-2
10. Goranitis I, Best S, Stark Z, Boughtwood T, Christodoulou J. The value of genomic sequencing in complex pediatric neurological disorders: a

discrete choice experiment. *Genet Med*. 2021;23(1):155-162. http://doi.org/10.1038/s41436-020-00949-2
11. Goranitis I, Best S, Christodoulou J, Boughtwood T, Stark Z. Preferences and values for rapid genomic testing in critically ill infants and children: a discrete choice experiment. *Eur J Hum Genet*. 2021;29(11):1645-1653. http://doi.org/10.1038/s41431-021-00874-1
12. Goranitis I, Wu Y, Lunke S, et al. Is faster better? An economic evaluation of rapid and ultra-rapid genomic testing in critically ill infants and children. *Genet Med*. 2022;24(5):1037-1044. http://doi.org/10.1016/j.gim.2022.01.013
13. Wu Y, Balasubramaniam S, Rius R, Thorburn DR, Christodoulou J, Goranitis I. Genomic sequencing for the diagnosis of childhood mitochondrial disorders: a health economic evaluation. *Eur J Hum Genet*. 2022;30(5):577-586. http://doi.org/10.1038/s41431-021-00916-8
14. Wu Y, Jayasinghe K, Stark Z, et al. Genomic testing for suspected monogenic kidney disease in children and adults: a health economic evaluation. *Genet Med*. 2023;25(11):100942. http://doi.org/10.1016/j.gim.2023.100942
15. Jayasinghe K, Stark Z, Kerr PG, et al. Clinical impact of genomic testing in patients with suspected monogenic kidney disease. *Genet Med*. 2021;23(1):183-191. http://doi.org/10.1038/s41436-020-00963-4
16. O'Rourke B, Oortwijn W, Schuller T, International Joint Task Group. The new definition of health technology assessment: a milestone in international collaboration. *Int J Technol Asses Health Care*. 2020;36(3):187-190. http://doi.org/10.1017/S0266462320000215
17. Norris S, Belcher A, Howard K, Ward RL. Evaluating genetic and genomic tests for heritable conditions in Australia: lessons learnt from health technology assessments. *J Community Genet*. 2022;13(5):503-522. http://doi.org/10.1007/s12687-021-00551-2
18. Smith HS. Genomic Medicine's critical outcome measure-utility. *JAMA Netw Open*. 2022;5(8):e2225988. http://doi.org/10.1001/jamanetworkopen.2022.25988
19. Hayeems RZ, Dimmock D, Bick D, et al. Clinical utility of genomic sequencing: a measurement toolkit. *npj Genom Med*. 2020;5(1):56. http://doi.org/10.1038/s41525-020-00164-7
20. Hayeems RZ, Luca S, Ungar WJ, et al. The Clinician-reported Genetic testing Utility Index (C-GUIDE): preliminary evidence of validity and reliability. *Genet Med*. 2022;24(2):430-438. http://doi.org/10.1016/j.gim.2021.10.005
21. Hayeems RZ, Luca S, Hurst ACE, et al. Applying the Clinician-reported Genetic testing Utility Index (C-GUIDE) to genome sequencing: further evidence of validity. *Eur J Hum Genet*. 2022;30(12):1423-1431. http://doi.org/10.1038/s41431-022-01192-w
22. Turbitt E, Kohler JN, Angelo F, et al. The PrU: development and validation of a measure to assess personal utility of genomic results. *Genet Med*. 2023;25(3):100356. http://doi.org/10.1016/j.gim.2022.12.003
23. Smith HS, Brothers KB, Knight SJ, et al. Conceptualization of utility in translational clinical genomics research. *Am J Hum Genet*. 2021;108(11):2027-2036. http://doi.org/10.1016/j.ajhg.2021.08.013
24. Smith HS, Morain SR, Robinson JO, et al. Perceived utility of genomic sequencing: qualitative analysis and synthesis of a conceptual model to inform patient-centered instrument development. *Patient*. 2022;15(3):317-328. http://doi.org/10.1007/s40271-021-00558-4
25. Mallett A, Stark Z, Fehlberg Z, Best S, Goranitis I. Determining the utility of diagnostic genomics: a conceptual framework. *Hum Genomics*. 2023;17(1):75. http://doi.org/10.1186/s40246-023-00524-1
26. Beiderbeck D, Frevel N, von der Gracht HA, Schmidt SL, Schweitzer VM. Preparing, conducting, and analyzing Delphi surveys: cross-disciplinary practices, new directions, and advancements. *MethodsX*. 2021;8:101401. http://doi.org/10.1016/j.mex.2021.101401
27. Hsu CC, Sandford B. The Delphi technique: making sense of consensus. *Pract Assess Res Eval*. 2007;12:1-8.
28. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014;67(4):401-409. http://doi.org/10.1016/j.jclinepi.2013.12.002
29. Houwink EJF, Henneman L, Westerneng M, et al. Prioritization of future genetics education for general practitioners: a Delphi study. *Genet Med*. 2012;14(3):323-329. http://doi.org/10.1038/gim.2011.15
30. Borle K, Kopac N, Dragojlovic N, et al. Where is genetic medicine headed? Exploring the perspectives of Canadian genetic professionals on future trends using the Delphi method. *Eur J Hum Genet*. 2022;30(5):496-504. http://doi.org/10.1038/s41431-021-01017-2
31. Stark Z, Dolman L, Manolio TA, et al. Integrating genomics into healthcare: a global responsibility. *Am J Hum Genet*. 2019;104(1):13-20. http://doi.org/10.1016/j.ajhg.2018.11.014
32. Manolio TA, Abramowicz M, Al-Mulla F, et al. Global implementation of genomic medicine: we are not alone. *Sci Transl Med*. 2015;7(290):290ps213. http://doi.org/10.1126/scitranslmed.aab0194
33. Best S, Stark Z, Phillips P, et al. Clinical genomic testing: what matters to key stakeholders? *Eur J Hum Genet*. 2020;28(7):866-873. http://doi.org/10.1038/s41431-020-0576-1
34. Callahan KP, Mueller R, Flibotte J, Largent EA, Feudtner C. Measures of utility among studies of genomic medicine for critically ill infants: a systematic review. *JAMA Netw Open*. 2022;5(8):e2225980. http://doi.org/10.1001/jamanetworkopen.2022.25980
35. Mighton C, Carlsson L, Clausen M, et al. Quality of life drives patients' preferences for secondary findings from genomic sequencing. *Eur J Hum Genet*. 2020;28(9):1178-1186. http://doi.org/10.1038/s41431-020-0640-x
36. Dotson WD, Bowen MS, Kolor K, Khoury MJ. Clinical utility of genetic and genomic services: context matters. *Genet Med*. 2016;18(7):672-674. http://doi.org/10.1038/gim.2015.153
37. Dive L, Holmes I, Newson AJ. Is it just for a screening program to give people all the information they want? *Am J Bioeth*. 2023;23(7):34-42. http://doi.org/10.1080/15265161.2023.2207510
38. Sherburn IA, Finlay K, Best S. How does the genomic naive public perceive whole genomic testing for health purposes? A scoping review. *Eur J Hum Genet*. 2023;31(1):35-47. http://doi.org/10.1038/s41431-022-01208-5
39. Pritchard D, Goodman C, Nadauld LD. Clinical utility of genomic testing in cancer care. *JCO Precis Oncol*. 2022;6:e2100349. http://doi.org/10.1200/PO.21.00349
40. Long JC, Gul H, McPherson E, et al. A dynamic systems view of clinical genomics: a rich picture of the landscape in Australia using a complexity science lens. *BMC Med Genomics*. 2021;14(1):63. http://doi.org/10.1186/s12920-021-00910-5
41. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care*. 2012;50(3):217-226. http://doi.org/10.1097/MLR.0b013e3182408812
42. Grosse SD, Khoury MJ. What is the clinical utility of genetic testing? *Genet Med*. 2006;8(7):448-450. http://doi.org/10.1097/01.gim.0000227935.26763.c6