Genetics in Medicine (2022) 24, 1–14

ELSEVIER

Genetics in Medicine

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

REVIEW

Systematic review of outcomes in studies of reproductive genetic carrier screening: Towards development of a core outcome set

Ebony Richardson*, Alison McEwen, Toby Newton-John, Ashley Crook, Chris Jacobs

University of Technology Sydney, Graduate School of Health, Broadway, NSW, Australia

ARTICLE INFO

Article history:
Received 3 March 2021
Revised 9 June 2021
Accepted 10 September 2021
Available online 30 November 2021

Keywords:
Reproductive genetic carrier screening
Outcomes
Core outcome set

ABSTRACT

Purpose: Current practice recommendations support the widespread implementation of reproductive genetic carrier screening (RGCS). These consensus-based recommendations highlight a research gap, with findings from current studies being insufficient to meet the standard required for more rigorous evidence-based recommendations. This systematic review assessed methodological aspects of studies on RGCS to inform the need for a core outcome set.

Methods: We conducted a systematic search to identify peer-reviewed published studies offering population-based RGCS. Study designs, outcomes, and measurement methods were extracted. A narrative synthesis was conducting using an existing outcome taxonomy and criteria used in the evaluation of genetic screening programs as frameworks.

Results: Sixty-five publications were included. We extracted 120 outcomes representing 24 outcome domains. Heterogeneity in outcome selection, measurement methods and time points of assessment was extensive. Quality appraisal raised concerns for bias. We found that reported outcomes had limited applicability to criteria used to evaluate genetic screening programs.

Conclusion: Despite a large body of literature, diverse approaches to research have limited the conclusions that can be cumulatively drawn from this body of evidence. Consensus regarding meaningful outcomes for evaluation of RGCS would be a valuable first step in working towards evidence-based practice recommendations, supporting the development of a core outcome set.

Introduction

Population-based reproductive genetic carrier screening (RGCS) identifies individuals and couples with an increased risk of having a child affected by a recessive or X-linked condition.[1] Practice recommendations support the widespread offer of RGCS to the general population, endorsing the discussion of RGCS with all women planning a pregnancy or during their first trimester and promoting informed choice to accept or decline the offer.[2-4] Such practice recommendations are guided by evidence from published research and expert consultation, with a grading

*Correspondence and requests for materials should be addressed to Ebony Richardson, Graduate School of Health, University of Technology Sydney, Building 20, 100 Broadway, Chippendale, NSW 2008, Sydney, Australia. *E-mail address:* ebony.j.richardson@student.uts.edu.au

doi: https://doi.org/10.1016/j.gim.2021.08.005

system used to indicate the type and quality of evidence available to support each recommendation.[5,6] In the case of RGCS, practice recommendations have utilised expert consensus as the primary source of evidence, indicating that the published literature has been insufficient to inform more rigorous evidence-based recommendations. Considering evidence-based practice as a key goal of medicine, there is a need for integration of clinical expertise with the best available evidence from systematic research.[7] As consensus-based recommendations drive the widespread implementation of RGCS in the general population it is crucial to assess the impacts, benefits and potential harms using rigorous methods. It is timely to consider the issues that exist in the current body of evidence and how these can be addressed to ensure that future studies can reliably inform evidence-based implementation of RGCS.

Previous systematic reviews have examined RGCS programs for specific conditions or focused on areas of particular interest, including reasons for uptake, informed choice, and reproductive decisions.[8-13] These systematic reviews all mentioned difficulty in synthesising data due to heterogeneity in study design, selection of outcomes, or measurement methods. Of particular note, two Cochrane systematic reviews identified no eligible studies due to stringent inclusion criteria requiring randomised or quasi-randomised study design, which were absent at the time of these reviews.[14,15] We hypothesised that diverse approaches to research as noted in previous reviews may account for the reliance on consensus-based recommendations for RGCS. If this is the case, the development of a core outcome set (COS) may be appropriate.

A COS is a minimum set of outcomes that should be measured and reported in all studies on a particular topic.[16] The development of a COS involves a multi-step consensus process incorporating key stakeholder groups. This systematic review is the first of these steps, and will be followed by a systematic review of qualitative literature, focus groups and interviews with patients, and a consensus process consisting of a Delphi survey and consensus meeting, details of which are outlined in full in the Core Outcome Development for Carrier Screening (CODECS) Study protocol.[17] When implemented into research, a COS ensures that a small number of outcomes are consistently available for comparison across studies, minimises outcome-reporting bias by ensuring that core outcomes are always reported regardless of significance, and maximises the relevance of outcomes due to the input of key stakeholders, including patients. This systematic review aims to assess the methodology used in studies that have implemented RGCS to inform the need for a core outcome set.

Material and Methods

This systematic review was registered with PROSPERO (CRD42019140793) and conducted per the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement and guidance from the Core Outcome Measurement in Effectiveness Trials (COMET) Initiative.[16,18] We searched the Cochrane Database of Systematic Reviews, the Joanna Briggs Institute Systematic Reviews database, MEDLINE, and PROSPERO to ensure that no similar systematic reviews were underway.

Search strategy

MEDLINE, EMBASE, CINAHL, and PsycINFO were searched on 1 December 2020 (illustrative search available in Supplemental File 1). We performed forward and backward searching using reference lists of included publications and forward citation through Google Scholar.

Study selection

All peer-reviewed published studies available in English that offered RGCS for recessive or X-linked conditions to participants through a population screening program were eligible for inclusion. Title and abstract screening, then full-text screening was performed in 10% increments by two independent reviewers (ER and AC) until >85% interrater reliability was achieved, with the remainder reviewed by the primary reviewer (ER) only. Any disagreements were resolved through discussion, and where required, by input from a third reviewer (CJ).

Quality assessment and risk of bias

The primary reviewer (ER) scored the quality of the included studies using the QualSyst tool.[19] "Quality" was defined in terms of the studies' internal validity or the extent to which the design, conduct, and analyses minimised errors and biases. As our aim was to determine all previously published outcomes regardless of study quality, the assessment of bias was not used as grounds for exclusion but rather to give an overall evaluation of study quality within the literature.

We assessed outcome reporting bias according to the Outcome Reporting Bias In Trials (ORBIT) study classification system for missing or incomplete outcome reporting.[20,21] Where available, we obtained published protocols for included studies and compared the outcomes to those reported in subsequent publications. We used discrepancies in outcomes between the protocol and subsequent publications to define a low or high risk of outcome reporting bias.

Data extraction

Due to the large number of studies identified through our search, data extraction was conducted in 5-year increments until outcome saturation was reached. This methodology is suitable for situations where the size of the review would be unmanageable if conducted in full.[16] Outcome saturation was defined as the point at which no new unique outcomes

were identified, and this occurred within two 5-year cycles (2010-2015, 2016-2020). This approach ensures that data extraction will continue until all relevant outcomes have been identified and prevents missing relevant outcomes from earlier research. For the purpose of this review, only quantitative data was extracted and analysed.

We extracted all outcomes that have been reported in studies of RGCS. Outcomes, and where supplied, their definition, measurement methods and time point were extracted verbatim using NVivo software.[22] The primary outcome was noted when specified, and basic study characteristics extracted. A coding guide was developed and piloted by ER and AC for 20% of studies to ensure consistency in data extraction for the remainder extracted by ER only. We defined study types within overarching categories of observational or experimental design, with further granularity defined by descriptive or analytic (inferential) statistics, single (cross-sectional) or multiple timepoints (cohort), and prospective or retrospective nature.[23]

Data analysis

We performed two approaches to narrative synthesis.[24] Firstly, a narrative synthesis was conducted to categorise study designs, outcomes and measurement methods. The COMET taxonomy was used as a high-level framework.[25] We elected not to define outcomes as adverse events/effects as there is currently no consensus definition for adverse outcomes in the context of genetic testing. Outcomes were grouped into more granular domains by ER, hereafter referred to as CODECS domains, and mapped to the COMET taxonomy. Definition of the domain and grouping of outcomes were developed iteratively with AC and taken to the study management group (CJ, AM, TNJ) for final review and consensus. Twenty-four CODECS study domains were defined (Supplemental Material 2). The number of outcomes with similar definitions or themes within each CODECS domain was used to indicate outcome heterogeneity. We analysed the frequency of outcome reporting at the level of individual outcomes, CODECS domains and COMET taxonomy domains. Measurement methods within each CODECS domain were captured and assessed for validation and piloting as an indication of quality. Meta-analysis was not appropriate for the goals of this review.

Secondly, a narrative synthesis was conducted using criteria defined in a review of RGCS for cystic fibrosis.[26] These criteria were used as a framework to determine whether outcomes reported in eligible studies would be applicable to criteria commonly used to evaluate genetic screening programs and inform evidence-based practice recommendations. Four criteria were defined; participation is voluntary with time allowed for consideration and based on consent, the target group is provided with good quality, comprehensible, and balanced information, there is enough evidence that psychological harm caused by the offer and/or participation is negligible, and there is enough evidence that social harm caused by the offer and/or participation is negligible. This approach was chosen as there are currently no consensus criteria for the assessment of genetic screening programs, and existing criteria used in other screening contexts have been recognised to have limited applicability for genetic screening programs.[26-28]

Results

Search strategy

Our literature search identified 2,923 records. After de-duplication and title and abstract screening, 430 publications remained. The remaining publications were separated into 5-year periods, and 230 full-texts published between 2010-2020 were screened. Sixty-five publications from 48 related studies were eligible for inclusion (Figure 1).[29-93]

Study characteristics

Study characteristics are summarised in Table 1. Eligible studies were from 15 countries, with the highest output from the USA (n = 14, 29%), Australia (n = 6, 13%) and Italy (n = 6, 13%). The most frequently reported RGCS programs were for haemoglobinopathies (n = 14, 31%), targeted panels in founder populations (n = 11, 21%), and expanded carrier screening panels (n = 11, 21%).

Quality assessment and risk of bias

Quality scores correlated to more rigorous study designs, with randomised controlled trials (RCTs) scoring highest (mean = 0.96, range = 0.92-1.0), followed by analytic studies (mean = 0.87, range = 0.61-1.0), and descriptive studies (mean = 0.79, range = 0.43-1.0). These results reflect the expected increase in potential bias that is introduced by less rigorous study designs. Scoring per study is available in Supplemental Material 3.

Outcome reporting bias could not be assessed for most studies (n = 45, 94%). Three protocols were available: two RCTs[44,94] and an analytic cross-sectional study.[95] The first protocol demonstrated consistency in the measurement, analysis and reporting of all outcomes that were defined in their published protocol.[37,44,45] No missing data were identified, and therefore the ORBIT classification was not applied. The second protocol defined ten outcomes, nine of which were published.[72,95] ORBIT Classification F was applied, indicating a low risk of outcome reporting bias for this study. The third protocol defined 16 outcomes; three of these were represented in publications included in this review, one was reported for a subset of patients only, and one was reported in a publication not included in this review but known to the authors.[41,50,62,94] Six published outcomes did not correspond to a defined protocol outcome. Eleven

outcomes defined in the protocol were not identified in publications to date and constitute missing data from this review. Due to inconsistencies between the protocol and publications, ORBIT classification E was applied, indicating a high risk of outcome reporting bias for this study.

Study designs

Most studies were observational in design (n = 46, 96%), with only two RCTs identified. A protocol for a third RCT was identified; however, no publications were available. Most studies provided descriptive statistics (n = 35, 73%), collected cross-sectional data (n = 42, 88%), and were retrospective in nature (n = 30, 63%). The most common study type was descriptive cross-sectional studies, representing audit-style summaries of a screening offer (n = 33, 69%). A detailed summary of included studies can be found in Supplemental Material 4.

Frequency of study outcomes

One hundred and twenty outcomes were extracted. The average number of outcomes reported per publication was seven (range 1-23). Only 8% (n = 5) of publications defined the primary outcome(s). The most frequently reported outcomes across studies were detection rate based on either DNA analysis or biochemical assays (n = 39, 81%), identification of increased risk couples (n = 26, 54%), uptake of prenatal diagnosis (n = 22, 46%), and results of prenatal diagnosis (n = 20, 42%). Outside of clinical outcomes directly related to test results or pregnancy outcomes, the most frequently reported outcomes were uptake of RGCS (n = 17, 35%), knowledge pertaining to the test offer (n = 10, 21%), and anxiety (n = 8, 17%).

Outcome domains and heterogeneity

Outcomes were grouped into 24 CODECS domains, with a range of 1-11 outcomes per domain, with higher numbers being indicative of outcome heterogeneity. CODECS domains were mapped to the COMET taxonomy (Figure 2), with the highest proportion of outcomes in the domain of 'delivery of care' (n = 48, 40%).

The frequency of reporting per CODECS domain is shown in Figure 3. The most frequently reported CODECS outcome domains were 'primary outcomes of RGCS'

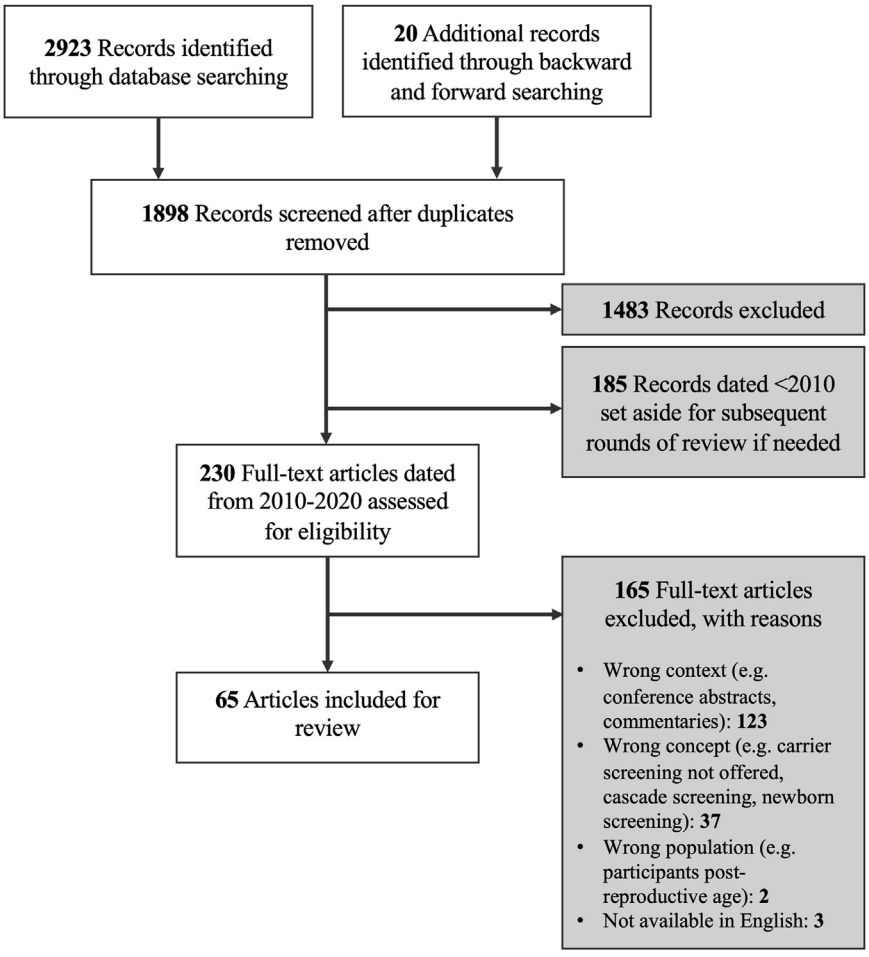

Figure 1 PRISMA Diagram.[18]

Table 1 Summary of included studies

Study Design ($N = 50$)[a,b]	Number of Studies
Observational ($n = 48$)	
Analytic Cohort, Prospective	1
Analytic Cohort, Retrospective	3
Analytic Cross-Sectional, Retrospective	9
Descriptive Cohort, Prospective	2
Descriptive Cross-Sectional, Prospective	14
Descriptive Cross-Sectional, Retrospective	19
Experimental ($n = 2$)	
Randomised Controlled Trial	2
Year of Publication ($N = 65$)	
2020-2016	33
2010-2015	32
Country of Study ($N = 48$)[a]	
Australia	6
China	2
Greece	2
India	2
Israel	3
Italy	6
Korea	1
Lebanon	1
Taiwan	2
Thailand	2
The Netherlands	3
Turkey	2
UAE	1
UK	1
USA	14
Population[b]	
Individuals undertaking RGCS during prenatal or preconception period	
Prenatal only	11
Preconception only (includes compulsory pre-marital screening)	6
Either	31
Increased risk couples identified through RGCS	2
Intervention[b]	
Haemoglobinopathies	16
Targeted panel in founder population	11
Expanded carrier screening (ECS)	11
Cystic fibrosis (CF)	4
Fragile X (FXS)	3
Spinal Muscular Atrophy (SMA)	4
3-gene panel (CF, FXS, SMA)	3

[a]65 publications from 48 studies.

[b]Some related studies included multiple study designs, populations, or interventions depending on sub-analyses published independently.

($n = 39$, 81%), 'intention and uptake' ($n = 34$, 71%), 'need for further testing' ($n = 29$, 60%), and 'pregnancy outcomes' ($n = 21$, 44%). Of the domains that included patient-reported outcomes, 'knowledge' was the most frequently reported. Most outcome domains demonstrated outcome heterogeneity. Two were most notable due to the degree of heterogeneity and the fact that they were frequently reported in studies on RGCS; psychological wellbeing and timeliness.

The outcome domain of 'psychological wellbeing' was reported in 20% of studies.[43,44,54,56,62,71,72,76,89,92] Ten different outcomes were used to measure psychological wellbeing; anxiety, concern, depression, feelings about results, perceived ability to cope, predicted negative feelings, reassurance, stress, subjective distress, and worry. Of these, the most frequently reported was anxiety. Most studies measured more than one psychological outcome (range: 1-3). Use of validated measures and timepoints of assessment were highly variable.

The outcome domain of 'timeliness' was reported in 20% of studies.[33,44,45,49,60,68,71,73,86,93] We defined timeliness as the provision of RGCS and follow-up testing, typically in the prenatal setting, in a manner that allowed sufficient time for deliberation and decision-making. Eleven different outcomes were reported pertaining to timeliness; gestational age when offered RGCS, gestational age at uptake, the time between pregnancy confirmation and RGCS, turnaround time for results, time between maternal results and partner testing, gestational age at the time of partner results, gestational age at the time of prenatal diagnosis, proportion screened by 10-, 12-, 16- and 26-weeks' gestation. There was a lack of consistency in defining gestations by which services were considered to have been delivered in a timely manner. Reporting was variable, with mean, median and range being used interchangeably.

Measurement methods

Various measurement methods were extracted from eligible studies, with most outcomes measured using an investigator-derived scale ($n = 66$, 55%) or extracted from clinical or laboratory databases ($n = 52$, 43%). Only a minority of outcomes ($n = 14$, 12%) were measured using a previously reported or validated patient-reported outcome measure, all of which were in the domains of psychological wellbeing, knowledge, decision satisfaction/regret and deliberation/informed choice (Supplemental Material 5).

Twelve publications from 10 studies assessed knowledge, each using a different measurement method. One used a validated knowledge scale that was designed specifically for their study,[72] four adapted a previously published scale that had been validated for use in a different context,[37,44,54,55,80] three adapted a previously published scale that had not undergone formal validation,[53,71,89] one developed a new scale and piloted it before use,[39] and one provided insufficient information regarding the measurement method.[48] Where a previously published scale was adapted, the integrity of the validation or piloting of the original scale was often compromised by the addition or removal of questions, changes in wording, or merging of

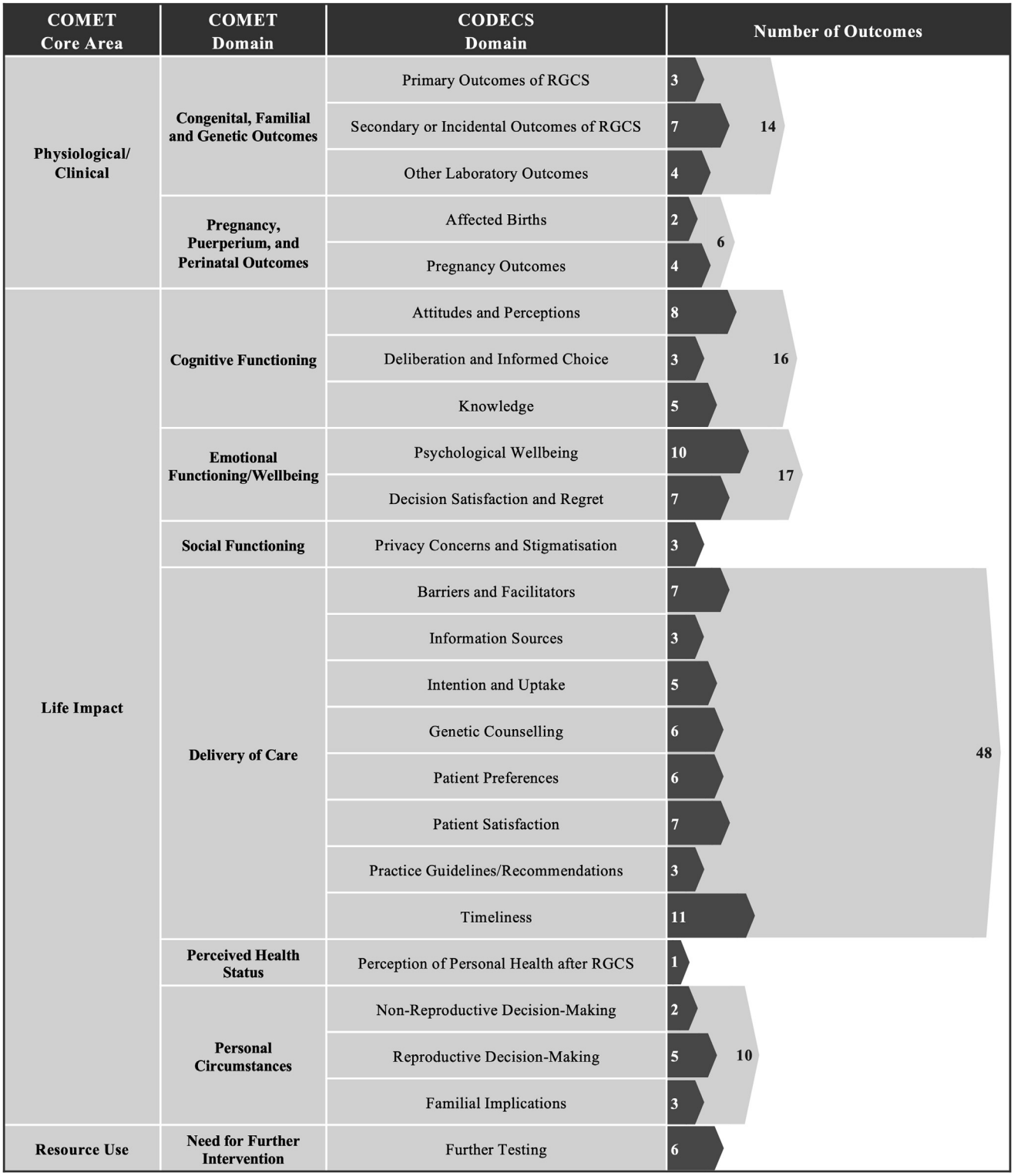

Figure 2 **Summary of outcomes per domain, mapped to applicable core area and domains from the COMET taxonomy.**

multiple previous scales into a new scale. Only one study performed formal validation of the adapted scale,[37,44] one study piloted the adapted scale,[71] and five studies did not report any piloting or validation of the adapted scale.[53-56,80,89]

Time points of measurement were also variable. A single time point was assessed by most studies (n = 41, 85%) and included audit data from databases between 1-30 years since screening (n = 34, 83%),[29,30,32-36,38,40,42,47-49,52,57,60,61,63-65,67-69,73,74,78,81,83-88,90,91,93] patient-reported outcomes at

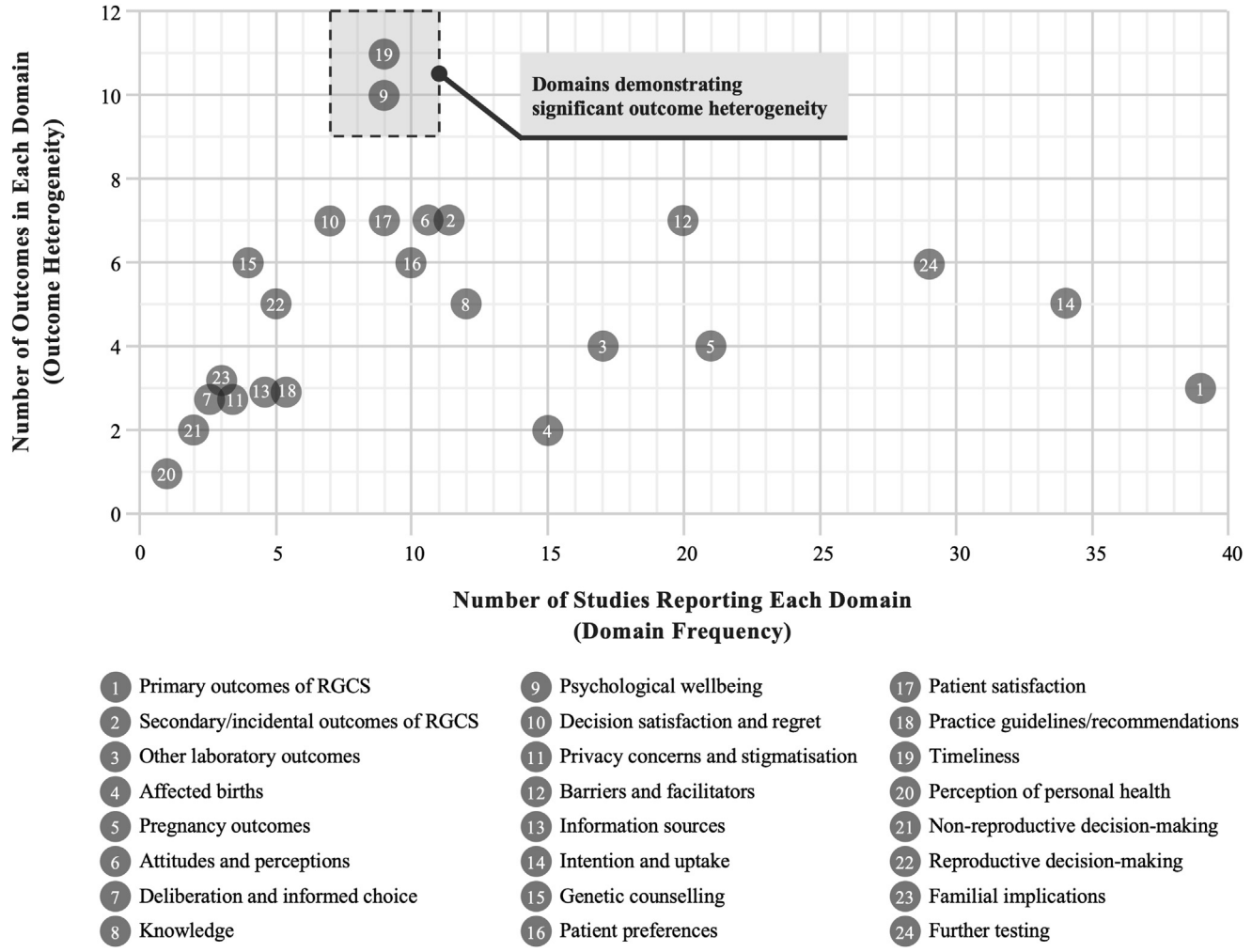

Figure 3 Domain Frequency and Outcome Heterogeneity. Visualising outcome heterogeneity in conjunction with frequency of domain reporting highlights domains that are most problematic when considering consistency and comparability in the research literature.

pre-test counselling after the decision to accept or decline was made (n = 3, 7%),[75,76,82] after maternal results but before partner results (n = 1, 2%),[39] and after results between 1-2 years since screening (n = 3, 7%).[46,51,53,55,56,58,59,66,92] Seven studies (15%) measured outcomes at multiple time points including before attending pre-test counselling (n = 2, 4%),[79,80] before and after education (n = 1, 2%),[89] at pre-test counselling when decision to accept or decline was made (n = 5, 12%),[37,41,43-45,50,54,62,70-72] and after results ranging from 2 weeks to >10 years after screening (n = 6, 15%).[37,41,43-45,50,54,62,70-72,77,79,80]

Criteria for assessment of genetic screening programs[26]

Participation is voluntary with time allowed for consideration and based on consent

CODECS outcome domains of 'intention and uptake', 'attitudes and perception', 'decision satisfaction and regret', and 'deliberation and informed choice', were mapped to the above criterion. Intended or actual uptake was reported in 71% (n = 34) of studies.[30,32-38,40-42,44,48-53,55-57,60,61,63,65,69-72,74,75,77-80,84-87,89,93] The outcome domain of 'attitudes and perceptions', which includes outcomes that assess how attitudes or perceptions influence test uptake, was reported in 23% (n = 11) of studies.[37,44,54,55,62,71,72,76,80,82,89,90,92] Outcome domains of 'decisional satisfaction and regret' and 'deliberation and informed choice' were reported by 15% (n = 7)[43,44,54,55,62,71,75,92] and 6% (n = 3)[37,44,72,79,80] of studies respectively.

The target group is provided with good quality, comprehensible, and balanced information

CODECS outcome domains of 'knowledge' and 'patient satisfaction' were mapped to this criterion. Knowledge was the most frequently reported of these outcome domains, with 25% (n = 12)[37,39,42-44,48,53-56,71,72,75,80,89] of studies assessing outcomes such as knowledge of the screening offer, recall, and understanding. Patient satisfaction was reported in 15%

Summary of Findings:

- 24 outcome domains were identified that capture the research landscape evaluating RGCS to date
- Selection of outcomes and measurement methods were heterogenous
- Indications of bias in the current literature were identified
- Cross-sectional and audit-style papers were over-represented
- Descriptive statistics were prevalent, limiting causal inferences or associations
- Patient perspective was lacking, with few studies including patient reported outcomes and no reported involvement of patients in study design
- Studies rarely defined primary outcome(s), highlighting a lack of consensus for how to assess clinical utility of RGCS
- Outcomes had limited applicability to criteria used to evaluate genetic screening programs
- Potential harms of RGCS, including psychological or social harms, cannot be robustly concluded from the current literature.

How Can A Core Outcome Set Help?

- Heterogeneity of outcomes and measurement methods will be minimised by defining a small number of outcomes that are recommended for all future studies.
- Involvement of key stakeholders in the identification and definition of core outcomes maximises relevance of research findings, with patients and policy-makers principally of concern from this review.
- Defining primary outcome(s) of RGCS will provide clarity in future studies aiming to evaluate RGCS offers.

Box 1 Overview of systemic review findings.

($n = 7$) of studies, assessing outcomes such as helpfulness of educational materials, feeling information needs were met, and satisfaction with pre-test genetic counselling.

There is enough evidence that psychological harm caused by the offer and/or participation is negligible

CODECS outcome domains of 'psychological wellbeing' and 'perception of personal health status after RGCS' were mapped to this criterion. Psychological wellbeing was reported in 20% ($n = 9$) of studies,[43,44,54,56,62,71,72,76,89,92] and perceptions of personal health status was reported in a single study.[71]

There is enough evidence that social harm caused by the offer and/or participation is negligible

CODECS outcome domains of 'affected births', 'reproductive decision-making', 'non-reproductive decision-making', 'familial implications', and 'privacy concerns and stigmatisation' were mapped to this criterion. The outcome domain of 'affected births' was reported in 31% ($n = 15$) of studies.[29,30,32,35,38,40,48,58,60,64,67,69,73,84,86] The outcome domain of 'reproductive decision-making' was reported in 10% ($n = 5$) studies.[46,58,59,62,70,71,76,92] Assessment of social impact or harms outside of reproductive decisions and birth rates was limited to a handful of studies and included assessing the impact of results on the couple's relationship, dissemination of results to family members, concerns regarding discrimination by insurance companies, and fear of stigmatisation within the community.[52,54,62,71,89] See Box 1 for a quick summary of the findings.

Discussion

A lack of consensus for 'what to measure' in research evaluating health interventions is a major challenge across the medical field and has been recognised to limit reliability of the conclusions that can be drawn from research evidence.[96] Significant inconsistency in the choice of outcomes, measurement methods, and a lack of outcomes informed by patients as end-users have been noted across medical specialities, including in reviews on clinical genetic service outcomes.[97,98] Increasingly, discussions within the genetics community focus on how we can best define healthcare outcomes to capture the value of genetic services, genetic counselling, and genetic testing.[99,100] This review is the first step in a structured approach to addressing this question.

Across studies of population-based RGCS, we identified potential biases introduced by study design, heterogeneity in outcome selection, and variability in measurement methods. We found that outcomes had limited applicability to criteria used to evaluate genetic screening programs. While consensus-based practice recommendations have led to increasing support for the widespread offer of RGCS to the general population, in order to achieve the goal of evidence-based medicine, it is imperative to address the issues highlighted in this review to generate research that can inform evidence-based practice recommendations as we move forward.[7]

We found that study designs compromised the quality of evidence from the current literature. Firstly, there were a large number of observational studies that have the potential

to introduce biases at the stage of design and conduct, with selection and measurement biases principle amongst these. We also found a high risk of outcome reporting bias in one of only two randomised controlled trials on this topic. Secondly, previous reviews have recognised the prevalence of cross-sectional studies as a methodological limitation of research on genetic testing and counselling.[100] Our findings are consistent with these reviews, with a predominance of cross-sectional studies and limited follow-up of outcomes over time, reiterating the necessity for longitudinal approaches to future research. Thirdly, a previous review has highlighted that a lack of analytic statistics impeded efforts to infer factors influencing decision-making and their relative contributions.[10] Our findings similarly revealed favouring of descriptive, as opposed to analytic, statistics that may limit the conclusions that can be drawn from data. Finally, we observed an oversaturation of audit-style studies drawing from clinical and laboratory databases. Whilst representing widespread international efforts to implement RGCS, these studies failed to contribute new findings and lacked patient-reported outcomes. As patient-centeredness is a core tenant of genetic counselling and medical practice, and it is well recognised that patient-reported data enriches information about relevant outcomes that reflect their experiences, we found the lack of patient voice in data collection concerning.[101-103] We identified a small number of well-designed studies that addressed biases, measured outcomes longitudinally, performed analytic statistics, and incorporated patient-reported outcomes; however, further work is needed to expand on the body of evidence they have created. Whilst randomised controlled trials are considered the gold-standard for generating unbiased research evidence, they are resource-intensive and may not be suitable for this context. Instead, efforts must be taken to ensure that future research on RGCS has clearly defined research questions that inform the study design, and that potential biases are addressed and minimised.

Capturing all reported outcomes from studies on RGCS provided insights into research priorities over the past decade. We identified an emphasis on delivery of care; focusing largely on barriers and facilitators to uptake, patient preferences, and satisfaction. This focus is not surprising considering that the context of RGCS has largely been in either increased risk populations with a public health imperative to reduce disease incidence, or broadly through commercial initiatives with a financial interest in uptake. Therefore, measuring uptake has been widely used to illustrate the acceptability of RGCS and provide a rationale for its continued offer. The skewing towards operational outcomes such as uptake however, results in a lack of insight into the patient experience and limits understanding of the benefits and harms of testing. It is evident that the relevance of outcomes being assessed needs further consideration and could benefit from the inclusion of patients at the inception of research design. Funding bodies are increasingly placing emphasis on consumer and community engagement, and systematic reviews have highlighted the positive impact that patient and public involvement can have on research quality, appropriateness and relevance.[104,105] Previous core outcome sets involving patients in the design and conduct of research led to the identification of outcomes that were not defined by health professionals or researchers alone.[16] We did not find any evidence of patient involvement in the design of research or selection of outcomes in this review. The absence of patient involvement at the outset of study design and under-representation of patient-reported outcomes in the RGCS literature emphasises the need for a clearer patient voice in future research.

Demonstrating the clinical utility of a health intervention is a central aim of research, however this review found that clinical utility has not been clearly illustrated for RGCS. When considering the goals of RGCS, two perspectives on clinical utility are apparent; a reduction in disease incidence, or the provision of information to allow reproductive autonomy and informed decision-making. Whilst most studies in this review did not define a primary outcome, reduction in disease incidence was frequently inferred as a primary outcome. This is problematic from an ethical perspective, as a focus on reducing incidence may be perceived as undervaluing the lives of those currently living with genetic conditions.[106] It is increasingly evident that the clinical utility of RGCS is more appropriately reflected by the latter perspective. Furthermore, an important element of clinical utility is timing, as the usefulness of information provided by RGCS is contingent upon whether patients have sufficient time to consider their options, are not precluded from options due to advanced gestation, and are not being put at risk for regret or poor psychological outcomes insofar as is possible in a prenatal setting. Most studies did not report any aspect of timeliness, and in those that did, we found a lack of consensus for how to do so. Previous systematic reviews have highlighted that despite RGCS being ideally conducted preconceptionally, testing during pregnancy remains prevalent.[11,12] Even as awareness of RGCS increases, many people may not appreciate its importance or be motivated to pursue it until they are pregnant, may have unexpected pregnancies, or be subject to health disparities that limit access. As long as RGCS continues to be offered in a prenatal setting, providing clarity around outcomes that account for timing will be crucial to evaluate screening programs and appropriately capture clinical utility for patients.

Evidence-based practice recommendations provide crucial guidance to practitioners regarding the safe and effective implementation of health interventions. We identified significant gaps in the body of evidence used to inform practice recommendations, which likely accounts for the reliance on consensus-based recommendations to date. The informed and voluntary nature of decision-making was compromised by a focus on uptake, which is an insufficient proxy for informed choice. A previous systematic review highlighted that many people accept screening simply because it is offered and not due to perceived benefits.[26] More informative measures of deliberation, informed choice, and decisional satisfaction/regret were identified in

this review, albeit less frequently, and should be a focus of future research to ensure that data representative of informed and voluntary screening is available for evaluation. Studies rarely assessed the quality of pre-test information and counselling. Some studies assessed patient satisfaction, which can be a valuable indicator of the quality of genetic counselling and information provision. Notably, despite validated satisfaction scales for genetic counselling being available, these were not utilised in any studies.[107] Ensuring that patients receive appropriate pre-test counselling will become more important as a diverse range of health professionals become involved in offering RGCS and as testing is scaled to encompass the general population. It is imperative to ensure that appropriate standards of knowledge are being fostered to meet evaluation criteria that strive towards informed decision-making.

Evaluation criteria also aim to understand the potential adverse outcomes of a health intervention in order to minimise harm to patients. In the context of RGCS, perceived risks include impacts on psychological wellbeing and possible social consequences, such as discrimination or stigmatisation. Previous reviews have indicated that psychological distress may occur at various stages of the screening process but is often not sustained.[12,108] However, the heterogeneity observed in the outcomes used to assess psychological wellbeing in this review, coupled with the variability in measurement methods and time points, detracts from how confidently we can draw conclusions about potential psychological harms. The impact of potential biases that we found in this review may also have flow-on implications for measuring psychological outcomes, including selective reporting biases which may skew published evidence towards favourable outcomes and selection bias which may limit generalisability of findings to the wider population. Many RGCS programs incur an out-of-pocket cost to participants and favour higher socioeconomic groups; as RGCS becomes accessible to the general population, it is crucial to establish the external validity of existing findings by evaluating psychological outcomes in more diverse populations.[100] Deciding on which outcome(s) best capture psychological wellbeing, minimising selection bias, and ensuring transparent reporting of all outcomes regardless of the results, will be necessary to provide greater certainty that RGCS results in negligible psychological harms. In regards to social consequences, little has been done to address these. We identified a small number of studies that considered impact of results on relationships and potential for stigmatisation or discrimination, however further work is needed to more fully understand the social consequences of RGCS.

At an overarching level, heterogeneity in a research dataset limits direct comparisons between studies on the same topic and indicates a lack of agreement for which outcomes can meaningfully represent the impact of an intervention. Where heterogeneity occurs, the ability to capture benefits and harms is compromised.[16,96] Future directions for research will involve clarifying what outcomes are valued by all key stakeholders in RGCS, including consumers, health professionals, researchers and policy-makers. Such research will need to address numerous issues highlighted in this review, starting with what outcomes should be measured, followed by how and when. Further exploration of outcomes related to limitations of RGCS, including patient understanding of residual risks, practical aspects of RGCS, such as ensuring appropriate storage and accessibility of results over time, and methodological aspects of research, including development of validated measurement tools, are areas of interest for future research. Researchers should strive to minimise bias in the design, conduct and reporting of their findings and consider making available a transparent protocol for their research that allows the methods to be clear and reproducible. Evaluation criteria used to assess genetic screening programs should be considered when designing research questions for future studies to ensure that findings are informative and work towards the goal of evidence-based practice recommendations.

Limitations

Publications not available in English were excluded due to a lack of resources for translation; nevertheless, we achieved representation from 15 countries. The iterative, inductive process used to extract outcomes and group them into domains may introduce biases from the reviewer; however, we minimised this by applying two independent reviewers and evaluating final domains and outcome groupings with the study management group. This review includes only quantitative data from the RGCS literature. More patient-centred outcomes will likely be evident in the qualitative literature, planned as a subsequent review by these authors.

Conclusion

Lack of consensus regarding outcomes to measure in the evaluation of RGCS perpetuates our inability to definitively demonstrate the impact, benefits and harms of RGCS at the standards required for evidence-based practice recommendations. Consensus on how to approach future research on this topic, including consideration of appropriate study designs that reduce bias, enrich understanding through the capture of longitudinal outcomes, and incorporate relevant outcomes informed by patients and other stakeholders, is needed. This review provides a strong rationale for the development of a core outcome set for RGCS.[17]

Data Availability

No datasets were generated or analysed during the current study.

Acknowledgements

This study is supported by the University of Technology Sydney and Graduate School of Health in the form of the Australian Research Training Program Fee Waiver Scholarship and Research Excellence Scholarship. The authors would like to thank Dr Karine Manera for her review and comments on this manuscript.

Author Contributions

Ebony Richardson (conceptualisation, data curation, investigation, formal analysis, project administration, visualisation, writing – original draft, writing – review and editing)
A/Prof Alison McEwen (supervision, writing - review and editing)
A/Prof Toby Newton-John (supervision, writing - review and editing)
Ashley Crook (data curation, writing – review and editing)
Dr Chris Jacobs (supervision, writing - review and editing)

Ethics Declaration

Ethics approval was not required for this review. The authors declare no conflicts of interest.

Conflict of Interest

No conflicts to disclose.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2021.08.005) contains supplementary material, which is available to authorized users.

References

1. Delatycki MB, Alkuraya F, Archibald A, et al. International perspectives on the implementation of reproductive carrier screening. *Prenat Diagn*. 2020;40(3):301–310.
2. Committee on Genetics. Committee Opinion No. 691: Carrier screening for genetic conditions. *Obstet Gynecol*. 2017;129(3):e41–e55.
3. Genomics Advisory Working Group & Women's Health Committee. Genetic carrier screening. In: RANZCOG; 2019.
4. Wilson RD, De Bie I, Armour CM, et al. Joint SOGC-CCMG Opinion for Reproductive Genetic Carrier Screening: An Update for All Canadian Providers of Maternity and Reproductive Healthcare in the Era of Direct-to-Consumer Testing. *J Obstet Gynaecol Can*. 2016;38(8):742–762 e743.
5. Bell N, Connor Gorber S, Tonelli M, et al. From ABCs to GRADE: Canadian task force on preventive health care's new rating system for clinical practice guidelines. *Can Fam Physician*. 2013;59(12):1282–1289.
6. National Health and Medical Research Council (NHMRC). *Additional levels of evidence and grades for recommendations for developers of guidelines*. Melbourne, VIC, Australia: NHMRC; 2009.
7. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71–72.
8. Ames AG, Metcalfe SA, Dalton Archibald A, Duncan RE, Emery J. Measuring informed choice in population-based reproductive genetic screening: a systematic review. *Eur J Hum Genet*. 2015;23(1):8–21.
9. Cannon J, Van Steijvoort E, Borry P, Chokoshvili D. How does carrier status for recessive disorders influence reproductive decisions? A systematic review of the literature. *Expert Rev Mol Diagn*. 2019;19(12):1117–1129.
10. Chen LS, Goodson P. Factors affecting decisions to accept or decline cystic fibrosis carrier testing/screening: a theory-guided systematic review. *Genet Med*. 2007;9(7):442–450.
11. Hill MK, Archibald AD, Cohen J, Metcalfe SA. A systematic review of population screening for fragile X syndrome. *Genet Med*. 2010;12(7):396–410.
12. Ioannou L, McClaren BJ, Massie J, et al. Population-based carrier screening for cystic fibrosis: a systematic review of 23 years of research. *Genet Med*. 2014;16(3):207–216.
13. Van Steijvoort E, Chokoshvili D, J WC, et al. Interest in expanded carrier screening among individuals and couples in the general population: systematic review of the literature. *Hum Reprod Update*. 2020;26(3):335–355.
14. Hussein N, Weng SF, Kai J, Kleijnen J, Qureshi N. Preconception risk assessment for thalassaemia, sickle cell disease, cystic fibrosis and Tay-Sachs disease. *Cochrane Database Syst Rev*. 2018;3:CD010849.
15. Kornman L, Chambers H, Nisbet D, Liebelt J. Pre-conception and antenatal screening for the fragile site on the X-chromosome. *Cochrane Database Syst Rev*. 2002;(1):CD001806.
16. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: Version 1.0. *Trials*. 2017;18(Suppl 3):280.
17. Richardson E, McEwen A, Newton-John T, Manera K, Jacobs C. The Core Outcome DEvelopment for Carrier Screening (CODECS) study: protocol for development of a core outcome set. *Trials*. 2021;22(1):480.
18. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Med*. 2009;3(3):e123–e130.
19. Kmet LM, Cook LS, Lee RC. *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*. Edmonton: Alberta Heritage Foundation for Medical Research; 2004.
20. Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One*. 2008;3(8):e3081.
21. Kirkham JJ, Dwan KM, Altman DG, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*. 2010;340:c365.
22. QSR International Pty Ltd. (2018) NVivo (Version 12). *https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home*.
23. Sut N. Study designs in medicine. *Balkan Med J*. 2014;31(4):273–277.
24. Popay J, Roberts H, Sowden A, et al. *Guidance on the conduct of narrative synthesis in sytematic reviews*. London: Institute for Health Research; 2006.
25. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol*. 2018;96:84–92.
26. Henneman L, Poppelaars FA, ten Kate LP. Evaluation of cystic fibrosis carrier screening programs according to genetic screening criteria. *Genet Med*. 2002;4(4):241–249.
27. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Geneva: World Health Organization; 1968. Available from: http://apps.who.int/iris/bitstream/10665/37650/17/WHO_PHP_34.pdf.

28. Molster CM, Lister K, Metternick-Jones S, et al. Outcomes of an international workshop on preconception expanded carrier screening: Some considerations for governments. *Front Public Health*. 2017;5:25.
29. Abi Saad M, Haddad AG, Alam ES, et al. Preventing thalassemia in Lebanon: successes and challenges in a developing country. *Hemoglobin*. 2014;38(5):308–311.
30. Aharoni S, Nevo Y, Orenstein N, et al. Impact of a national population-based carrier-screening program on spinal muscular atrophy births. *Neuromuscular Disorders*. 2020;30(12):970–974. http://doi.org/10.1016/j.nmd.2020.10.005.
31. Akler G, Birch AH, Schreiber-Agus N, et al. Lessons learned from expanded reproductive carrier screening in self-reported Ashkenazi, Sephardi, and Mizrahi Jewish patients. *Mol Genet Genomic Med*. 2020;8(2):e1053.
32. Amato A, Cappabianca MP, Lerone M, et al. Carrier screening for inherited haemoglobin disorders among secondary school students and young adults in Latium, Italy. *J Community Genet*. 2014;5(3):265–268.
33. Archibald AD, Smith MJ, Burgess T, et al. Reproductive genetic carrier screening for cystic fibrosis, fragile X syndrome, and spinal muscular atrophy in Australia: outcomes of 12,000 tests. *Genet Med*. 2018;20(5):513–523.
34. Baxi A, Manila K, Kadhi P, Heena B. Carrier Screening for beta thalassemia in pregnant indian women: Experience at a single center in Madhya Pradesh. *Indian J Hematol Blood Transfus*. 2013;29(2):71–74.
35. Belhoul KM, Abdulrahman M, Alraei RF. Hemoglobinopathy carrier prevalence in the United Arab Emirates: first analysis of the Dubai Health Authority premarital screening program results. *Hemoglobin*. 2013;37(4):359–368.
36. Ben-Shachar S, Orr-Urtreger A, Bardugo E, Shomrat R, Yaron Y. Large-scale population screening for spinal muscular atrophy: clinical implications. *Genet Med*. 2011;13(2):110–114.
37. Brown K, Dormandy E, Reid E, Gulliford M, Marteau T. Impact on informed choice of offering antenatal sickle cell and thalassaemia screening in primary care: a randomized trial. *J Med Screen*. 2011;18(2):65–75.
38. Canatan D, Delibas S. Report on ten years' experience of premarital hemoglobinopathy screening at a center in Antalya, Southern Turkey. *Hemoglobin*. 2016;40(4):273–276.
39. Carlotti K, Hines K, Weida J, Lah M, Schwantes-An TH. Perceived barriers to paternal expanded carrier screening following a positive maternal result: To screen or not to screen. *J Genet Couns*. 2021;30(2):470–477. http://doi.org/10.1002/jgc4.1333.
40. Castellani C, Picci L, Tridello G, et al. Cystic fibrosis carrier screening effects on birth prevalence and newborn screening. *Genet Med*. 2016;18(2):145–151.
41. Clarke EV, Schneider JL, Lynch F, et al. Assessment of willingness to pay for expanded carrier screening among women and couples undergoing preconception carrier screening. *PLoS One*. 2018;13(7):e0200139.
42. Coiana A, Faa V, Carta D, Puddu R, Cao A, Rosatelli MC. Preconceptional identification of cystic fibrosis carriers in the Sardinian population: A pilot screening program. *J Cyst Fibros*. 2011;10(3):207–211.
43. Curd H, Lewis S, Macciocca I, et al. High school Tay-Sachs disease carrier screening: 5 to 11-year follow-up. *J Community Genet*. 2014;5(2):139–146.
44. Dormandy E, Bryan S, Gulliford MC, et al. Antenatal screening for haemoglobinopathies in primary care: a cohort study and cluster randomised trial to inform a simulation model. The Screening for Haemoglobinopathies in First Trimester (SHIFT) trial. *Health Technol Assess*. 2010;14(20):1–160.
45. Dormandy E, Gulliford M, Bryan S, et al. Effectiveness of earlier antenatal screening for sickle cell disease and thalassaemia in primary care: cluster randomised trial. *BMJ*. 2010;341:c5132.
46. Ghiossi CE, Goldberg JD, Haque IS, Lazarin GA, Wong KK. Clinical utility of expanded carrier screening: reproductive behaviors of at-risk couples. *J Genet Couns*. 2018;27(3):616–625.
47. Ghosh M, Basu K, Karmakar S, et al. Thalassaemia carrier detection during antenatal period: Single centre experience from Eastern India. *J Clin Diagn Res*. 2019;13(12):EC01–EC05.
48. Giambona A, Damiani G, Vinciguerra M, et al. Incidence of haemoglobinopathies in Sicily: the impact of screening and prenatal diagnosis. *Int J Clin Pract*. 2015;69(10):1129–1138.
49. Giles Choates M, Stevens BK, Wagner C, Murphy L, Singletary CN, Wittman AT. It takes two: uptake of carrier screening among male reproductive partners. *Prenat Diagn*. 2020;40(3):311–316.
50. Gilmore MJ, Schneider J, Davis JV, et al. Reasons for declining preconception expanded carrier screening using genome sequencing. *J Genet Couns*. 2017;26(5):971–979.
51. Grinzaid KA, Page PZ, Denton JJ, Ginsberg J. Creation of a national, at-home model for Ashkenazi Jewish carrier screening. *J Genet Couns*. 2015;24(3):381–387.
52. Guler E, Garipardic M, Dalkiran T, Davutoglu M. Premarital screening test results for beta-thalassemia and sickle cell anemia trait in east Mediterranean region of Turkey. *Pediatr Hematol Oncol*. 2010;27(8):608–613.
53. Hardy MW, Kener HJ, Grinzaid KA. Implementation of a carrier screening program in a high-risk undergraduate student population using digital marketing, online education, and telehealth. *Public Health Genomics*. 2018;21(1-2):67–76.
54. Ioannou L, Massie J, Collins V, McClaren B, Delatycki MB. Population-based genetic screening for cystic fibrosis: attitudes and outcomes. *Public Health Genomics*. 2010;13(7-8):449–456.
55. Ioannou L, Massie J, Lewis S, McClaren B, Collins V, Delatycki MB. 'No thanks'-reasons why pregnant women declined an offer of cystic fibrosis carrier screening. *J Community Genet*. 2014;5(2):109–117.
56. Ioannou L, Massie J, Lewis S, et al. Evaluation of a multi-disease carrier screening programme in Ashkenazi Jewish high schools. *Clin Genet*. 2010;78(1):21–31.
57. Jiang F, Chen GL, Li J, et al. Pre-gestational thalassemia screening in mainland China: the first two years of a preventive program. *Hemoglobin*. 2017;41(4-6):248–253.
58. Johansen Taber K, Lim-Harashima J, Naemi H, Goldberg J. Fragile X syndrome carrier screening accompanied by genetic consultation has clinical utility in populations beyond those recommended by guidelines. *Mol Genet Genomic Med*. 2019;7(12):e1024.
59. Johansen Taber KA, Beauchamp KA, Lazarin GA, Muzzey D, Arjunan A, Goldberg JD. Clinical utility of expanded carrier screening: results-guided actionability and outcomes. *Genet Med*. 2019;21(5):1041–1048.
60. Kaufmann JO, Demirel-Gungor G, Selles A, et al. Feasibility of nonselective testing for hemoglobinopathies in early pregnancy in The Netherlands. *Prenat Diagn*. 2011;31(13):1259–1263.
61. Kim MJ, Kim DJ, Kim SY, et al. Fragile X carrier screening in Korean women of reproductive age. *J Med Screen*. 2013;20(1):15–20.
62. Kraft SA, Schneider JL, Leo MC, et al. Patient actions and reactions after receiving negative results from expanded carrier screening. *Clin Genet*. 2018;93(5):962–971.
63. Kuhl A, van Calcar S, Baker M, Seroogy CM, Rice G, Scott Schwoerer J. Development of carrier testing for common inborn errors of metabolism in the Wisconsin Plain population. *Genet Med*. 2017;19(3):352–356.
64. Ladis V, Karagiorga-Lagana M, Tsatra I, Chouliaras G. Thirty-year experience in preventing haemoglobinopathies in Greece: achievements and potentials for optimisation. *Eur J Haematol*. 2013;90(4):313–322.
65. Larsen D, Ma J, Strassberg M, Ramakrishnan R, Van den Veyver IB. The uptake of pan-ethnic expanded carrier screening is higher when offered during preconception or early prenatal genetic counseling. *Prenat Diagn*. 2019;39(4):319–323.
66. Lazarin GA, Haque IS, Nazareth S, et al. An empirical estimate of carrier frequencies for 400+ causal Mendelian variants: results from

an ethnically diverse clinical sample of 23,453 individuals. *Genet Med*. 2013;15(3):178–186.
67. Lew RM, Proos AL, Burnett L, Delatycki M, Bankier A, Fietz MJ. Tay Sachs disease in Australia: reduced disease incidence despite stable carrier frequency in Australian Jews. *Med J Aust*. 2012;197(11):652–654.
68. Liao C, Pan M, Han J, et al. Prenatal control of Hb Bart's hydrops fetalis: a two-year experience at a mainland Chinese hospital. *J Matern Fetal Neonatal Med*. 2015;28(4):413–415.
69. Macarov M, Zlotogora J, Meiner V, et al. Genetic screening for Krabbe disease: learning from the past and looking to the future. *Am J Med Genet A*. 2011;155A(3):574–576.
70. Mathijssen IB, Henneman L, van Eeten-Nijman JM, et al. Targeted carrier screening for four recessive disorders: high detection rate within a founder population. *Eur J Med Genet*. 2015;58(3):123–128.
71. Mathijssen IB, Holtkamp KCA, Ottenheim CPE, et al. Preconception carrier screening for multiple disorders: evaluation of a screening offer in a Dutch founder population. *Eur J Hum Genet*. 2018;26(2):166–175.
72. Metcalfe SA, Martyn M, Ames A, et al. Informed decision making and psychosocial outcomes in pregnant and nonpregnant women offered population fragile X carrier screening. *Genet Med*. 2017;19(12):1346–1355.
73. Monni G, Peddes C, Iuculano A, Ibba RM. From prenatal to preimplantation genetic diagnosis of beta-thalassemia. Prevention model in 8748 cases: 40 years of single center experience. *J Clin Med*. 2018;7(2).
74. Picci L, Cameran M, Marangon O, et al. A 10-year large-scale cystic fibrosis carrier screening in the Italian population. *J Cyst Fibros*. 2010;9(1):29–35.
75. Prior TW, Snyder PJ, Rink BD, et al. Newborn and carrier screening for spinal muscular atrophy. *Am J Med Genet A*. 2010;152A(7):1608–1616.
76. Propst L, Connor G, Hinton M, Poorvu T, Dungan J. Pregnant women's perspectives on expanded carrier screening. *J Genet Couns*. 2018;27(5):1148–1156.
77. Punj S, Akkari Y, Huang J, et al. Preconception carrier screening by genome sequencing: results from the clinical laboratory. *Am J Hum Genet*. 2018;102(6):1078–1089.
78. Robson SJ, Caramins M, Saad M, Suthers G. Socioeconomic status and uptake of reproductive carrier screening in Australia. *Aust N Z J Obstet Gynaecol*. 2020;60(6):976–979. http://doi.org/10.1111/ajo.13206.
79. Schuurmans J, Birnie E, Ranchor AV, et al. GP-provided couple-based expanded preconception carrier screening in the Dutch general population: who accepts the test-offer and why? *Eur J Hum Genet*. 2020;28(2):182–192.
80. Schuurmans J, Birnie E, van den Heuvel LM, et al. Feasibility of couple-based expanded carrier screening offered by general practitioners. *Eur J Hum Genet*. 2019;27(5):691–700.
81. Scott SA, Edelmann L, Liu L, Luo M, Desnick RJ, Kornreich R. Experience with carrier screening and prenatal diagnosis for 16 Ashkenazi Jewish genetic diseases. *Hum Mutat*. 2010;31(11):1240–1250.
82. Shao Y, Liu S, Grinzaid K. Evaluation of two-year Jewish genetic disease screening program in Atlanta: insight into community genetic screening approaches. *J Community Genet*. 2015;6(2):137–145.
83. Simone L, Khan S, Ciarlariello M, et al. Reproductive male partner testing when the female is identified to be a genetic disease carrier. *Prenat Diagn*. 2021;41(1):21–27. http://doi.org/10.1002/pd.5824.
84. Singer A, Sagi-Dain L. Impact of a national genetic carrier-screening program for reproductive purposes. *Acta Obstet Gynecol Scand*. 2020;99(6):802–808.
85. Su YN, Hung CC, Lin SY, et al. Carrier screening for spinal muscular atrophy (SMA) in 107,611 pregnant women during the period 2005-2009: a prospective population-based cohort study. *PLoS One*. 2011;6(2):e17067.
86. Theodoridou S, Prapas N, Balassopoulou A, et al. Efficacy of the national thalassaemia and sickle cell disease prevention programme in Northern Greece: 15-year experience, practice and policy gaps for natives and migrants. *Hemoglobin*. 2018;42(4):257–263.
87. Tongsong T, Charoenkwan P, Sirivatanapa P, et al. Effectiveness of the model for prenatal control of severe thalassemia. *Prenat Diagn*. 2013;33(5):477–483.
88. Tzeng CC, Tsai LP, Chang YK, et al. A 15-year-long southern blotting analysis of FMR1 to detect female carriers and for prenatal diagnosis of fragile X syndrome in Taiwan. *Clin Genet*. 2017;92(2):217–220.
89. Warsch JR, Warsch S, Herman E, et al. Knowledge, attitudes, and barriers to carrier screening for the Ashkenazi Jewish panel: a Florida experience : Education and Barriers assessment for Jewish Genetic Diseases. *J Community Genet*. 2014;5(3):223–231.
90. Westemeyer M, Saucier J, Wallace J, et al. Clinical experience with carrier screening in a general population: support for a comprehensive pan-ethnic approach. *Genet Med*. 2020;22(8):1320–1328.
91. Yamsri S, Sanchaisuriya K, Fucharoen G, Sae-Ung N, Ratanasiri T, Fucharoen S. Prevention of severe thalassemia in northeast Thailand: 16 years of experience at a single university center. *Prenat Diagn*. 2010;30(6):540–546.
92. Yip T, Grinzaid KA, Bellcross C, Moore RH, Page PZ, Hardy MW. Patients' reactions and follow-up testing decisions related to Tay-Sachs (HEXA) variants of uncertain significance results. *J Genet Counsel*. 2019;28(4):738–749.
93. Zhang J, Wang Y, Ma D, et al. Carrier screening and prenatal diagnosis for spinal muscular atrophy in 13,069 Chinese pregnant women. *J Mol Diagn*. 2020;22(6):817–822.
94. Kauffman TL, Wilfond BS, Jarvik GP, et al. Design of a randomized controlled trial for genomic carrier screening in healthy patients seeking preconception genetic testing. *Contemp Clin Trials*. 2017;53:100–105.
95. Martyn M, Anderson V, Archibald A, et al. Offering fragile X syndrome carrier screening: a prospective mixed-methods observational study comparing carrier screening of pregnant and non-pregnant women in the general population. *BMJ Open*. 2013;3(9):e003660.
96. Gargon E, Gurung B, Medley N, et al. Choosing important health outcomes for comparative effectiveness research: a systematic review. *PLoS One*. 2014;9(6):e99111.
97. Athens BA, Caldwell SL, Umstead KL, Connors PD, Brenna E, Biesecker BB. A systematic review of randomized controlled trials to assess outcomes of genetic counseling. *J Genet Couns*. 2017;26(5):902–933.
98. Payne K, Nicholls S, McAllister M, Macleod R, Donnai D, Davies LM. Outcome measurement in clinical genetics services: a systematic review of validated measures. *Value Health*. 2008;11(3):497–508.
99. Senter L, Austin JC, Carey M, et al. Advancing the genetic counseling profession through research: Identification of priorities by the National Society of Genetic Counselors research task force. *J Genet Couns*. 2020;29(6):884–887. http://doi.org/10.1002/jgc4.1330.
100. Wang C, Gonzalez R, Merajver SD. Assessment of genetic testing and related counseling services: current research and future directions. *Soc Sci Med*. 2004;58(7):1427–1442.
101. Brown JB, Freeman T, McWilliam CL, McWhinney IR, Weston W, Stewart M. *Patient-Centered Medicine: Transforming the Clinical Method*. 3rd ed. CRC Press; 2013.
102. Clarke A. The Evolving Concept of Non-directiveness in Genetic Counselling. In: Petermann HI, Harper PS, Doetz S, eds. *History of Human Genetics: Aspects of Its Development and Global Perspectives*. Springer International Publishing; 2017:541–566.
103. Austin E, LeRouge C, Hartzler AL, Segal C, Lavallee DC. Capturing the patient voice: implementing patient-reported outcomes across the health system. *Qual Life Res*. 2020;29(2):347–355.
104. National Health and Medical Research Council. *Statement on Consumer and Community involvement in Health and Medical Research*. Consumers Health Forum of Australia; 2016.

105. Brett J, Staniszewska S, Mockford C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. *Health Expect*. 2014;17(5):637–650.
106. Dive L, Newson AJ. Ethical issues in reproductive genetic carrier screening. *Med J Aust*. 2021;214(4):165–167, e161.
107. Kasparian NA, Wakefield CE, Meiser B. Assessment of psychosocial outcomes in genetic counseling research: an overview of available measurement scales. *J Genet Couns*. 2007;16(6):693–712.
108. Henneman L, Borry P, Chokoshvili D, et al. Responsible implementation of expanded carrier screening. *Eur J Hum Genet*. 2017;25(11):1291.