Health Policy Analysis

Defining a Core Data Set for the Economic Evaluation of Precision Oncology

Samantha Pollard, PhD, Deirdre Weymann, MA, Brandon Chan, MSc, Morgan Ehman, MPH, Sarah Wordsworth, PhD, James Buchanan, PhD, Timothy P. Hanna, MD, PhD, Cheryl Ho, MD, Howard J. Lim, MD, PhD, Paula K. Lorgelly, PhD, Adam J.N. Raymakers, PhD, Christopher McCabe, PhD, Dean A. Regier, PhD

ABSTRACT

Objectives: Precision oncology is generating vast amounts of multiomic data to improve human health and accelerate research. Existing clinical study designs and attendant data are unable to provide comparative evidence for economic evaluations. This lack of evidence can cause inconsistent and inappropriate reimbursement. Our study defines a core data set to facilitate economic evaluations of precision oncology.

Methods: We conducted a literature review of economic evaluations of next-generation sequencing technologies, a common application of precision oncology, published between 2005 and 2018 and indexed in PubMed (MEDLINE). Based on this review, we developed a preliminary core data set for informal expert feedback. We then used a modified-Delphi approach with individuals involved in implementation and evaluation of precision medicine, including 2 survey rounds followed by a final voting conference to refine the data set.

Results: Two authors determined that variation in published data elements was reached after abstraction of 20 economic evaluations. Expert consultation refined the data set to 83 unique data elements, and a multidisciplinary sample of 46 experts participated in the modified-Delphi process. A total of 68 elements (81%) were selected as required, spanning demographics and clinical characteristics, genomic data, cancer treatment, health and quality of life outcomes, and resource use.

Conclusions: Cost-effectiveness analyses will fail to reflect the real-world impacts of precision oncology without data to accurately characterize patient care trajectories and outcomes. Data collection in accordance with the proposed core data set will promote standardization and enable the generation of decision-grade evidence to inform reimbursement.

Keywords: core data set, economic evaluation, precision medicine, precision oncology.

Introduction

Precision oncology uses multiomic data such as genome and transcriptome analysis to tailor treatment and prevention to individual pathophysiology.[1] Fundamental to this process is next-generation sequencing (NGS), a term for massively parallel DNA sequencing including whole-genome or exome sequencing and multigene panels to identify targetable genomic aberrations and candidate pathways.[2] Despite the ability of NGS to produce rapid results at decreasing cost, clinical use of NGS varies. This is partly due to insufficient evidence demonstrating cost-effectiveness, a prerequisite to implementation guidance across jurisdictions.[3-5]

Cancer is characterized as a collection of individually rare diseases.[6] Rarity of individual genomic aberrations presents challenges for evaluations of patient and system impacts and informing timely decision making.[7] Rather than undertake lengthy patient accrual periods powered to detect small effects

and account for heterogeneity, investigators often pursue non-randomized, tumor-agnostic studies powered on short-term outcomes.[8,9] Although able to support timely reporting, evidence generated from nonrandomized designs is ill equipped to inform robust economic evaluations and corresponding reimbursement decisions.

In response to high costs associated with randomized trials, NGS evaluations are turning to real-world methods.[1,10] Recognizing both immediate and downstream impacts of precision oncology interventions, evaluations reliant on real-world data are limited in their ability to generate robust evidence. This is due to the fact that clinical studies and administrative data sets do not collect information on all relevant long-term endpoints or confounding factors.[11] Therefore, economic evaluations lack the requisite data to estimate economic value. Given that precision oncology has the potential to incur both immediate- and long-term patient and system impacts, decision makers require

evidence for each stage of the intervention and follow-up care pathway. With the emergence of genomic data generated through large-scale precision oncology trials[12,13] alongside routinely collected administrative data, guidance is urgently needed to determine what data are required to generate valid and reliable impact estimates. Although guidelines for comparative- and cost-effectiveness analyses exist,[14,15] there is a lack of specificity regarding data fields necessary to support NGS evaluations. The consistent capture of data elements collected from the point of cancer diagnosis throughout the entire patient care and follow-up trajectory will enable reliable estimates of value for money. Building upon existing frameworks and investigations, we develop a core data set to facilitate economic evaluations of precision oncology.[14,16-20]

Methods

We applied a multiphased approach to generate a set of core data elements. Investigators conducted a literature review of recent economic evaluations of precision medicine, elicited stakeholder feedback, and mapped data elements to 3 clinical data sets. Finally, we conducted a modified-Delphi process with an international sample of experts. Our approach was guided by previous core data set development procedures.[17,21-23]

Phase 1: Literature Review

The search strategy is described in Appendix Table S1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.01.005. Article selection was limited to evaluations of precision oncology and rare diseases, representing clinical contexts within which NGS has been most frequently applied to research settings.[4] We supplemented the search with a manual and key author search and informal expert consultation. Two coauthors (SP and DW) sequentially reviewed titles and abstracts followed by full review of potentially eligible articles. Final selection was based on citation frequency per year and representation across oncology and rare diseases, technologies, and authorship. This approach was designed to include a diverse sample of highly cited evaluations likely to drive future evaluations decisions and to identify variation in data sources and evaluation inputs. Three reviewers (SP, DW and ME) (SP sequentially abstracted study characteristics including evaluation type, inputs and outcome measures, methods, reported and nonreported limitations, and results. Abstracted inputs and outcomes formed the basis for the preliminary core data elements list.

The preliminary data element list was mapped to clinical data sets including the Marathon of Hope Cancer Centres Network (MOHCCN), the Minimal Common Oncology Data Elements list, and the American Association for Cancer Research's Genomics Evidence Neoplasia Information Exchange.[8,24] Finally, the list was circulated among our research team for discussion. This process continued until the team agreed the spectrum of relevant elements was included.

Phase 2: Core Data Set Refinement

Eligible participants were involved in the implementation or evaluation of precision medicine, including oncologists, clinician scientists, economists, health services researchers, and decision makers. A list of potential participants was generated by the research team based on current employment roles and institutions. Recruiting via email invitation, we applied a purposive sampling strategy to include a diversity of perspectives.[22] Using snowball sampling, we prioritized variation in expertise and location. Recruitment continued until 2 authors (SP and DAR)

agreed variation was achieved according to geography and expertise.

The modified-Delphi process included 2 online survey rounds followed by a virtual video conference.[23] Online surveys were programmed using REDCap.[25] Participants provided a written informed consent before round 1. No participants were recruited beyond round 1. Round 1 generated qualitative and quantitative feedback about the proposed list of data elements. Participants categorized elements according to whether they should be included as part of the core data set. Response options included "required"; "preferred," defined as outside the scope of a required element; or "unable to answer," for elements external to participant expertise. Participants could suggest additional data elements for consideration.[21] Consistent with previous processes,[21,23] agreement threshold was set at 70%. If 70% of participants agreed that an element should be included as either required or not required, it would be excluded from subsequent rounds. Items suggested by > 10% of participants in round 1 were considered for inclusion during round 2.[21]

Round 2 provided clarification on elements and integrated feedback from round 1. In round 2, participants categorized elements in the same manner as round 1, for which agreement was not previously reached. Based on round 1 feedback, specific elements were defined to enhance understanding regarding the proposed use as part of the core data set. Responses were summarized and reported in aggregate after round 2.[22]

The objective of the round 3 video conference was to review the core elements list, discuss elements for which agreement was not reached in previous rounds, and integrate feedback.[23] During the 1-hour conference, the facilitator (SP) introduced each element and invited discussion. Participants voted anonymously, and aggregate results were reported. Two note takers (DW and BC) documented discussion points in lieu of audio recording. As part of the round 1 survey and after the final conference, participants completed a brief demographic questionnaire. The modified-Delphi process was approved by the BC Cancer Behavioral Research Ethics Board (H20-00464).

Results

Characteristics of Included Evaluations

The MEDLINE search identified 643 articles. A total of 75 evaluations were reviewed in full. A total of 3 additional evaluations were identified after hand searching and expert consultation. Full-text review identified 52 eligible evaluations. Based on criteria described earlier, data were abstracted for 20 evaluations, after which reviewers (SP and DW) agreed that variation in inputs and outcomes had been identified, as described in Appendix Figure S2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.01.005. Detailed characteristics of included studies are provided in Appendix Table S3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.01.005.

Methodologic and Measurement Heterogeneity

The literature review found methodologic and reporting variation among highly cited economic evaluations of precision medicine, as shown in Appendix Table S1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.01.005. Evaluations ranged in terms of the number and type of comparators used, clinical trajectories modeled, stated perspective (eg, payer or societal), time horizons (eg, 1 year to lifetime), cost inputs (eg, screening, treatment, surveillance, downstream consequences for family members [spillover]), and data sources used (eg, literature,

administrative data, prospective data). Outcomes varied considerably across individual evaluations (eg, quality-adjusted life-years [QALYs], diagnostic yield, cost per diagnosis, interventions avoided, life-years gained). Evaluations captured a range of data sources including primary clinical data, published literature, and administrative health insurance claims data. Costing inputs varied from short-term limited costs of testing, screening, and treatment (eg, sequencing, validation, analysis, chemotherapy) to long-term healthcare costs (eg, consultations, surveillance, treatment, and surveillance after cascade testing). We further identified variation in reporting of economic modeling decisions as recommended by accepted reporting guidelines.[14]

The breadth of data-related challenges identified is presented in Table 1. Data sources informing model inputs and costing sources were often underreported, with heterogeneity identified in terms of cost inputs used to populate economic models.

Participant Characteristics

Recruitment for the modified-Delphi process began in April 2020 and continued until September 2020. A total of 97 invitations were emailed, with 61 potential participants providing an informed consent and 46 completing the round 1 survey. Round 2 was conducted between October and December 2020, with 35 completed surveys returned. A total of 14 voting participants attended the final consensus conference in March 2021.

One participant did not complete the round 3 demographics survey, as described in Table 2.

A Core Data Set for Precision Oncology

After the round 1 survey, 28 of 83 elements were selected as required and 5 elements as not required. The round 2 survey included 50 remaining elements for which consensus had not been reached. After round 2, 31 elements were determined to be required and no elements were determined to be not required. The round 3 conference included 13 element categories for which agreement had not been reached. After the round 3 conference, 10 unique elements were selected as required and 1 was not required. No new elements were added throughout rounds.

The 13 element categories discussed during the round 3 conference alongside final voting results, depicting persistent disagreement, are presented in Table 3.

Appendix Table S4 in Supplemental Materials found at https:// doi.org/10.1016/j.jval.2022.01.005 describes the complete list of elements deliberated upon in rounds 1 to 3. Included are elements and timing for collection. Of 83 elements considered, 68 (81%) are required, as shown in Table 4. Demographic characteristics include a patient identifier, age or date of birth, sex, and geographic location. Clinical elements span from initial cancer diagnosis, recurrence and progression, performance status, and treatment response. Genomic sequencing elements consider biopsy site and date and sequencing report encapsulating variant classification, pathogenicity, actionability, and costs of sequencing, interpretation, and presequencing and postsequencing genetic counseling. Cancer treatment information includes all previous systemic, surgical, and radiologic interventions, in addition to genomics-informed treatment. Patient outcomes include preference-based patient-reported outcome measures using validated questionnaires (eg, EQ-5D, Health Utilities Index). Additionally, data elements estimating survival, progression, metastasis, and tumor-specific secondary endpoints are required, where appropriate. Costs related to cascade hereditary cancer testing and subsequent interventions were selected as required to account for downstream spillover effects. Finally, required resource utilization

Table 1. Data challenges for conducting economic evaluations in precision oncology.

Features	Challenge
Uncertainty and biological heterogeneity	Complex clinical pathways introduce model uncertainty
	Models do not consistently or comprehensively capture spillover effects (eg, impact of test results on proband family members)
	Economic models rely predominantly on estimating health outcomes without strong evidence of comparative effectiveness, and inconsistently account for health-related quality of life and personal utility estimates
	Downstream costs are poorly characterized and inconsistent across evaluations
Methodologic transparency	Simplifying assumptions fail to reflect real-world uptake and downstream (spillover) effects, and resultant estimates risk bias
	NGS technologies that are relevant to multiple clinical conditions may fail to apply appropriate or relevant comparator(s)
	Cost-effectiveness thresholds may be selected arbitrarily or may differ across studies/laboratories

elements capture hospitalizations, physician visits, imaging, non-genomic lab-tests, and noncancer prescription drugs.

Table 5 illustrates the sequential workflow within which multiple sources of resource utilization is accumulated. Data collection able to overcome limitations identified in the published literature necessitates historic health resource use, baseline demographic and clinical characteristics, sequencing information, as well as ongoing collection of resource use and health outcomes, including patient-reported quality of life (QOL). The sequencing process alone involves detailed collection covering initial consultation, sample acquisition and preparation, sequencing, analysis, validation, interpretation, and the return of findings to providers and patients.

Our core data set enables resource utilization estimates to characterize this complex and multifaceted workflow from a healthcare payer perspective. To support routine data collection throughout enrollment, sequencing, and patient follow-up, Figure 1 describes the recommended timeline for data capture, developed by the interdisciplinary research team after data set refinement.

Discussion

Precision oncology cohort studies and clinical trials present an opportunity to collect high-quality data supporting clinical and economic evaluations to inform decision making. For example, the 100 000 Genomes Project in the United Kingdom, the All of Us

Table 2. Participant demographics.

Participant characteristic	Round 1, n = 46, % (n)	Final, n = 13, % (n)
Age band		
25-34	9 (4)	15 (2)
35-44	39 (18)	31 (4)
45-54	41 (19)	38 (5)
55-64	9 (4)	8 (1)
65+	2 (1)	8 (1)
Gender		
Male	59 (27)	77 (10)
Female	41 (19)	23 (3)
Other	—	—
Area of expertise (not mutually exclusive)		
Clinician	35 (16)	31 (4)
Health economics	52 (24)	54 (7)
Epidemiology	15 (7)	15%(2)
Health technology assessment	37 (17)	46 (6)
Policy	24 (11)	8 (1)
Precision medicine	54 (25)	51 (7)
Other	4 (2)	8 (1)
Experience (years)		
0-5	11 (5)	—
5-10	28 (13)	31 (4)
10-20	35 (16)	38 (5)
20+	26 (12)	31 (4)
Location		
Canada	57 (26)	70 (9)
United States	11 (5)	8 (1)
United Kingdom	11 (5)	15 (2)
New Zealand	2 (1)	—
France	7 (3)	—
Australia	9 (4)	—
Ireland	2 (1)	—
The Netherlands	2 (1)	8 (1)

Research Program in the United States, and Canada's MOHCCN are endeavoring to catalog patient genomes and acquire data on patients' entire cancer care trajectories.[12,13] With potential for large amounts of information generated through these types of initiatives, understanding data requirements enabling health and QOL measurement for comparative evaluations is critical. Adoption of

Table 3. Round 3 data elements for discussion.

Category	Variables	Vote (% selecting required)	Required indicator
Demographic and socioeconomic factors	Location	79	R
Genomic elements	Confirmation of histologic diagnosis	64	NR
Cancer treatment	Surgical intent	79	R
	Radiotherapy body site, dose, and fractionation	79	R
	Reason why genomics-informed treatment was not given	79	R
Patient outcomes	Disease-specific, clinically relevant secondary endpoints	71	R
	Dates and severity of adverse events	50	NR
	Number, size, location of sites, metastases	50	NR
	Categorization of involved site	50	NR
Cascade testing	Number of family members eligible for cascade testing	43	NR
	Number of family members accepting interventions	50	NR
	Health/nonhealth outcomes	29	NR
	Costs	86	R

NR indicates not required; R, required.

Table 4. Required core data set for the economic evaluation of precision oncology.

Category	Data element		Recommended timeline for data collection
Demographic and socioeconomic factors	Unique patient identifier		
	Date of birth or age		
	Sex		
	Location (eg, region, postal code, local health authority)		
Clinical characteristics	Tumor group		Baseline
	Tumor subgroup		
	Histology (eg, tumor grade)		
	Date(s) of all primary cancer diagnoses established through pathology or imaging (or both)		
	Site specific staging criteria (eg, TNM)		
	At least one performance status measure (eg, ECOG performance status)		
	Date(s) of cancer recurrence established through pathology or imaging (or both)		At diagnosis, baseline, and ongoing at regular intervals
	Date(s) of cancer metastasis established through pathology and/or imaging		
Genomic elements	Date(s) and type(s) of previous genetic testing received		Baseline (historic)
	All historic genetic test reports (including single gene and germline mutation tests)		
	Date(s) of patient's tumor biopsy		At occurrence
	Date(s) patient's normal DNA comparator was collected (eg, blood sample)		
	Flag for whether biopsy site was metastatic		
	Flag for whether biopsy site was radiated		
	Pathology tumor content from biopsy		
	Genomic tumor content from biopsy sample and sufficiency to undergo sequencing		
	Date of report		
	Full sequencing report	Aggregate sequencing information (eg, tumor mutation burden, immune signature)	
		Actionable findings (eg, OncoKB, ESMO scale, etc.)	
		Informative findings (eg, mutations that may not have prognostic or therapeutic relevance at the time of analysis but are deemed informative)	
		Relevant genes for which a germline variant was identified and corresponding pathogenicity	
		Sequencing type (eg, genome, transcriptome, exome, multi gene expression testing)	
	Date(s) that a genetic diagnosis was established		
	Cost of clinical consult		
	Cost of sample acquisition and preparation (eg, anesthesia, sample collection, pathology reagents)		
	Cost of next-generation sequencing		

Table 4. Continued

Category	Data element		Recommended timeline for data collection
	Cost of bioinformatics analysis (including computation, analyst time)		
	Cost of validation and confirmatory testing		
	Cost of interpretation by committee		
	Number of pre- and post-NGS genetic counseling appointments		At first occurrence, ongoing
	Cost of genetic counseling appointments		
Cancer treatment	Systemic therapy	Number of lines of therapy received	
		Date(s) lines were received	
		Treatment protocol(s)	
		Drug name(s)	
		Treatment intent (eg, curative or palliative)	
		Access indicator, if applicable (eg, off-label, clinical trial, out of pocket)	
	Surgical treatment	Date(s) of surgical treatment	Historic and ongoing
		Body site of surgical resection	
		Treatment intent of surgery (eg, curative or palliative)	
	Radiotherapy	Date(s) of radiotherapy treatment	
		Radiotherapy body site, dose, and fractionation	
		Modality of radiotherapy (eg, SABR, IMRT, VMAT, 3DCRT, brachytherapy)	
		Treatment intent of radiotherapy (eg, curative or palliative)	
	Indicator if treatment was provided presequencing or postsequencing		
	Indicator if treatment was genomics informed		At first occurrence, ongoing
	Reason why genomics-informed treatment was not given, if applicable		
Patient outcomes	At least one preference-based measure (eg, EQ-5D, HUI, EORTC QLQ C3015)		Baseline and ongoing at routine intervals
	Death date		At occurrence
	Disease-specific, clinically relevant secondary endpoints, as applicable		
	Date(s) of disease progression, established through (eg, Response Evaluation Criteria in Solid Tumours [RECIST] and Immunotherapy [iRECIST] criteria, clinician assessment)		At first occurrence and ongoing
	Clinician assessed best response on genomics-informed and usual care cancer treatment, (eg, stable disease, complete response, partial response, or progression, not evaluable)		
	Costs of cascade genetic testing and intervention(s)		
Resource utilization	Type and dates of hospitalizations pre- and post-NGS (admissions and discharges, including ER and ICU)		
	Costs of hospitalizations pre- and post-NGS		
	Type and dates of physician visits pre- and post-NGS (eg, general practitioner, Oncologist, other specialist)		
	Costs of physician visits pre- and post-NGS		Historic and ongoing

Table 4. Continued

Category	Data element	Recommended timeline for data collection
	Type and dates of imaging (eg, CT, MRI, PET, ultrasound, x-ray)	
	Costs of imaging pre- and post-NGS	
	Type and dates of nongenomic laboratory tests	
	Costs of nongenomic laboratory tests	
	Type and date of noncancer prescription drugs	
	Costs of noncancer prescription drugs	

CT indicates computed tomography; ECOG, Eastern Cooperative Oncology Group; ER, emergency room; ESMO, European Society for Medical Oncology; HUI, Health Utility Index; ICU, intensive care unit; iRECIST, immunotherapy Response Evaluation Criteria in Solid Tumours; MRI, magnetic resonance imaging. PET, positron emission tomography; RECIST, Response Evaluation Criteria in Solid Tumours.

the core data set will improve standardization across large-scale initiatives; accurately characterize heterogenous care patterns and outcomes; and enable robust real-world cost-effectiveness evidence generation for precision oncology.

Key Findings and Practice Implications

The ability to generate robust causal effect estimates suitable to inform reimbursement deliberations for precision oncology remains challenging.[1,10] Evaluating the spectrum of immediate and downstream impacts of precision oncology requires data access spanning diagnosis to NGS, to the integration of sequencing results into care, to subsequent resource utilization and patient outcomes. For example, variation in resource use before sequencing may correlate with downstream cost-effectiveness. Without information related to historic, baseline and downstream resource use and patient outcomes, economic evaluations will fail to characterize real-world system-level impacts.

Our literature review highlighted variation in published economic evaluations to account for selection bias stemming from these complex disease and outcomes trajectories. This variation poses challenges for decision makers when cost-effectiveness evidence is hampered by nonstandardized inputs and outcomes. Heterogeneity further limits comparability across evaluations, adding to decisional complexity. Results illustrate an unmet need to design evaluations guided by an understanding of the complex, variable downstream impacts of precision oncology, to reduce the potential for inappropriate reimbursement.[4,5,7,26] To overcome existing limitations, we established a core data set to generate evidence promoting appropriate access to high-quality precision oncology while protecting patients and health systems from technologies that do not deliver sufficient value to justify their cost.

Ensuring Comprehensive Outcomes Measurement

The primary outcome resulting from NGS is information. The extent to which that information impacts patients and health systems is a function of how patients and providers understand, value, and use genomic information to make clinical decisions. Decisions related to uptake of precision medicine are preference sensitive.[27] Clinical utility, or medical actionability, reflects availability and accessibility of interventions corresponding to identified therapeutic targets. Decisions to engage with precision oncology require patients to trade-off potential risks (eg, cost) against uncertain benefits. If individuals are unwilling to accept specific risks or trade-offs in favor of potential but unknown benefit, uptake will be limited, value for money will be lost, and reimbursement will be unsustainable.[4]

In most jurisdictions, the recommended endpoint for cost-effectiveness analysis is the incremental cost-effectiveness ratio, the incremental cost of an intervention divided by incremental effectiveness. The recommended measure of effectiveness is the QALY, combining health-related QOL (on a 0-1 scale, 0 reflecting death and 1 is perfect health) with survival.[28] Our core data set allows for collection of survival and health-related QOL outcomes after sequencing through validated instruments (eg, EQ-5D[29]). This will enable decision makers to consider health and QOL outcomes, providing patients and clinicians with information they need to make clinical decisions.[14,16]

Data, Analytic, and Implementation Considerations

Integrating prospective data collection into practice may require expansion of health system data infrastructure capacity. Although many elements may be available through administrative and claims data, routine collection of QOL may require consent and ethical approvals if conducted within research. Additional resources may be needed for prospective collection of patient outcomes regarding treatment response and progression (eg, Response Evaluation Criteria in Solid Tumours [RECIST]), where accurate measurement depends on best practices for data entry and monitoring but will reduce error through analysis of accurate and complete data sets.

Although comprehensive collection will support evaluations of varying types, the analytic method chosen will influence data needs. For example, regression-based analyses require access to individual-level costs and outcomes collected over a long-term horizon. These data may be used to populate input parameters for decision models but are not mandatory for decision model estimation. Models can instead draw on literature to overcome missing data and project costs and outcomes beyond observed values. Frameworks and study perspectives will also affect data requirements for individual studies. For example, cost-effectiveness analyses from a system or payer perspective rely solely on measured health outcomes (eg, life-years gained [QALYs]) to quantify benefits and identify relative technology efficiency. Studies applying cost-effectiveness analyses from a societal perspective may further incorporate nonhealth outcomes valued by society.

Implementation of the core data set necessitates infrastructure supporting comprehensive, routine collection across data elements. Healthcare systems are heterogenous in what data are generated routinely. Therefore, data systems will require improvements to data architecture and curation that bring together siloed data sources and improve resources enabling data scientists

Table 5. Next-generation sequencing services.

Component	Details
Clinical consult	Clinician reviews the appropriateness of sequencing for patient, provides pretest genetic counseling, and obtains informed consent.
Sample acquisition and preparation	Patients provide blood or tissue samples for clinical laboratory. Samples are prepared for sequencing.
Next-generation sequencing	Sequencing and production of raw data for the bioinformatics team
Bioinformatics analysis	Bioinformatician generates a report with genomic aberrations as well and candidate pathways. Conduct routine bioinformatics reanalysis of the sequence data (eg, variant of uncertain significance reclassification)
Validation and confirmatory testing	When necessary, validate sequencing results using additional testing (eg, IHC, FISH, or targeted sequencing)
Interpretation	Bioinformatician presents report to committee of clinicians, laboratory and research staff, senior scientists, clinical geneticists, subspecialists, molecular geneticists, and/ or referring physician. Committee reviews report and reaches consensus on recommendations and summarizes informative and actionable findings.
Results discussed with patients	Results discussed with patients and/or families with the treating physician and genetic counselor

FISH indicates fluorescence in situ hybridization; IHC, immunohistochemistry.

to use natural language processing or artificial intelligence to abstract elements unusable in the current form.[30]

Each component described earlier is critical to the implementation of the core data set and the ability to generate evidence enhancing economic evaluations of precision oncology. Stakeholder engagement and interdisciplinary collaborations are crucial to meeting this core data set. For example, clinical trialists will need to partner with health economists and other scientists to meet data requirements. Investigators designing and conducting real-world evaluations are encouraged to bring together experts in biostatistics and data science methods to consider approaches to data linkage, counterfactual cohort definition, and missing data imputation, where required.[31] As demonstrated by the current investigation, there is support to address data deficiencies and pave a path toward a more comprehensive and systematized approach to economic evaluation of precision oncology.

Efforts to broaden buy in from institutional and external stakeholders may begin with transparent communication about implementation efforts. For example, we encourage sharing approaches used to define individual data elements for the purposes of standardized collection, as well as methods to integrate prospective data collection (eg, QOL) into routine patient care. Sharing learnings around partnerships, infrastructure, and resources required for adherence to the core data set will support external institutions and jurisdictions in their efforts to increase support and build capacity.

Within health systems facing resource constraints, system-wide transformations necessary to support data set collection may be especially challenging. This limited ability for internal evaluations to capture representative patient populations or required elements may introduce inequities in evidence generation to support timely decision making. In the absence of immediate capacity to adhere to the core data set, this work presents a framework within which institutions and health systems can begin to develop requisite infrastructure.
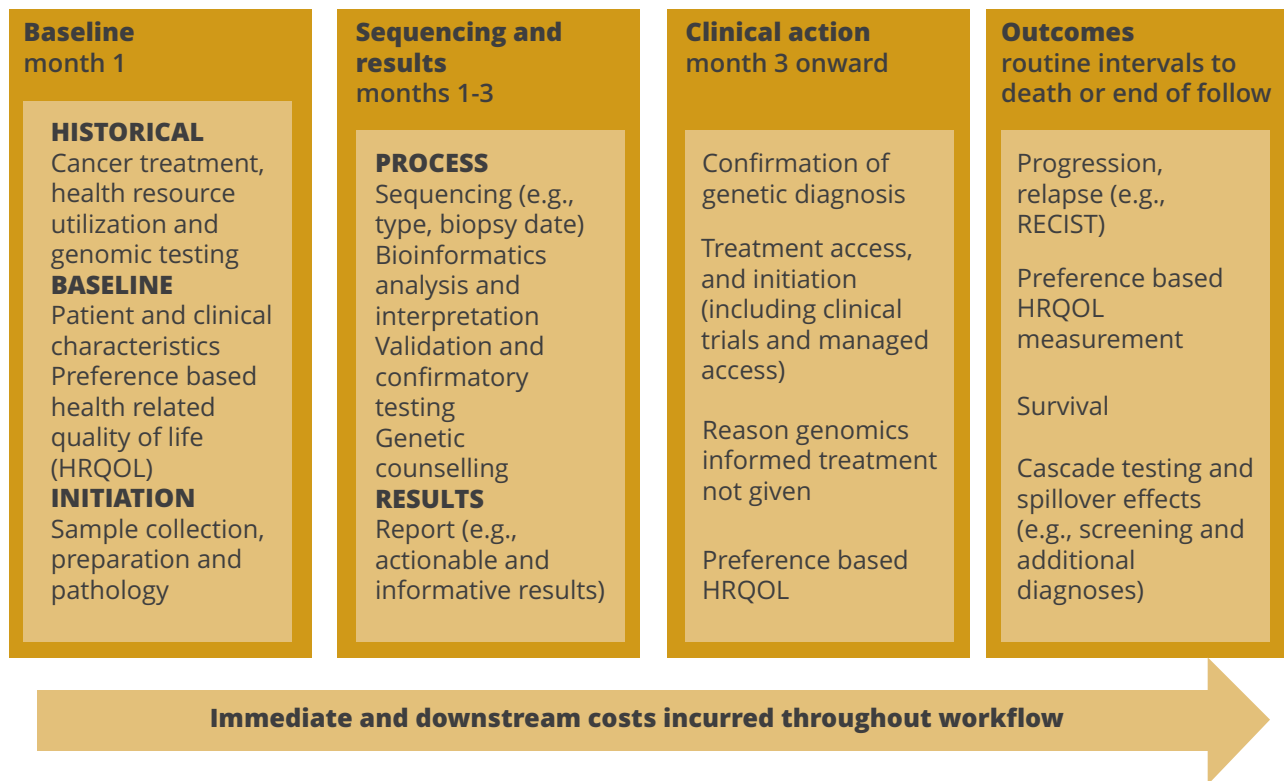
Limitations

This work has several limitations. First, the literature review obtained a selection of highly cited economic evaluations of precision oncology and rare diseases. Although the objective of the review was not to report on all published economic evaluations within these clinical contexts, we recognize that there are limitations associated with selecting only highly cited publications. Although article selection continued until 2 reviewers determined the spectrum of variation was identified, the review may not have yielded all relevant elements for consideration in the core data set. Focusing on commonly cited publications risks perpetuating a pattern of economic evaluation that fails to account for the spectrum of costs, inputs, and outcomes relevant to robust evaluation of precision oncology. Subsequent methods, including informal stakeholder feedback, the international and multidisciplinary modified-Delphi process allowing for the suggestion of new elements for consideration, and the addition of a data collection timeline, endeavored to identify and address data deficiencies and methodologic heterogeneity identified through the literature review. Despite the completion of a comprehensive review alongside stakeholder engagement, there remains a possibility that relevant data elements are not included in our core data set.

Second, we mapped preliminary iterations of the core data set to 3 external data elements lists to ensure inclusion of relevant fields. The MOHCCN data set is specific to Canada, whereas Minimal Common Oncology Data Elements and American Association for Cancer Research's Genomics Evidence Neoplasia Information Exchange are international. Further mapping to international data sets may have identified additional elements relevant for consideration. Despite this, engagement of international experts through the modified-Delphi process did not yield additional elements.

Third, using a snowball sampling approach, we endeavored to integrate a diverse and international sample of experts. Our sample was overrepresented by individuals from Canadian institutions, most of whom identified as male. Relatedly, only 30% of round 1 participants (14 of 46) contributed to all rounds of the modified-Delphi process. Furthermore, we did not collect participant-reported information related to ethnicity or cultural background and therefore are unable to comment the extent to which these perspectives are represented in the core data set. Stakeholder feedback is unlikely to reflect the spectrum of diversity because of underrepresented perspectives.

Figure 1. Precision oncology workflow and timeline for data collection.

Baseline month 1	**Sequencing and results** months 1-3	**Clinical action** month 3 onward	**Outcomes** routine intervals to death or end of follow
HISTORICAL Cancer treatment, health resource utilization and genomic testing **BASELINE** Patient and clinical characteristics Preference based health related quality of life (HRQOL) **INITIATION** Sample collection, preparation and pathology	**PROCESS** Sequencing (e.g., type, biopsy date) Bioinformatics analysis and interpretation Validation and confirmatory testing Genetic counselling **RESULTS** Report (e.g., actionable and informative results)	Confirmation of genetic diagnosis Treatment access, and initiation (including clinical trials and managed access) Reason genomics informed treatment not given Preference based HRQOL	Progression, relapse (e.g., RECIST) Preference based HRQOL measurement Survival Cascade testing and spillover effects (e.g., screening and additional diagnoses)

Immediate and downstream costs incurred throughout workflow

HRQOL indicates health-related quality of life; RECIST, Response Evaluation Criteria in Solid Tumours.

Finally, patients or patient representatives were not included in our sampling frame. Therefore, the elements of our core data set may not reflect the entire spectrum of data valued by patients. Patient-reported QOL is a required element, but patient-valued outcomes can also include, for example, indirect costs such as productivity loss, caregiver time, and transportation costs related to clinical appointments.[32] Future research generating core data elements should include patient participants. Furthermore, our investigation generated a minimum set of data elements. We emphasize that this is not a maximum data set and that prospective data collection should consider collection of additional patient-reported outcomes, patient characteristics such as gender or ethnicity, and indirect costs. Furthermore, engagement of patients, members of the public, or patient representatives as research partners within individual economic evaluations will assist in the identification of patient valued inputs and outcomes.

Conclusions

In the absence of data to characterize patient care patterns and outcomes, future economic analyses cannot comprehensively reflect real-world impacts of precision oncology. Data deficiencies and evidentiary uncertainty impose challenges for decision makers seeking to allocate scarce resources to interventions likely to maximize population benefit. Our work responds to a justified need for long-term clinical and costing data to support rigorous evaluations of precision oncology. The core data set proposed in this study will guide future database design and management for applications of NGS technologies in research and clinical settings.

Standardizing data collection will provide necessary inputs for robust clinical and economic evaluations, improve consistency across studies, and ensure decision makers have access to reliable health technology assessment evidence when making resource allocation decisions throughout the technology life cycle.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2022.01.005.

Article and Author Information

Author Affiliations: Canadian Centre for Applied Research in Cancer Control, Cancer Control Research, BC Cancer, Vancouver, Canada (Pollard, Weymann, Chan, Ehman, Raymakers, Regier); Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, England, UK (Wordsworth, Buchanan); Oxford NIHR Biomedical Research Centre, Oxford, England, UK (Wordsworth, Buchanan); Department of Oncology, Queen's University, Kingston, Canada (Hanna); Division of Medical Oncology, BC Cancer, Vancouver, Canada (Ho, Lim); Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, Canada (Ho, Lim); Department of Applied Health Research, University College London, London, England, UK (Lorgelly); Institute of Health Economics, Edmonton, Alberta, Canada (McCabe); School of Population and Public Health, University of British Columbia, Vancouver, Canada (Regier).

Correspondence: Dean A. Regier, PhD, Canadian Centre for Applied Research in Cancer Control, Cancer Control Research, BC Cancer, 601 West Broadway, Vancouver, BC V5Z 4C2, Canada. Email: dregier@bccrc.ca

REFERENCES

1. Prasad V, Fojo T, Brada M. Precision oncology: origins, optimism, and potential. *Lancet Oncol.* 2016;17(2):e81–e86.
2. Phillips KA, Deverka PA, Hooker GW, Douglas MP. Genetic test availability and spending: where are we now? Where are we going? *Health Aff (Millwood).* 2018;37(5):710–716.
3. Weymann D, Dragojlovic N, Pollard S, Regier DAJ. Allocating healthcare resources to genomic testing in Canada: latest evidence and current challenges [published online July 5, 2019]. *J Community Genet.* https://doi.org/10.1007/s12687-019-00428-5.
4. Regier DA, Weymann D, Buchanan J, Marshall DA, Wordsworth S. Valuation of health and nonhealth outcomes from next-generation sequencing: approaches, challenges, and solutions [published correction appears in Value Health. 2019;22(4):502]. *Value Health.* 2018;21(9):1043–1047.
5. Rogowski W, Payne K, Schnell-Inderst P, et al. Concepts of 'personalization' in personalized medicine: implications for economic evaluation. *Pharmacoeconomics.* 2015;33(1):49–59.
6. Renfro LA, An MW, Mandrekar SJ. Precision oncology: a new era of cancer clinical trials. *Cancer Lett.* 2017;387:121–126.
7. Regier DA, Veenstra DL, Basu A, Carlson JJ. Demand for precision medicine: a discrete-choice experiment and external validation study. *Pharmacoeconomics.* 2020;38(1):57–68.
8. Laskin J, Jones S, Aparicio S, et al. Lessons learned from the application of whole-genome analysis to the treatment of patients with advanced cancers. *Mol Case Stud.* 2015;1(1):a000570.
9. Park JJ, Hsu G, Siden EG, Thorlund K, Mills EJ. An overview of precision oncology basket and umbrella trials for clinicians. *Cancer J Clin.* 2020;70(2):125–137.
10. Le Tourneau C, Delord J-P, Gonçalves A, et al. Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. *Lancet Oncol.* 2015;16(13):1324–1334.
11. Weymann D, Pataky R, Regier DA. Economic evaluations of next-generation precision oncology: a critical review. *JCO Precis Oncol.* 2018;2:1–23.
12. Turnbull C, Scott RH, Thomas E, et al. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS [published correction appears in *BMJ.* 2018;361:k1952]. *BMJ.* 2018;361:k1687.
13. All of Us Research Program Investigators. The "All of Us" research program. *N Engl J Med.* 2019;381(7):668–676.
14. Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Eff Resour Alloc.* 2013;11(1):6.
15. Carias C, Chesson HW, Grosse SD, et al. Recommendations of the second panel on cost effectiveness in health and medicine: a reference, not a rule book. *Am J Prev Med.* 2018;54(4):600–602.
16. Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA task force report. *Value Health.* 2005;8(5):521–533.
17. Stone E, Rankin N, Phillips J, et al. Consensus minimum data set for lung cancer multidisciplinary teams: results of a Delphi process. *Respirology.* 2018;23(10):927–934.
18. Abbasi M, Ahmadian L, Amirian M, Tabesh H, Eslami S. The development of a minimum data set for an infertility registry. *Perspect Health Inf Manag.* 2018;15:1b.
19. Casarett DJ, Teno J, Higginson I. How should nations measure the quality of end-of-life care for older adults? Recommendations for an international minimum data set. *J Am Geriatr Soc.* 2006;54(11):1765–1771.
20. Choquet R, Maaroufi M, de Carrara A, Messiaen C, Luigi E, Landais P. A methodology for a minimum data set for rare diseases to support national centers of excellence for healthcare and research. *J Am Med Inform Assoc.* 2015;22(1):76–85.
21. Sigurdardottir KR, Kaasa S, Rosland JH, et al. The European Association for Palliative Care basic dataset to describe a palliative care cancer population: results from an international Delphi process. *Palliat Med.* 2014;28(6):463–473.
22. Howell D, Fitch M, Bakker D, et al. Core domains for a person-focused outcome measurement system in cancer (PROMS-Cancer Core) for routine care: a scoping review and Canadian Delphi Consensus. *Value Health.* 2013;16(1):76–87.
23. Hirsch M, Duffy JM, Barker C, et al. Protocol for developing, disseminating and implementing a core outcome set for endometriosis. *BMJ Open.* 2016;6(12):e013998.
24. mCODE™: minimal common oncology data elements. American Society of Clinical Oncology. https://mcodeinitiative.org. Accessed 2020.
25. REDCap: Research Electronic Data Capture. REDCap. https://www.project-redcap.org. Accessed March 1, 2020.
26. Buchanan J, Wordsworth S, Schuh A. Issues surrounding the health economic evaluation of genomic technologies. *Pharmacogenomics.* 2013;14(15):1833–1847.
27. Regier DA, Peacock SJ, Pataky R, et al. Societal preferences for the return of incidental findings from clinical genomic sequencing: a discrete-choice experiment. *CMAJ.* 2015;187(6):E190–E197.
28. The Canadian Agency for Drugs and Technologies in Health. *Guidelines for the Economic Evaluation of Health Technologies: Canada.* 4th ed. Ottawa, Canada: The Canadian Agency for Drugs and Technologies in Health (CADTH); 2017.
29. EuroQol Research Foundation. EQ-5D. https://euroqol.org/eq-5d-instruments/. Accessed 2021.
30. Fahr P, Buchanan J, Wordsworth S. A review of the challenges of using biomedical big data for economic evaluations of precision medicine. *Appl Health Econ Health Policy.* 2019;17(4):443–452.
31. Weymann D, Laskin J, Jones SJM, et al. Matching methods in precision oncology: an introduction and illustrative example. *Mol Genet Genom Med.* 2021;9(1):e1554.
32. Iragorri N, de Oliveira C, Fitzgerald N, Essue B. The indirect cost burden of cancer care in Canada: a systematic literature review. *Appl Health Econ Health Policy.* 2021;19(3):325–341.