

The Core Outcome Development for Carrier Screening (CODECS) Study: Protocol for Development of a Core Outcome Set

Ebony Joanne Richardson (✉ 13446531@student.uts.edu.au)

University of Technology Sydney <https://orcid.org/0000-0003-4260-7356>

Alison McEwen

University of Technology Sydney Graduate School of Health

Toby Newton-John

University of Technology Sydney Graduate School of Health

Karine Manera

The University of Sydney

Chris Jacobs

University of Technology Sydney Graduate School of Health

Study protocol

Keywords: Core outcome set, Reproductive genetic carrier screening, Genetic counselling, Patient-reported outcomes, Qualitative research, Delphi Survey, Outcome reporting

Posted Date: December 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-115947/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Trials on July 22nd, 2021. See the published version at <https://doi.org/10.1186/s13063-021-05439-7>.

Abstract

Background: Reproductive genetic carrier screening is a type of genetic testing available to those planning a pregnancy, or during their first trimester, to understand their risk of having a child with a severe genetic condition. There is a lack of consensus for 'what to measure' in studies on this intervention, leading to heterogeneity in choice of outcomes and methods of measurement. Such outcome heterogeneity has implications for the quality and comparability of these studies and has led to a lack of robust research evidence in the literature to inform policy and decision-making around the offer of this screening. As reproductive genetic carrier screening becomes increasingly accessible within the general population, it is timely to investigate the outcomes of this intervention.

Objectives: The development of a core outcome set is an established methodology to address issues with outcome heterogeneity in research. We aim to develop a core outcome set for reproductive genetic carrier screening to clarify and standardise outcomes for research and practice.

Methods: In accordance with guidance from the COMET (Core Outcome Measures in Effectiveness Trials) Initiative, this study will consist of five steps: i) A systematic review of quantitative studies, using narrative synthesis to identify previously reported outcomes, their definitions, and methods of measurement ii) A systematic review of qualitative studies using content analysis to identify excerpts related to patient experience and perspectives that can be interpreted as outcomes iii) Semi-structured focus groups and interviews with patients who have undertaken reproductive genetic carrier screening to identify outcomes of importance to them iv) Delphi survey of key stakeholders, including patients, clinicians, and researchers, to refine and prioritise the list of outcomes generated from the previous steps v) A virtual consensus meeting with a purposive sample of key stakeholders to finalise the core outcome set for reporting.

Discussion: This protocol outlines the core outcome set development process and its novel application in the setting of genetic testing. This core outcome set will support the standardisation of outcome reporting in reproductive carrier screening research and contribute to an evolving literature on outcomes to evaluate genetic testing and genetic counselling as health interventions.

COMET core outcome set registration: <http://www.comet-initiative.org/Studies/Details/1381>

Background

Genetic testing is an increasingly common health intervention across numerous clinical settings and is recognised to hold vast potential for improving patient care. Genetic counselling is the process that surrounds the consideration of genetic risk, encompassing many aspects directly related to genetic testing. Genetic counselling can be performed by genetic counsellors educated at the post-graduate level and regulated by a professional body. Genetic counselling provided by genetic counsellors includes a number of aspects separate to whether a genetic test is undertaken, for example, facilitating understanding, providing individual and family support, and assisting clients with adjustment to genetic

risk.⁽¹⁾ However, aspects of genetic counselling are also undertaken by a range of other health professionals involved in patient care. As the processes of genetic testing and genetic counselling are intricately entwined, so too are the health-related outcomes that are measured in research in this area. Most research in the field of genetics does not distinguish between outcomes of genetic testing and genetic counselling; therefore, both will be considered and referred to herein.

One of the most widely adopted genetic tests worldwide is reproductive genetic carrier screening (RGCS), which provides individuals and couples with information about their risk of having a child with a genetic condition before or during early pregnancy.⁽²⁾ RGCS identifies carriers of recessively inherited conditions (autosomal recessive or X-linked), such as cystic fibrosis, spinal muscular atrophy, and Fragile X syndrome. These conditions often arise unexpectedly and carriers are, in most instances, asymptomatic. Since there is usually only an increased reproductive risk if a carrier chances to partner with a carrier of the same condition, most couples that have an affected child will not have an existing family history that could have forewarned of their risk. Recessively inherited conditions are individually rare but when combined, are estimated to affect at least 30 in every 10,000 or 0.3% of births.⁽³⁻⁵⁾ Based on this birth prevalence, it is estimated that 1–2% of couples will be at risk for having a child affected with a genetic condition, and this number can be much higher in consanguineous populations.⁽⁶⁾ Of those at an increased risk, the likelihood of having an affected child ranges from 25%-50% in each pregnancy, depending on the specific condition. The intent of RGCS is to provide couples who are at increased risk with information to allow them to make informed reproductive decisions. Those who are aware of their risk can choose to pursue prenatal diagnosis during pregnancy, opt for in-vitro fertilisation (IVF) with preimplantation genetic diagnosis or the use of a donor gamete, consider adoption, or pursue pregnancy without any intervention and diagnose postnatally if desired. For this protocol, individuals and couples undertaking RGCS will be referred to as patients, however we acknowledge that these will be largely healthy adults, most of which will not go on to require significant medical follow-up as a result of their carrier screening results.

Carrier screening programs have been implemented since the 1970s in populations that have increased carrier frequencies for certain conditions, with targeted testing of only the conditions relevant to that population. Such conditions include but are not limited to Tay Sachs disease in the Ashkenazi Jewish population, and thalassaemia and other inherited haemoglobinopathies across a range of ethnicities.⁽⁷⁻⁹⁾ These programs pre-dated our ability to identify carriers through genetic testing, instead relying on biochemical assays, with cystic fibrosis being one of the first conditions to have a screening program based on molecular methods introduced in the 1990's.⁽¹⁰⁾ Early carrier screening programs typically focused on one genetic condition; however recent advancements in genetic technologies have enabled a shift in the breadth of RGCS. Next generation sequencing has facilitated the development of 'expanded' panels that analyse hundreds to thousands of genetic conditions in a single laboratory test. These expanded panels are broadly available to the general population, and whilst they have predominantly been commercial offerings to date, largely limiting their uptake to high-income groups, there are emerging efforts internationally to support equitable access to expanded screening.⁽¹¹⁾ There are now a range of

ways in which individuals and couples may access RGCS, including community screening programs in increased risk populations, attending public or private prenatal services during early pregnancy, or accessing preconception care through general practitioners or genetic counsellors in the public or private sectors.

There is increasing support for RGCS to be offered widely. In 2016, the Society of Obstetricians and Gynaecologists of Canada Genetics Committee and the Canadian College of Medical Geneticists Clinical Practice Committee (SOGC-CCMG) released a joint practice recommendation supporting the discussion of RGCS with all women/families considering pregnancy or at their first prenatal visit.⁽¹²⁾ This advice was closely followed by a similar practice recommendation from the American College of Obstetricians and Gynaecologists (ACOG) in 2017.^(13, 14) These international organisations were amongst the first to support the widespread offer of RGCS outside of increased risk populations, with The Royal Australian and New Zealand College of Obstetricians and Gynaecologists (RANZCOG) following suit in March 2019.⁽¹⁵⁾ With this building momentum, it is a pivotal time to address research efforts to evaluate the impact of RGCS.

As with other areas of medicine, one of the aims of research in the field of genetics is to understand the benefits and harms of genetic testing as a health intervention. This is most often achieved by measuring the impact of a genetic test on patient outcomes when it is utilised in clinical practice; however this is acknowledged to be challenging.^(16, 17) There is an established literature aiming to demonstrate the effectiveness of clinical genetics services, genetic counselling, and genetic testing, and systematic reviews have overall demonstrated a modest positive impact.^(1, 18–23) A problem that has arisen frequently in the genetics literature is comparability across studies, with heterogeneity in the choice of outcomes and method of measurement. Often studies measure the same or similar concepts, such as psychological impact, but vary in the specific outcome that they report within this broad domain, utilise different measurement tools, and measure the outcome at variable time points. When outcome heterogeneity exists, the ability to directly compare and contrast the results of studies is hindered, and combining results, such as in a meta-analysis, becomes unreliable. This issue has been highlighted in research and commentary on the outcomes of genetic testing and genetic counselling and is becoming a focus of many discussions within the field.^(1, 18)

Another issue noted in genetics research is the propensity for observational study designs due to the challenges of including a comparison group in the clinical setting. Very few studies on RGCS are experimental in design, with only a handful of randomised controlled trials. Observational study designs are well-recognised to have a lower standard of methodological rigour, with a number of potential problems that may lead to biasing of results.⁽²⁴⁾ One such issue is that there is not a requirement or tendency to publish a protocol outlining the outcomes that will be measured. This introduces a risk of reporting bias as there is a lack of accountability for publishing all outcomes, regardless of whether they support the author's position or reach statistical significance. There is also a great deal of variability in the inclusion of patient-reported outcomes, which are important for ensuring that the results of the

research are relevant to patients. A small number of systematic reviews have been conducted in the field of RGCS, focusing on carrier screening for specific conditions. Those reviews that address data analysis and risk of bias in their methods, identified issues with outcome heterogeneity, study design, and overall quality of evidence, whilst others that didn't specifically address these issues performed narrative syntheses, which is indicative that a meta-analysis was not possible with the available data.^(23, 25–28)

We propose developing a core outcome set (COS) for RGCS. A COS is an agreed minimum set of outcomes that should be measured and reported in all studies on a particular topic.⁽²⁹⁾ The development of a COS applies a rigorous approach to defining outcomes that are relevant to all key stakeholders of a health intervention. This approach aims to minimise the heterogeneity in outcomes that are measured by different researchers, and as a result, maximise the ability to compare and combine studies in meta-analysis or other data synthesis approaches. Defining a COS also reduces the likelihood of reporting bias by ensuring that, at the very least, the core outcomes would be reported in all studies on an intervention. The incorporation of individuals who have had RGCS, clinicians involved in their care, and researchers and policy-makers guiding practice in this area in the development of this COS will ensure that outcomes are relevant to all stakeholders.

The Core Outcome DEvelopment in Carrier Screening (CODECS) study will apply the methodology outlined by the COMET (Core Outcomes Measures in Effectiveness Trials) Initiative to develop a COS for RGCS. To our knowledge, this study will be the first example of a COS aimed at standardising the reporting of outcomes in studies on a genetic testing intervention.

Methods/design

Scope:

The methodology defined by the COMET (Core Outcomes Measures in Effectiveness Trials) Initiative and the Core Outcome Set-STAndardised Protocol Items (COS-STAP) Statement will inform this protocol.^(29, 30) The COMET database was searched to confirm that there were no overlapping projects and the CODECS study subsequently registered (<http://www.comet-initiative.org/Studies/Details/1381>). The PICO framework is recommended by the COMET initiate for defining the scope of a core outcome development study, using the first three elements of population, intervention, and comparator.⁽³¹⁾ The population that this COS is being developed for incorporates any individual or couple that is offered genetic carrier screening to inform their current or future reproductive decisions. This may be offered as population screening in increased risk populations as well as the general population and includes school, community, preconception, and prenatal programs. This COS is not intended to cover cascade carrier screening in family members following the diagnosis of a genetic disease in a family member.

The definition of the intervention includes RGCS via targeted single-gene or small gene panels, through to pan-ethnic expanded carrier screening panels and virtual panels from whole-genome sequencing. The intervention encompasses pre-test genetic counselling, genetic testing, and post-test results

management. Molecular genetic testing methodologies are the predominant laboratory method of carrier screening currently. However, some programs do remain reliant on biochemical methods to triage access to molecular genetic panels. An example of this is haemoglobinopathies screening programs, which use results of mean corpuscular volume (MCV) and mean corpuscular haemoglobin (MCH) ± anaemia, to triage which individuals will be screened using molecular methods. This COS is intended to be applicable to molecular and combined biochemical/molecular methods. A comparator is not expected to be appropriate for most RGCS programs. However where appropriate, we will include comparators such as control populations (no RGCS testing) and targeted versus expanded interventions (single gene or small panels compared to expanded panels). This COS is intended to be applicable to all population-based RGCS scenarios and is being developed to take into account the significant variability in screening approaches used internationally.

This COS is being developed for use in research on RGCS, as well as in clinical practice. The majority of research in this area involves observational study designs assessing the impact of RGCS after it has already been implemented into clinical practice, and it is only recently that there has been a shift in the literature towards more rigorous study methodologies using randomised controlled designs. Therefore, it was decided that separating out the research and clinical contexts was not possible for this COS.

The CODECS study involves five steps: a systematic review of quantitative literature, a systematic review of qualitative literature, semi-structured focus groups/interviews with patient stakeholders, an international online Delphi survey, and a virtual consensus meeting (Fig. 1).

Step I – Systematic Review of Outcomes Reported in Quantitative Studies on Reproductive Genetic Carrier Screening

A systematic review of the literature will be conducted to identify outcomes and their method of measurement in studies evaluating an offer of RGCS. These will form the basis of the preliminary list of outcomes that will be reviewed and refined during the consensus process. The full protocol for this systematic review is published on the international prospective register of systematic reviews, PROSPERO (CRD42019140793).

Study Selection

MEDLINE, CINAHL, PsycINFO, and EMBASE will be searched for quantitative, qualitative, and mixed-methods studies. The quantitative and qualitative studies identified will be reviewed separately to account for the different approaches needed to extract the outcomes. Step I will include studies that are solely quantitative, or for mixed methods studies, the portion of quantitative data. A percentage of title and abstract screening will be performed independently by two reviewers until inter-rater reliability of > 85% is achieved, after which the remainder will be screened by ER only due to resource limitations. The

full-text screening will be similarly performed. Any disagreement on the eligibility of studies will be resolved through discussion with a third reviewer.

All peer-reviewed published studies where reproductive genetic carrier screening has been offered as a health intervention will be eligible for inclusion. Studies that are primarily evaluating laboratory test methodology, are not primary research, where the context of testing is not primarily related to RGCS (for example, newborn screening, cascade carrier screening), or that are not available in English, will be excluded.

Data extraction

Outcomes will be extracted from eligible studies from the last five years to form a preliminary list; the review will then proceed to the previous five years and compare outcomes to the preliminary list. If no additional outcomes are identified, the review will be considered complete after this 10-year period; however further cycles will be conducted if additional outcomes continue to be identified. This methodology is per the guidance of the COMET Handbook, suitable for situations where the size of the review would be unmanageable if conducted in full.⁽²⁹⁾ A guideline for data extraction has been developed and will be piloted with two independent reviewers for 20% of studies. Outcomes, and where supplied, their definition, method of measurement and time point, will be extracted verbatim from studies using NVivo software. The primary outcome, if specified, will be noted. In addition, study type, target population, intervention type, screening approach, and other basic study characteristics will be extracted.

Quality Assessment

The quality of the included studies will be scored using the QualSyst tool by the primary reviewer (ER).⁽³²⁾ In the context of a systematic review of outcomes where the aim is to determine all published outcomes regardless of study quality, the assessment of bias will not be used as grounds for exclusion but rather to give an overall evaluation of the quality of studies within the RGCS literature.

Analysis and Presentation of Results

A narrative synthesis will be conducted on data extracted from quantitative studies, with outcomes to be grouped within domains and mapped to the COMET taxonomy.^(33, 34) The domains will be reviewed and discussed by the CODECS study management group. The number of different outcomes (including the method of measurement and time points) and the number of studies that assessed each outcome will be evaluated. Subgroup analyses will be considered to identify outcomes that may be specific to targeted and expanded carrier screening approaches.

Step II: Systematic Review of Outcomes Reported in Qualitative Studies on Reproductive Genetic Carrier Screening

The inclusion of patient perspectives is considered a minimum standard of the COS development process in accordance with the Core Outcome Set-STAndards for Development (COS-STAD).⁽³⁵⁾ The COMET handbook suggests that the preliminary list of outcomes generated by a systematic review of the quantitative literature may be supplemented with additional domains derived from a review of the qualitative literature if time and resources allow.⁽²⁹⁾ We will apply the methodology developed by Gorst et al. to extract outcomes of importance to patients from the qualitative literature on RGCS.⁽³⁶⁾ We will compare the extracted outcomes identified from the qualitative literature with those identified from the quantitative literature. These outcomes will be used to identify gaps in knowledge or representation of patient groups to guide focus groups/interviews in Step 3.

Study Selection

The initial steps of the systematic review will be conducted in the same manner as the quantitative studies described above, diverging at the point of data extraction. Eligible studies will be those that utilise qualitative methodology or mixed methods studies involving a qualitative component.

Data Extraction

It is not anticipated that any existing studies will have conducted qualitative research specifically for the purpose of identifying outcomes. Therefore, our approach to data extraction will be deductive, taking excerpts verbatim from the qualitative literature and deducing the outcome that they represent. Excerpts will be any text relating to how patients felt or were impacted by their experience of undertaking RGCS, including quotations and author's interpretation of themes. Each relevant excerpt will be extracted as a node using NVivo software by the primary reviewer (ER) and checked by a second reviewer (CJ) to ensure that all relevant text excerpts have been extracted. Both reviewers will then independently interpret outcomes from 20% of studies and check these for agreement. A coding guideline will be developed and used by the primary reviewer (ER) to interpret the remaining excerpts.

Quality Assessment

The quality of the included studies will be scored as described above for quantitative studies.

Analysis and Presentation of Results

We will draw on content analysis to conduct a narrative synthesis of the data from eligible studies.⁽³⁴⁾ This method will allow us to convert qualitative findings into frequency counts, permitting comparison with the findings of the quantitative literature. Outcomes will be grouped within domains and mapped to the COMET taxonomy as above. Each text excerpt and its deduced outcome will be independently categorised into the taxonomy by two reviewers (ER and CJ). All categorisations will be discussed until 100% agreement is reached, with reference back to the original article for context as needed. Some outcomes are expected to be relevant to multiple domains within the taxonomy, and where this occurs, they will be categorised under two domains as recommended in the taxonomy guidance.⁽³³⁾

Step III: Semi-Structured Patient Focus Groups/Interviews

We will conduct primary qualitative research with patients who have undertaken RGCS to identify outcomes important to those with a lived experience of this process. We will give participants a choice of attending a focus group or one-on-one in-depth interview. Focus groups are a valuable way to encourage participant interaction and enrich the sharing of their experiences; however there are a number of factors that may influence the appropriateness of conducting focus groups.⁽³⁷⁾ It is possible that recruitment may be limited by factors related to the sensitivity of the research topic, in particular amongst participants that fall into the increased risk group and may feel uncomfortable relaying their experience in a group setting. Therefore, the option of one-on-one interviews will also be available and decided upon once recruitment is underway.

Participants and Recruitment

We will recruit individuals or couples who have had RGCS in order to inform their reproductive decisions. Participants groups will be defined by two characteristics; their level of risk prior to RGCS (a priori) and their level of risk following results (a posteriori). A priori risk will be either average or increased. Average a priori risk will be defined as *the participant having no existing health concerns or family history to indicate an increased risk of being a carrier*. Increased a priori risk will be defined as *the participant having an existing factor such as ethnicity with a known increased frequency of carriers or a known family of a genetic condition that is included in the screening*. There are a number of potential outcomes of RGCS; however, a posteriori risk will be grouped into either low or increased reproductive risk. Low reproductive risk results are defined as those where neither member of a reproductive couple are found to be carriers of the same genetic condition, or where one member of a reproductive couple is found to be a carrier of an autosomal recessively inherited genetic condition but their reproductive partner is not a carrier of the same condition. Increased reproductive risk is defined as those where both members of a reproductive couple are found to be carriers of the same autosomal recessively inherited genetic conditions, or where the female reproductive partner is found to be a carrier of an X-linked genetic condition. We will aim to recruit 30 participants, with equal representation from each group.

To recruit an international and diverse sample with a range of experiences, we will circulate an expression of interest to participate in the research through a number of social media channels and parenting forums. Respondents who follow the link will be directed to an online survey on the Qualtrics platform to receive more information about the research and fill out basic demographic and screening questions to confirm their eligibility.⁽³⁸⁾ Eligible participants will be from countries that score > 50 on the corruption perceptions index, indicating that they are not vulnerable populations.⁽³⁹⁾

Data Collection

Focus groups/interviews will be conducted using Zoom, an online audio- and video-conferencing platform.⁽⁴⁰⁾ Recent research has indicated the viability of Zoom as a tool for the collection of qualitative data due to its ease of use, cost-effectiveness, data management, and security features.⁽⁴¹⁾ Focus groups will be approximately 90 minutes, depending on the number of participants. One-on-one interviews will be

up to 60 minutes in duration. An online platform has been chosen to facilitate international participation and reduce the inconvenience of travelling for participants. Focus groups/interviews will be audio- and video-recorded and transcribed verbatim. The focus group/interview schedule will begin with open questions to elicit patient experiences, after which more specific questions related to outcomes will be informed from the list of outcome domains generated from the systematic review steps above. Our data collection will draw on grounded theory, with data collection and analysis occurring concurrently and utilising constant comparison to refine data collection as the study progresses. ⁽⁴²⁾

Data Analysis

Using thematic analysis, the transcripts will be reviewed line by line and inductively coded to identify concepts.⁽⁴³⁾ Similar concepts will be grouped into themes and corresponding subthemes. These concepts/themes will reflect the perspectives, beliefs, and values of participants in regards to outcomes of reproductive carrier screening. To ensure that the complete range and depth of the data are included, at least two investigators will be involved in coding the data. Data collection will continue until data saturation is reached (the point at which no new themes are identified).

Step IV: Delphi Consensus Survey

We will follow published principles of applying the Delphi process in the context of COS development.⁽⁴⁴⁾ This process will utilise a sequential, two to three round online survey with an internationally representative sampling of key stakeholders in the field.

Developing the survey

The preliminary list of outcomes generated from the previous steps will be reviewed by the research team to form the basis of the Delphi survey. Lay language summaries will be developed and presented together with the medical definitions to facilitate the participation of patients in this step of the COS development process. The Delphi survey will be generated using Qualtrics software and piloted with the study advisory group. Feedback will be incorporated into the survey structure, definitions and lay-language summaries, and overall usability of the survey.

Participants and Recruitment

Five key stakeholder groups with current or recent personal, clinical, research, or policy experience of RGCS will be targeted for the Delphi survey; patients (including both carriers and non-carriers identified through targeted or expanded screening), genetic health professions (genetic counsellors and clinical geneticists), non-genetics health professionals (obstetrician/gynaecologists, midwives, general practitioners), and policy-makers.

No recommendations currently exist regarding sample size for a Delphi survey, with wide variability in the number of participants across Delphi surveys for COS development.⁽⁴⁵⁾ Decisions regarding sample size are based on the area of practice and feasibility of recruiting sufficient representation from each

stakeholder group. As this is the first COS that we are aware of in the setting of genetic testing, we do not have a guide for how many stakeholders may be willing to participate in this process. We have adopted the approach of COS developers in the obstetric setting.⁽⁴⁶⁾ Equal representation of patient and professional perspectives is desirable; as such, we will aim to recruit at a minimum 50 patient participants and 50 participants from other professional stakeholder groups to the first round of the Delphi survey. In recognising natural rates of attrition in subsequent rounds, this number should allow sufficient representation through the three rounds of the Delphi process.

Our recruitment strategy incorporates diverse methods of identifying and recruiting participants to account for the range of key stakeholders that we are aiming to include. Expressions of interest to participate will be distributed through various channels: 1) patient participants from focus groups/interviews will be invited to participate in the Delphi process. We will also recruit through social media to reach our goal of 50 total patient participants; 2) researchers in the field will be purposively sampled based on first and last authors of papers included in our systematic reviews; 3) genetic and non-genetic health professionals will be purposively sampled based on professional networks of the research team and members lists of relevant professional organisations; 4) policy-makers will be purposively sampled from listed committee members on major practice recommendations related to RGCS. All participants will be encouraged to snowball information about the study to their networks to ensure a broad range of participants and experiences are captured. Participants who respond to expressions of interest through any of the above channels will be directed to an online survey on the Qualtrics platform to receive more information about the research and fill out basic demographic and screening questions to confirm their eligibility. Once eligibility is confirmed, they will receive the link to the Delphi survey, where they will be required to electronically indicate their consent before proceeding to the consensus questions. During the Delphi survey, recruitment may be targeted to groups that are under-represented to ensure balanced representation.

Data Collection

In round 1, participants will be asked to rate each outcome on a nine-point Likert scale based on the degree of importance as recommended by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group.⁽⁴⁷⁾ Rating 1 to 3 will be interpreted as "limited importance"; 4 to 6 as "important, but not critical"; and 7 to 9 as "critical importance." An option of "unsure" will also be available. The sequence of questions will be randomised to minimise ordering bias. For each outcome, a free text box will be available for participants to provide feedback or comments. New outcomes can be suggested by participants at the end of round 1 and will be reviewed by the research team to determine if they are unique and not overlapping, wholly or partially, with an existing outcome. Those that are deemed to be suitable will be carried over to round 2.

There is a lack of agreement on the definition of consensus to be used when deciding which outcomes to include in the second round of a Delphi survey, and a wide range of thresholds have been utilised in COS development. Per the guidance in the COMET handbook, using less stringent criteria in earlier rounds and

considering responses from individual stakeholder groups minimises the likelihood that outcomes that may have been rated higher in subsequent rounds after receiving feedback are not dropped too soon in the Delphi process.⁽²⁹⁾ We will adopt the definition utilised in a recent COS developed for surgery in oesophageal cancer, whereby criteria for inclusion in round 2 will be any outcomes that are rated 7 to 9 (critically important) by > 50% of participants and 1 to 3 (limited importance) by no more than 15% of any single stakeholder group.⁽⁴⁸⁾ Results will be presented graphically to participants at the time of the second round of the survey along with their rating of each outcome and any representative comments provided by participants that indicate their reasoning. This will allow participants to compare their ratings to other participants and consider whether they would change their rating in the next round. Participants will then be directed to re-rate the outcomes that have been carried over from the first round, with a free text box once again available for them to explain their rating or respond to comments from other participants from round 1. More stringent criteria for consensus will be applied to determine if there is a need for a third round of the Delphi survey, with outcomes that are rated 7 to 9 (critically important) by > 70% of participants and 1 to 3 (limited importance) by no more than 15% of any single stakeholder group.

A third round of the Delphi survey will be conducted if the number of outcomes remaining after the second round is too high to reasonably discuss at a consensus meeting. Criteria for inclusion may need to be adjusted at the time of the Delphi if sufficient reduction in outcomes is not being achieved, changes to which will be reported alongside the results of the Delphi survey. Following the final round of the Delphi survey (whether that is the second or third), outcomes that were rated 7 to 9 (critically important) by > 70% of participants and 1 to 3 (limited importance) by no more than 15% of participants will be taken forward to the consensus meeting for consideration of inclusion in the final COS.

Each round of the survey will be open for four weeks to provide participants with sufficient time to complete it. A maximum of 3 reminders will be sent to participants when two-weeks, one-week, and one-day are remaining to complete the survey. At the end of round 2 of the Delphi survey, participants will be asked to indicate if they are interested in participating in the final consensus meeting.

Data Analysis

We will summarise the overall distribution in ratings for outcomes across the rounds of the Delphi survey and the points at which outcomes were excluded from consideration. The mean and median will be calculated for each outcome. Data will be analysed in sub-groups to allow comparison between prioritisation of outcomes between health consumer participants and other stakeholder participants, and also between different subsets of the other stakeholder groups (for example, genetics health professionals versus non-genetic health professionals).

Step V: Consensus Meeting

We will host a virtual consensus meeting using Zoom. The purpose of this meeting will be to review the findings from steps 1–4 and discuss the outcomes for inclusion in an agreed-upon COS. Methods of measurement, implementation, and directions for further research will also be discussed if time allows.

Participants and Recruitment

From the participants that complete the Delphi survey, a maximum of 15 (2–3 from each participant group) will be purposively selected from those that have indicated an interest in participating in the final virtual consensus meeting, ensuring equal representation across stakeholder groups. We may consider contacting Delphi participants that did not express interest at the conclusion of the Delphi, where there is insufficient representation from specific groups. Purposively selected participants will be sent an email per the contact details they have provided at an earlier stage to invite them to participate. Verbal consent will be obtained at the beginning of the session. The number of participants selected to participate in the final virtual consensus meeting is a pragmatic decision based on balancing sufficient representation to incorporate all perspectives with a manageable number of participants that gives everyone a chance to contribute.

Data Collection and Analysis

The virtual consensus meeting will be convened over 1–3 hours, depending on the number of outcomes that need to be discussed based on the results of the Delphi survey. Consideration will be given to holding two consensus meetings based on time-zones to facilitate international participation in this final stage of the COS development. The virtual consensus meeting will consist of a voting system as well as open sections of discussion. An overview of the CODECS projects results to date will be presented at the start of the meeting, followed by proposal of each outcome that satisfied the inclusion criteria set out in the Delphi survey. Participants will have the opportunity to discuss each outcome before lodging a vote for its inclusion or exclusion, with outcomes that achieve > 75% consensus being included. The core outcome set literature varies in its approach to consensus thresholds for the consensus meeting, with the majority setting a minimum of 70%. A slightly higher threshold of 75% is favoured by some authors as it allows for increased stringency in the final step of the consensus process and we have elected this approach.^(49, 50) Results will be presented after the voting is complete, outlining which outcomes reach consensus for inclusion in the core outcome set. Those that reach consensus for exclusion or that have no consensus will be reviewed, with panel members having an opportunity to provide an opinion if they see a fundamental reason why they disagree with the exclusion of these outcomes. Should the number of outcomes that reach consensus for inclusion be unwieldy, we may consider a tiered approach for reporting of the core outcome set, as has been done by previous COS developers.⁽⁵¹⁾ Should time allow, we will finish the meeting with a discussion focusing on the definition of outcomes, appropriate outcome measures, and reporting and implementation of the COS; however these will need to be addressed in more detail in subsequent research. The discussion sections of the meeting will be transcribed verbatim and analysed using thematic analysis as described for the focus groups/interviews above.

Dissemination and Implementation

This COS will be reported according to the COS-STAR reporting guidelines.⁽⁵²⁾ Efforts for dissemination and implementation will include publishing the COS in an open-access journal, presenting the findings at

conferences of relevant professional organisations, sharing with clinical trial registries, and encouraging stakeholder participants to circulate the final COS to their professional networks internationally.

Discussion

RGCS is one of the most widely available genetic tests internationally and has the potential to provide families with information about their reproductive risks and allow them to make informed decisions in family planning and pregnancy. As with many other types of genetic testing, it is not clear what outcomes are best to assess when considering the impact of RGCS, which has led to marked heterogeneity within the literature and hindered policy-makers in their attempts to utilise high-quality research evidence to support its implementation into routine clinical practice. Contingencies exist from a policy perspective in such cases, allowing expert consensus to be used to make practice recommendations; however this does not address the underlying issues.

This study will provide researchers with guidance on which outcomes to include, at a minimum, in any study evaluating RGCS. As has been seen in other contexts, the application of a COS ensures that a minimum set of outcomes are routinely reported in all studies on a particular topic, allowing reliable comparisons across studies to be achieved. It also facilitates the combining of data where appropriate for use in meta-analyses to quantify outcomes. Reporting bias is minimised by requiring that the COS is always reported as a minimum, meaning that even non-significant changes will be represented in the literature. Differences that do not reach significance in studies with small sample sizes may reach significance when combined in a meta-analysis. The COS will ensure that outcomes that are relevant to patients are incorporated in future studies. The development of this COS will have implications beyond RGCS, to other forms of genetic testing, and assist in ongoing efforts to define outcomes of genetic services and genetic counselling.

Abbreviations

CODECS – Core Outcome Development for Carrier Screening

COMET – Core Outcome Measures in Effectiveness Trials

COS – Core Outcome Set(s)

RGCS – Reproductive Genetic Carrier Screening

Declarations

TRIAL STATUS

Protocol version 1.0, October 2020

The development of this core outcome set is active and ongoing, with systematic reviews of the quantitative and qualitative literature currently underway.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study and relevant consent documents have received approval from the University of Technology Sydney Medical Research Ethics Board [UTS HREC REF NO. ETH20-5179]. Consent will be obtained verbally for patient focus groups/interviews and the virtual consensus meeting, and electronically for the Delphi survey.

CONSENT FOR PUBLICATION

Not applicable

AVAILABILITY OF DATA AND MATERIALS

Not applicable

COMPETING INTERESTS

The authors declare that they have no competing interest

FUNDING

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.

AUTHOR'S CONTRIBUTION

All authors have made substantial contribution to the conception and writing of this protocol, and have seen and approved the final draft. ER is the principal investigator and is undertaking this project to meet the requirements for award of a Doctor of Philosophy (PhD) in genetic counselling at the University of Technology Sydney. CJ is the primary research supervisor, and AM and TNJ are co-supervisors, for ER. CJ, AM and TNJ provided expert advice and guidance on the project design, research methodologies, critical revision of drafts of the manuscript and overall doctoral supervision. KM provided detailed advice on COS development methodology and experience with applying this methodology and critically reviewed drafts of the manuscript. ER led the writing of the protocol.

ACKNOWLEDGEMENT

This project is supported by an Australian Research Training Program Fee Waiver Scholarship and Research Excellence Scholarship, with thanks to the Graduate School of Health, University of Technology Sydney.

References

1. Resta RG. What Have We Been Trying To Do and Have We Been Any Good At It? A History of Measuring the Success of Genetic Counseling. *Eur J Med Genet.* 2019;62(5):300-7.
2. Delatycki MB, Alkuraya F, Archibald A, Castellani C, Cornel M, Grody WW, et al. International Perspectives on the Implementation of Reproductive Carrier Screening. *Prenat Diagn.* 2020;40(3):301-10.
3. Sankaranarayanan K. Ionizing Radiation and Genetic Risks IX. Estimates of the Frequencies of Mendelian Diseases and Spontaneous Mutation Rates in Human Populations: A 1998 Perspective. *Mutat Res.* 1998;411(2):129-78.
4. Carter CO. Contribution of Gene Mutations to Genetic Disease in Humans. Bora KC, Douglas GR, Nestmann ER, editors. Amsterdam: Elsevier; 1982. 1-8 p.
5. Henneman L, Borry P, Chokoshvili D, Cornel MC, van El CG, Forzano F, et al. Responsible implementation of expanded carrier screening. *Eur J Hum Genet.* 2016;24(6):e1-e12.
6. Ropers HH. On The Future of Genetic Risk Assessment. *J Community Genet.* 2012;3(3):229-36.
7. Cao A, Saba L, Galanello R, Rosatelli MC. Molecular Diagnosis and Carrier Screening for Beta Thalassemia. *JAMA.* 1997;278(15):1273-7.
8. Davies SC, Cronin E, Gill M, Greengross P, Hickman M, Normand C. Screening for Sickle Cell Disease and Thalassaemia: A Systematic Review with Supplementary Research. *Health Technol Assess.* 2000;4(3):i-v, 1-99.
9. Lew RM, Burnett L, Proos AL, Barlow-Stewart K, Delatycki MB, Bankier A, et al. Ashkenazi Jewish Population Screening for Tay-Sachs Disease: The International and Australian Experience. *J Paediatr Child Health.* 2015;51(3):271-9.
10. Watson EK, Mayall E, Chapple J, Dalziel M, Harrington K, Williams C, et al. Screening for Carriers of Cystic Fibrosis through Primary Health Care Services. *BMJ.* 1991;303(6801):504-7.
11. Australian Genomics. Mackenzie's Mission: The Australian Reproductive Genetic Carrier Screening Project [Available from: <https://www.australiangenomics.org.au/our-research/disease-flagships/mackenzies-mission/>.]
12. Wilson RD, De Bie I, Armour CM, Brown RN, Campagnolo C, Carroll JC, 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):742-62 e3.

13. Committee on Genetics. Committee Opinion No. 690: Carrier Screening in the Age of Genomic Medicine. *Obstet Gynecol.* 2017;129(3):e35-e40.
14. Committee on Genetics. Committee Opinion No. 691: Carrier Screening for Genetic Conditions. *Obstet Gynecol.* 2017;129(3):e41-e55.
15. Genomics Advisory Working Group & Women's Health Committee. *Genetic Carrier Screening.* RANZCOG; 2019.
16. An Evidence Framework for Genetic Testing. Committee on the Evidence Base for Genetic Testing, editor. Washington DC: The National Academies Press; 2017.
17. Sun F, Bruening W, Erinoff E, Schoelles KM. Addressing Challenges in Genetic Test Evaluation: Evaluation Frameworks and Assessment of Analytic Validity. AHRQ Methods for Effective Health Care. Rockville MD 2011.
18. 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):902-33.
19. Madlensky L, Trepanier AM, Cragun D, Lerner B, Shannon KM, Zierhut H. A Rapid Systematic Review of Outcomes Studies in Genetic Counseling. *J Genet Couns.* 2017;26(3):361-78.
20. 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.
21. Pitini E, De Vito C, Marzuillo C, D'Andrea E, Rosso A, Federici A, et al. How is Genetic Testing Evaluated? A Systematic Review of the Literature. *Eur J Hum Genet.* 2018;26(5):605-15.
22. Horne J, Madill J, O'Connor C, Shelley J, Gilliland J. A Systematic Review of Genetic Testing and Lifestyle Behaviour Change: Are We Using High-Quality Genetic Interventions and Considering Behaviour Change Theory? *Lifestyle Genom.* 2018;11(1):49-63.
23. Kessels SJM, Carter D, Ellery B, Newton S, Merlin TL. Prenatal Genetic Testing for Cystic Fibrosis: A Systematic Review of Clinical Effectiveness and an Ethics Review. *Genet Med.* 2020;22(2):258-67.
24. Hammer GP, du Prel JB, Blettner M. Avoiding Bias in Observational Studies: Part 8 in a Series of Articles on Evaluation of Scientific Publications. *Dtsch Arztebl Int.* 2009;106(41):664-8.
25. 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):1117-29.
26. 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):442-50.
27. 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.

28. Ioannou L, McLaren BJ, Massie J, Lewis S, Metcalfe SA, Forrest L, et al. Population-Based Carrier Screening for Cystic Fibrosis: A Systematic Review of 23 Years of Research. *Genet Med.* 2014;16(3):207-16.
29. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: Version 1.0. *Trials.* 2017;18(Suppl 3):280.
30. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core Outcome Set-STAndardised Protocol Items: the COS-STAP Statement. *Trials.* 2019;20(1):116.
31. Melnyk BM, Fineout-Overholt ES. Evidence-Based Practice in Nursing & Healthcare: A Guide to Best Practice. Philadelphia: Lippincott Williams & Wilkins; 2005.
32. 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.
33. Dodd S, Clarke M, Becker L, Maverganes 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.
34. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. London: Institute for Health Research; 2006.
35. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med.* 2017;14(11):e1002447.
36. Gorst SL, Young B, Williamson PR, Wilding JPH, Harman NL. Incorporating Patients' Perspectives into the Initial Stages of Core Outcome Set Development: A Rapid Review of Qualitative Studies of Type 2 Diabetes. *BMJ Open Diabetes Res Care.* 2019;7(1):e000615.
37. Barbour R. Doing Focus Groups. 2nd ed. London: Sage Publications; 2018.
38. Qualtrics. Provo, Utah, USA 2005. Available at: <https://www.qualtrics.com>.
39. Corruption Perceptions Index: Transparency International. www.transparency.org/cpi (2019). Accessed on 10 June 2020.
40. Zoom Video Communications Inc. Security Guide. Zoom Video Communications Inc. Retrieved from <https://d24cgw3uvb9a9h.cloudfront.net/static/81625/doc/Zoom-Security-White-Paper.pdf2016>.
41. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants. *Int J Qual Methods.* 2019;18:1-8.
42. Corbin JM, Strauss AL. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park, California: Sage Publications; 1990.
43. Braun V, Clarke V. Thematic Analysis. APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological. APA handbooks in psychology®. Washington, DC, US: American Psychological Association; 2012. p. 57-71.
44. Sinha IP, Smyth RL, Williamson PR. Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials: Recommendations for the Future Based on a Systematic Review of

Existing Studies. PLoS Med. 2011;8(1):e1000393.

45. Akins RB, Tolson H, Cole BR. Stability of Response Characteristics of a Delphi Panel: Application of Bootstrap Data Expansion. BMC Med Res Methodol. 2005;5:37.
46. D'Souza R, Hall C, Sermer M, Siu S, Silversides C. Development of a Core Outcome Set for Studies on Cardiac Disease in Pregnancy (COSCarP): a study protocol. Trials. 2020;21(1):300.
47. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE Guidelines: 2. Framing the Question and Deciding on Important Outcomes. J Clin Epidemiol. 2011;64(4):395-400.
48. Blazeby JM, Macefield R, Blencowe NS, Jacobs M, McNair AG, Sprangers M, et al. Core Information Set for Oesophageal Cancer Surgery. Br J Surg. 2015;102(8):936-43.
49. Bennett WL, Robinson KA, Saldanha IJ, Wilson LM, Nicholson WK. High Priority Research Needs for Gestational Diabetes Mellitus. J Womens Health (Larchmt). 2012;21(9):925-32.
50. Tyler N, Wright N, Grundy A, Waring J. Developing a Core Outcome Set for Interventions to Improve Discharge from Mental Health Inpatient Services: A Survey, Delphi and Consensus Meeting with Key Stakeholder Groups. BMJ Open. 2020;10(5):e034215.
51. Manera KE, Tong A, Craig JC, Brown EA, Brunier G, Dong J, et al. Standardized Outcomes in Nephrology-Peritoneal Dialysis (SONG-PD): Study Protocol for Establishing a Core Outcome Set in PD. Perit Dial Int. 2017;37(6):639-47.
52. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, et al. Core Outcome Set-STAndards for Reporting: The COS-STAR Statement. PLoS Med. 2016;13(10):e1002148.

Figures

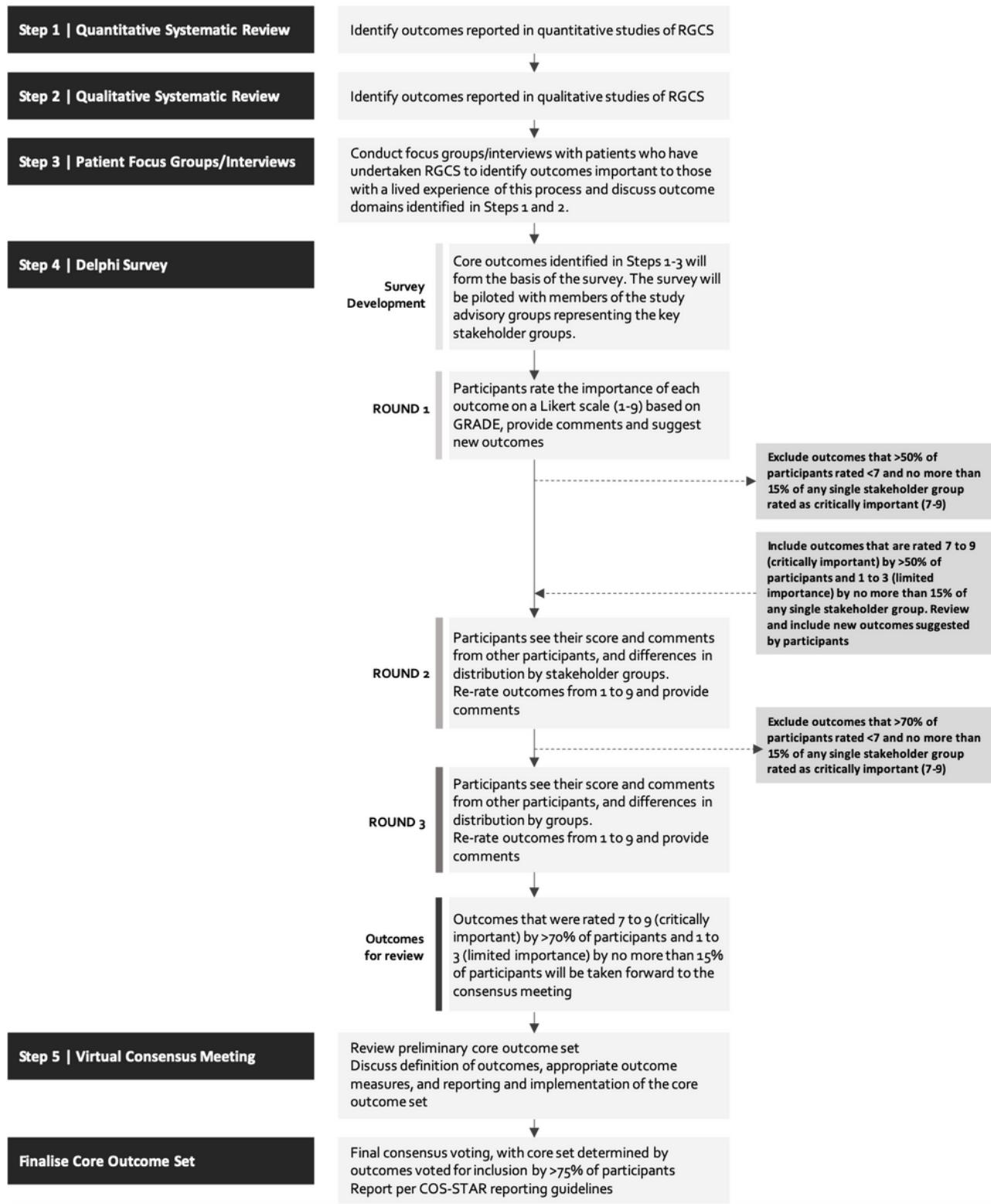


Figure 1

Study Overview Diagram