ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/117256/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Grant, Peter E., Pampaka, Maria, Payne, Katherine, Clarke, Angus and McAllister, Marion 2019. Developing a short-form of the Genetic Counselling Outcome Scale: The Genomics Outcome Scale. European Journal of Medical Genetics 62 (5) , pp. 324-334. 10.1016/j.ejmg.2018.11.015

Publishers page: https://doi.org/10.1016/j.ejmg.2018.11.015

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Title: Developing a short-form of the Genetic Counselling Outcome Scale: The Genomics Outcome Scale

Authors and affiliation: Peter E. Grant [a], Maria Pampaka [b], Katherine Payne [c], Angus Clarke [d], Marion McAllister [a]

[a] Centre for Medical Education, School of Medicine, Cardiff University, Cardiff, UK

[b] Departments of Social Statistics (School of Social Science) and Education (School of Environment, Education and Development), The University of Manchester, Manchester, UK

[c] Division of Population Health, Health Services Research and Primary Care, The University of Manchester, Manchester, UK

[d] Division of Cancer & Genetics, School of Medicine, Cardiff University, Cardiff, UK

Corresponding Author: Dr Marion McAllister, Centre for Medical Education, School of Medicine, Cardiff University, Cardiff, CF14 4XN. Email: mcallistermf@cardiff.ac.uk; T: +44 (0) 2922 510 811

Abstract

The Genetic Counselling Outcome Scale (GCOS-24) is a 24-item patient reported outcome measure for use in evaluations of genetic counselling and testing services. The aim of this study was to develop a short form of GCOS-24. The study comprised three phases. Phase I: Cognitive interviews were used to explore interpretability of GCOS-24 items and which GCOS-24 items were most valued by the target population. Phase II: The Graded Response Model was used to analyse an existing set of GCOS-24 responses (n= 395) to examine item discrimination. Phase III: Item Selection. Three principles guided the approach to item selection (i) Items with poor discriminative properties were not selected; (ii) To avoid redundancy, items capturing a similar outcome were not selected together; item information curves and cognitive interview findings were used to establish superior items. (iii) Rasch analysis was then used to determine the optimal scale. In Phase I, ten cognitive interviews were conducted with individuals affected by or at risk for a genetic condition, recruited from patient support groups. Analysis of interview transcripts identified twelve GCOS-24 items which were highly valued by participants. In Phase II, Graded Response Model item characteristic curves and item information curves were produced. In Phase III, findings from Phases I and II were used to select ten highly-valued items that perform well. Finally, items were iteratively removed and permutated to establish optimal fit statistics under the Rasch model. A six-item questionnaire with a five-point Likert Scale was produced (The Genomics Outcome Scale (GOS)). Correlation between GCOS-24 and GOS scores is high (r=.838 at 99% confidence), suggesting that GOS maintains the ability of GCOS-24 to capture empowerment, whilst providing a less burdensome scale for respondents. This study represents the first step in developing a preference-based measure which could be used in the evaluation of technologies and services used in genomic medicine.

Keywords:

GCOS-24; GOS; Genetic Counselling; Patient reported outcome measure; Item response theory

Introduction

Genetic counselling and associated genomic testing services (hereafter shortened to 'clinical genetics services' (CGS)) have the potential to offer a number of benefits to individuals and families affected by conditions that may have a genetic aetiology. Recent studies have provided evidence that patients are seeking information and a supportive relationship, and that the benefits of genetic counselling include relief of uncertainty and feelings of vulnerability, as well as adaptation to the genetic condition in the family (Bernhardt *et al.,* 2000; MacLeod *et al.,* 2002; McAllister *et al.,* 2008; Payne *et al.,* 2007; Skirton, 2001). Robust and validated measures of these benefits are needed to provide evidence to service commissioners about the outcomes of investing in existing CGS or future service developments.

Evaluations of CGS have examined outcome variables such as knowledge, information recall, reproductive intentions, decisions made, anxiety or distress, patient satisfaction, perceived risk, perceived personal control, health behaviours, and decisional conflict (Payne *et al.,* 2008; Madlensky et al., 2017). There is some evidence that genetic counselling can result in increased knowledge, perceived personal control, positive health behaviours, and accuracy of perceived risk amongst patients, and decreased anxiety, worry, and decisional conflict. There is also evidence that patients are typically very satisfied with genetic counselling (Madlensky et al., 2017).

Measures of process such as waiting times and numbers of patients seen have also been used, as well as the performance characteristics of genetic tests (e.g. sensitivity, specificity and predictive values) (Clarke *et al.,* 1996; Payne *et al.,* 2008). Little attention has been paid to exploring outcomes relevant to the population of individuals who use CGS (McAllister *et al.*, 2008), and there have been calls for research to identify outcomes that are most important to patients (Madlensky et al., 2017).

Moreover, many of the measures which have been used to evaluate CGS have not undergone rigorous psychometric validation, with many having been assessed for internal consistency only, and few measures assessed for important characteristics such as reliability and responsiveness to change (Payne *et al.,* 2008; McAllister & Dearing 2015).

In 2011, the Genetic Counselling Outcome Scale (GCOS-24) (Figure 1) was developed to provide an English language patient-reported outcome measure (PROM), specific to clinical genetics services (CGS) (McAllister, 2011b). GCOS-24 items are grounded in extensive qualitative research with CGS patients and providers, capturing an emergent theoretical construct labelled 'empowerment', comprising five sub-dimensions that summarise the outcomes valued by those stakeholders: cognitive, decisional and behavioural control, emotional regulation and hope (McAllister *et al.,* 2008; McAllister, 2011a). 'Empowerment' was chosen as the construct name because it appeared to capture the 'meaning' across the five sub-dimensions. Despite 'patient empowerment' having gained considerable importance in healthcare policy globally, there is no universally accepted definition of the term. Whilst most definitions are consistent with the approaches and principles of patient-centred care, patient empowerment has been conceptualised in many different ways, including as an underpinning ethos (e.g. that patients have rights relating to autonomy, self-determination and power within their healthcare relationships), as empowering interventions (e.g. shared decision-making) and as an indicator (e.g. a patient 'state' ranging from low to high levels of the variable 'empowerment') (Bravo *et al*, 2015). A range of patient empowerment constructs have been operationalised in published measures of empowerment, including constructs that reflect patient states, patient experiences and capacities, patient actions and behaviours, patient self-determination and patient skills development (Barr *et al,* 2015). The 'empowerment' construct operationalised in the GCOS-24 is most consistent with a patient 'state'.

GCOS-24 has been demonstrated to be valid, reliable and responsive, with no floor or ceiling effects observed (McAllister *et al.,* 2011b), and has been used for service evaluation (Inglis *et al.,* 2015; McAllister *et al.,* 2016) and quality improvement (Costal-Tirado *et al.,* 2017) in genetic counselling services. It has also received international attention, having been translated into Danish (Diness *et al.,* 2017) and Spanish (Munoz-Caballo *et al.,* 2018).

<<insert Figure 1 here>>

GCOS-24 has 24 items each with 7 response options (Figure 1). GCOS-24 generates an overall 'empowerment' score, however it is not clear what interpretation can be attached to differences in score and between items. Further work is needed to attach 'preference weights' to the measure, reflecting the value or priority which is placed on each item by the target population (Sinnott *et al.,* 2007). This will make it clear what interpretation can be attached to changes in score. In its current form, however, GCOS-24 produces a substantial number of possible response permutations (1.92×10^{20}). A shorter version of the scale would make it possible to design a study to elicit such preference weights, thereby facilitating future use of the shorter scale in economic evaluations of genetic and genomic testing with and without genetic counselling.

Additionally, the wording of GCOS-24 items 1, 14 and 23, which refer specifically to CGS, means that the measure is unsuitable for use outside of CGS. Genetic testing is increasingly being performed outside the traditional models of service provision within CGS and is now moving into other specialities. This process is referred to as 'mainstreaming genetic testing' and is occurring, for example, in the context of cancer predisposition genes (Rahman, 2014), paediatrics (Valente *et al.,* 2008), and neurogenetic testing (Lo *et al.,* 2014). It is therefore becoming ever more important to have a valid and reliable PROM which can be used to evaluate genetic and genomic counselling and testing both within and outside of CGS. A

further added benefit of a shorter measure would be to reduce completion time, which may also facilitate integration into clinical care.

In summary, a shorter version of GCOS-24 would be useful because (1) GCOS-24 is a thoroughly validated PROM for genetic counselling and testing services, since most other available CGS-specific PROMs have not been assessed for both reliability and responsiveness to change (2) genetic testing is increasingly being done outside the context of clinical genetics services, with no thoroughly validated PROM available (because GCOS-24 has items that refer specifically to clinical genetics services) and (3) most available CGS-specific measures have been developed for use in cancer genetics only, and are not suitable for general genetics services (4) there is no available PROM with attached preference weights that could be developed for use in economic evaluations of genetic counselling and testing services and (5) a shorter scale would reduce respondent burden and facilitate integration into clinical care.

Over recent years, the growing emphasis on patient-centred care has accelerated the demand for high-quality PROM data, leading to a rise in popularity for modern psychometric methods such as item response theory (IRT) (Alonso *et al.*, 2013; Nguyen, 2014). IRT methods enable the creation of item banks for measuring specified health status domains, which in turn allows for item comparison and computerised adaptive testing (CAT) tools for tailored assessments without loss of scale precision or content validity (Bjorner *et al.,* 2003; Cella *et al.,* 2007; Haley *et al.,* 2004; Harniss *et al.,* 2007). The recognised value of IRT methods is demonstrated by the Patient-Reported Outcome Measurement Information System (PROMIS) initiative in the US, which aims to catalogue validated PROMs and build accessible item banks for measuring key health concepts applicable to a range of conditions.

This study aims to take the first step towards establishing a PROM which would be appropriate for routine use in audit and clinical evaluations of genetic services. The specific

aim is to develop a short form of the GCOS-24 (using both qualitative and IRT methods), suitable for use both within and outside the context of CGS and in research, which still appropriately captures the empowerment construct.

Participant data

For Phase I, participants were identified and recruited by Genetic Alliance UK (GAUK: https://www.geneticalliance.org.uk/), a national charity comprising over 180 support groups for genetic conditions. Phase II and Phase III used an existing dataset, comprising a set of responses to GCOS-24 (n=395), collected in 2010 for the original psychometric validation (McAllister, 2011b). Specific details (e.g. gender, ethnicity, condition type, reason for referral) can be found in McAllister *et al.* (2011b).

Methods

There were three phases to this study. Phase I used qualitative cognitive interviews (Ericsson & Simon, 1980) to explore the relevance of the existing GCOS-24 items from the perspective of the target population. Phase II involved analysis of an existing data set of GCOS-24 responses (n=395) using Samejima's Graded Response Model (GRM) (Samejima, 1969) to examine item discrimination. Phase III combined the results from Phases I & II to inform item selection, and employed the Rasch model to explore potential item combinations and functional problems with the seven-point Likert Scale.

Ethics approval for the recruitment of human participants was granted by Cardiff University School of Medicine, 12th May 2017. Ethics approval for the secondary use of GCOS-24 responses was granted by the National Research Ethics Service (NRES) Committee North West.

Phase I: Cognitive Interviews

Phase I used open-ended, semi structured think-aloud cognitive interviews (Ericsson & Simon, 1980), conducted on an individual basis. Potential participants were contacted by

GAUK using the following inclusion criteria: individuals who (i) are at risk of, or affected by, a genetic condition; (ii) are over 18 years old; (iii) have expressed an interest in participating in research. A Participant Information Sheet was sent to participants, and was provided again prior to interview. Informed consent was confirmed immediately prior to interview through a written, signed consent form. For interviews conducted by telephone or Skype, this process was done by post.

The interview guide was adapted from the cognitive interview guide described by Irwin *et al.* *(*2009) and was designed to explore participants' perceptions of the meaning and interpretability of GCOS-24 items, and which items were considered most important or relevant. Interviews were audio-recorded and then transcribed in full. Qualitative data was analysed using framework analysis (Ritchie & Spencer, 1994), a method which uses a pre-determined theoretical structure to assist in the process of identifying, analysing, and reporting patterns (themes) within qualitative data. McAllister's empowerment construct (McAllister *et al.*, 2011a) was chosen as an appropriate framework, since GCOS-24 was specifically developed to capture that construct.

Phase II: Quantitative Analysis

An essential property for any measurement scale is the ability to discriminate between individuals. A scale intending to measure empowerment would be of little use if it produced the same results regardless of whether a person was empowered or not. One of the aims of reducing GCOS-24 was to retain those items which can discriminate between degrees of empowerment. Samejima's GRM (Samejima, 1969) was used to analyse GCOS-24 responses (n=395), having separated the variables into the sub-dimensions of empowerment that each was designed to capture (e.g. cognitive control, behavioural control).

GRM item characteristic curves can be used to assess the likelihood of respondents selecting a certain response option at various degrees of the latent trait. An item is better at

discriminating between individuals when the curves are peaked and dispersed across all levels of the latent trait. GRM item information curves show how well and precisely each item measures the latent trait across various levels of that trait. Certain items may display a skew towards lower levels of the trait, while others may skew towards higher levels. Comparing item information curves allows a comparison to be made between items on how well the latent trait is represented by the item. The R package ltm (Rizopoulos, 2006) was used to perform the GRM.

Phase III: Item Selection

Three principles guided the approach to item selection. (i) Items with poor discriminative properties were not selected; (ii) To avoid redundancy, items capturing a similar outcome were not selected together; item information curves and cognitive interview findings were used to establish superior items. (iii) Rasch analysis was then used to determine the optimal scale.

The Rasch Rating Scale model (Andrich, 1978) was used to examine model fit and the reliability of potential item combinations, as well as functional problems with the 7-point Likert scale. The Winsteps Rasch Measurement software (Linacre, 2018) was used for Rasch Analysis.

Results

Phase I: Cognitive Interviews

Of the 35 individuals contacted, ten (28.6%) replied and were successfully recruited to participate in think-aloud cognitive interviews. Participant characteristics are summarised in Table 1. For anonymity, participants are identified with the letter P followed by a number. Evidence confirming a diagnosis of a genetic condition was not sought, but all participants believed that their condition was genetic.

<<insert Table 1 here>>

Table 2 summarises the items which were most valued by participants. For simplicity, items have been grouped according to which sub-dimension of McAllister's empowerment framework each was designed to capture: (i) cognitive control; (ii) decisional control; (iii) behavioural control; (iv) emotional regulation; (v) hope (McAllister, 2011a). The qualitative framework analysis findings are presented below. Empowerment was sufficient to integrate all themes which arose from the cognitive interviews, and item numbering will be referred to in GCOS-24 (Figure 1).

<<insert Table 2 here>>

Cognitive Control

Part of feeling empowered in relation to a genetic condition in the family is having a belief that you have sufficient knowledge and understanding about the condition (cognitive control) (McAllister *et al.*, 2008). This could be knowledge about how the condition is inherited, what causes it, what the signs and symptoms are, and what the implications are for the rest of the family, both at present and in the future. All ten participants spoke of their desire to learn more about their condition, both at the time of diagnosis and as an ongoing pursuit, and of the benefits that this knowledge could have on their lives. On an item level, six GCOS-24 items had been designed to capture cognitive control: items 1, 3, 12, 14, 18, and 23 (Table 2). Of these, items 18 ('I don't know who else in my family might be at risk for this condition') and 12 ('I don't know if this condition could affect my other relatives (brothers, sisters, aunts, uncles, cousins)') appeared to be valued most highly by participants. Knowing how the condition might affect one's relatives was judged to be very useful information. This participant spoke of item 18:

> "That's a really good question because if you, if this was day one, so you ask someone before their first session, they're probably going to answer that quite high. If you ask them after 5 sessions, then actually the answer could be completely the

> opposite, so I think that's a valuable question to ask, because you can show the progress they've made and what they've learnt from the session. I think that's a really good question. And before I had the genetic counselling I would have answered I don't know, and now I can answer I do know, because I had the service and got the information that I needed." (P4)

Another participant, when considering item 12, stated:

> "I think it is an important piece of knowledge to have. If I didn't know that information I would be worried, and I could see how people would get worried about that type of thing" (P5).

Items 12 and 18 emerged as strong candidates for retention (Table 2).

Decisional control

Decisional control within the empowerment framework is not restricted only to decisions made about healthcare. It can include any major or minor decision which is influenced by having a genetic condition within the family (McAllister *et al.,* 2008). This might involve decisions on marriage, whether or when to have children, or on seemingly unrelated decisions such as buying a car or whether to take on a mortgage. Decisional control was discussed by participants, however the corresponding GCOS-24 items (10, 13 and 24) were problematic. More specifically, items 13 ('In relation to the condition in my family, nothing I decide will change the future for my children / any children I might have) and 24 (I can make decisions about the condition that may change my child(ren)'s future / the future of any child(ren) I may have) suffered because they were not seen as relevant by those participants who did not have children. Item 10 ('I don't know what could be gained from each of the options available to me') was unclear to some:

... [Interviewer: "Are any items difficult to understand?"] "I suppose number 10: I don't know what could be gained from each of the options available to me. That's a little bit, what options are we talking about?" (P1)

... [Interviewer: "What does item 10 mean to you?"] "Umm. Well the first thing that comes to mind after reading that question is, I don't know what options it means. Umm. As far as I'm concerned I had genetic counselling, and now I've just got to see my consultant, take my medication... and that's it. I don't know any options that are available to me at all. So, it doesn't mean a lot to me."

Item 24 emerged as the strongest candidate for retention, as demonstrated by the following:

... "I like the one about decision making [item 24], 'cause it's the main thing; you get the information you want so you can make decisions. Sometimes we have to make decisions without all the information, so to actually have a service available to inform all your decision making, that's gold you know. Wish we had it for more things. So yeah, if you're only going to keep one of them, I think the decision making one is the best one there." (P4)

... "To me reading that [item 24], it's what I do every day; I make decisions for her. If I feel she [the daughter] can't do something in the normal way, then I find other routes so it enables her to do everything anybody else is doing. To me that is making a decision. So, you're always decision making, always. You can never stop decision making for your child."

Item 24 emerged as the strongest candidate for retention (Table 2).

Behavioural control

Behavioural control is perhaps the most diverse dimension of empowerment, representing the perception of an individual that they are able to take action to improve their situation. This includes making effective use of the health and social care systems which are available,

managing the condition day to day, or communicating about genetic risks with relatives (McAllister *et al.,* 2008). All participants spoke at length about the importance of outcomes corresponding to behavioural control. Topics included their experiences with the NHS, the vital importance of both medical and non-medical services following diagnosis with a genetic condition, and how important it is to be able to communicate with others about the condition, whether that be with family, work colleagues, or with a school on behalf of their child. This participant, for example, spoke of her experience with local support groups:

> ... "The [support] groups are a massive help. I forced my sister to join. I do think that having a network of people going through the same thing, it doesn't matter what your situation is, whether it's, you know, cancer, depression, or anything that's happened, if you've got a group of people going through the same thing you are, it's ultimately just support and it will always help." (P5)

On an item level, almost all items designed to capture behavioural control were valued by participants. Especially popular were items 2 ('I can explain what the condition means to people in my family who may need to know') and 16 ('I can explain what the condition means to people outside my family who may need to know'), with all ten expressing the benefits of being able to talk about the condition. This participant contextualised the outcome within social situations:

> ... "Most people I think are naturally inquisitive. If they can see or know that an individual has a condition, disability, call it what you want, and if you're able to talk comfortably about it, and other people around you can talk comfortably about it, and answer what may sometimes seem ignorant or silly questions, and you're happy to take those questions, then that's in the best interest of everybody. Whether it is family, friends, or work colleagues, whatever, it makes life easier for everybody." (P1)

A problem was observed with respect to items 5 ('I don't know where to go to get the medical help I/my family needs') and 15 ('I know how to get the non-medical help I / my family needs (e.g. educational, financial, social support)'). A common perception amongst even the most well-informed participants was of not being sure whether there might be more services out there, and as such they were reluctant to agree or disagree with these items.

Item 7 produced contrasting interpretations over the meaning of 'control':

> ... "I've got control over how people react to it, over how much people need to know, or how they act around my child. Obviously I can't control how ill he'll get. [...] I think that's the bit of control I've got, he will be very confident. He's going to be brought up very confident, and very, you know, not embarrassed about anything." (P3)
>
> ... "See I don't like, I would take out control and I would put manage. I can't, we have no control. We can only do things to lessen the impact, or try to lessen the impact. [...]So I don't think you can control, ever control it. I think you can try and manage the condition. But control, no." (P7)
>
> ... "You've got no control over it [the condition]. Get over yourself." (P2)
>
> ... "I don't know about control. [...] It's more empowerment and advocacy of ownership, those are the things. Those are the terms I would be more likely to use over control. I don't use control, or very rarely." (P1)

Aside from item 7, all items in behavioural control emerged as candidates for retention (Table 2).

Emotional regulation

Emotional regulation in the empowerment framework refers to the ability to manage the emotional aspects of a genetic condition, both individually and within the family (McAllister *et al.,* 2011a). The diagnosis of a genetic condition can raise significant emotional challenges,

and it is important that these emotions are addressed by any provider of genetic counselling or testing services. In this study, outcomes such as guilt, anxiety, blame, helplessness, powerlessness, shame and sadness were brought up by participants as well as feeling "lost" (P6), "damaged"(P4), or "broken" (P4).

Items 4, 11 and 21 were designed to capture emotional regulation. Of these, item 4 ('When I think about the condition in my family I get upset') emerged as the strongest candidate. Item 21 ('I feel guilty because I (might have) passed this condition on to my children') was not seen as relevant to those without children. Item 11 ('Having this condition in my family makes me feel anxious') was highly valued by some, but overall the findings suggest that anxiety levels do not necessarily reduce in the long term, and instead fluctuate depending on the situation. This would not be a desirable item to have in an instrument which measures patient benefits from CGS.

> ... "Having the condition in my family makes me feel anxious... I don't think that will ever fully go away. I'm anxious for him [the son] when he starts a family. What if he gets someone pregnant and he hasn't stopped the hydroxycarbamide? It does happen. I'm anxious for my future grandchildren. [...] Knowing more about it [the condition] has made me less anxious in one respect, but more anxious in other respects. I am very anxious because, even though I know a lot, I don't feel I know everything." (P5)
>
> ... "[Interviewer: Does the anxiety improve over time?]" "The anxiety? ... no. Because the minute she has another episode you sort of take a deep breath and you hold it until she comes out of this episode."

Item 4 emerged as the strongest candidate for retention (Table 2).

Hope

Positivity, or a positive mind-set, was the chief manifestation of hope in this study. Item 19 ('I am hopeful that my children can look forward to a rewarding family life') is only

applicable to those with children, and even within that demographic appeared to be of questionable relevance:

> ... "I don't see why you're asking that as part of genetic counselling [laughs]. I just think that everybody, who would say no to that? [Continues laughing] 'I'm going to have kids and I hope that they have a c**p life!' So I don't, if you're trying to evaluate the results of these questions, I can't see how that would help the service at all; because everybody always wants the best for their children. I can't imagine anybody not answering positively to that." (P4)

Item 6 ('I can see that good things have come from having this condition in my family') was criticised for being irrelevant:

> ... "I can see that good things have come from having this condition? No, no I don't see that. Because we're a close family anyway and whether this condition was there or not there we would still be the same close family, so that to me, that's an irrelevant question. I can see good things that have come? No. (P6)

> ... "Ah well we've got Cal [the son], ummm, I guess it makes you feel thankful in different ways doesn't it, makes you appreciate little things. But... no I wouldn't think that was very relevant to this sort of thing, personally." (P3)

Item 8 ('I feel positive about the future') and item 20 ('I am able to make plans for the future') emerged as the strongest candidates. Both received some criticism for being vague, but nevertheless were highly valued (Table 2).

> ... "I mean number 8 is good, I feel positive about the future. Possibly a little bit vague. It doesn't actually specify somebody might think that their horse is going to come in tomorrow at some race and they're going to win a whole lot of money. They might be positive because of that." (P1)

... “I like number 8, it’s a nice all encompassing statement. But what if somebody has some other issue or some other hope in their mind and they think, ‘oh yeah I’m positive’, it doesn’t have anything to do with this. (P1)

... “I think it’s a good question to have in there [item 8], but it’s just a case of, I feel positive about the future, it’s like what aspect of the future? I feel positive about the future of my health, or the future of my mental health, or just feeling positive about the future in general. I think it’s just a bit too open ended.” (P5)

... “[Item 20] That’s why we went really. You know, we were thinking about having a family. We needed the information before we put the plan in place. So yeah, that’s quite a valid question. (P4)

... “[Item 20] How far in the future do they mean? Do they mean a few weeks, or a few years into the future? Our lives are constant planning, everything has to be risk assessed and planned in advance, there is very little that we can do spontaneously. Umm. I don’t know. It’s a difficult one because how far in advance are they asking you to look? Am I looking to plan 5 years, or am I looking to when my kids are adults? It’s difficult, I can’t really answer it.” (P7)

Table 2 presents those items highly valued by participants for each dimension of empowerment. Considering the diversity of outcomes within behavioural control, and the high value given to them by participants, it was observed that multiple items may merit inclusion if they capture different aspects of the dimension.

One noticeable trend throughout the interviews was the confusion experienced by participants when answering items beginning with ‘I don’t know’. Of the ten participants, seven selected a response option contrary to what they meant when asked about such an item. One individual (P4) recognised this when asked about item 18:

> ... "Erm... so I would disagree to that [item 18]. Because it's like a negative isn't it. Umm if I've read it correct. So it says I don't know, but actually I'm saying I do know, so I would have to disagree with that statement. So that might be slightly confusing to someone. You may get a couple of false positives, if someone misunderstands the question. It is common for these, I've done it myself when I've had to write these kind of evaluations, as soon as you put in the word 'I do not' or 'I don't', you know those kind of things, you sometimes get people who misunderstood the question. So I would I would say strongly disagree or disagree. Because after the service I did know [who else in my family might be at risk for this condition]." (P4)

The consequence of this finding was an agreement within the research team that any items containing 'I don't know' would be reworded to 'I know' if retained for the reduced scale.

Phase II: Quantitative Analysis

The Graded Response Model (GRM)

The GCOS-24 GRM item characteristic curves (shown in Supplemental Data) demonstrate the likelihood of respondents selecting a certain response option at various levels of the latent trait. Figure 2a presents the GRM output for item 15 as an example of good discriminative ability. Clear peaks can be seen ordered from 'Strongly Disagree' at low levels of the latent trait to 'Strongly Agree' at high levels of the latent trait; a measurement scale would benefit from inclusion of such an item. In contrast, Figure 2b presents the GRM output for item 22, a poor discriminator showing that 'Strongly Disagree' was the primary choice across a large range of the underlying trait. Those items which did not discriminate well between different levels of latent trait were removed in Phase III. Problems included poorly defined peaks, meaning failure to differentiate between individuals; excessive skew, meaning a significant floor or ceiling effect; and excessive dominance of one response option. Such items would

make poor candidates for inclusion in the reduced scale. An example item information curve is provided in Fig 3. The plot includes items designed to capture 'emotional regulation', and shows that item 4 would be the best candidate. Complete GRM results can be found in the Supplemental Data, and their application to item selection is described in Phase III.

<<insert Figure 2 here>>

Phase III: Item Selection

One aim of this study was to develop a measure which could be used outside the context of CGS, where genetic and genomic tests are done e.g. oncology, paediatrics. Items 1, 14 and 23 were therefore not considered for selection because they specifically refer to 'clinical genetics services'. The three principles of item selection will now be addressed.

(i) Items with poor discriminative properties were not selected:

Item characteristic curves were visually assessed. A flexible approach was used for data interpretation, since it was recognised that a number of factors may contribute to an item's quantitative properties. For example, because the cognitive interview findings indicated that a significant proportion of respondents experience confusion when answering items beginning with 'I don't know', such items were not immediately rejected for displaying poor discrimination. Additionally, items asking specifically about children were expected to show a prominent peak for Option 4 ('Neither Agree nor Disagree / Not Applicable'). Following consideration, item 2; item 3; item 5; item 6; item 7; item 10; item 13; item 19; item 21; and item 22 showed an unjustifiable inability to discriminate and were therefore removed from further consideration.

(ii) To avoid redundancy, items capturing a similar outcome were not retained together; item information curves and cognitive interview findings were used to establish superior items:

Cognitive control: Items 12 ('I don't know if this condition could affect my other relatives (brothers, sisters, aunts, uncles, cousins)') and 18 ('I don't know who else in my family might be at risk for this condition'). Neither item could be differentiated based on results at this point. Both were highly valued by cognitive interview participants (Table 2), and GRM results were indivisible.

All remaining items within decisional control, behavioural control and hope were considered sufficiently distinct. In emotional regulation, Item 11 was removed for having a far inferior item information curve to item 4 (Figure 3) and because cognitive interview results suggested that anxiety levels may not reduce over time, and instead fluctuate depending on the situation. This is not a desirable property in a scale designed to measure outcomes.

Ten items remained in consideration at this stage (Table 3).

<<insert Table 3 here>>

<<insert Figure 3 here>>

(iii) Rasch Analysis was used to determine and validate the optimal scale:

Rasch model fit was assessed using various item combinations, with the only restriction/assumption being that each dimension of empowerment should be represented. A series of Rasch rating scale models were run starting with the selected 10 items and removing iteratively based on results (Table 4). Optimal performance was observed using a six-item scale containing items 4, 16, 17, 18, 20 and 24 (Table 4). Pearson correlation of the measure with this 6-item instrument with GCOS-24 was r=.838 at 99% confidence, indicating that the new scale retains the ability to capture the empowerment construct. The Rating scale characteristic curve for the short form instrument's items (Figure 4) shows only five peaks, suggesting that 'Slightly Disagree' and 'Slightly Agree' could be collapsed with 'Disagree'

and ‘Agree’, respectively, without compromising scale quality. This is supported by further rating scale statistics (Tables 6 & 7).

<<insert Table 4 here>>

<<insert Figure 4 here>>

<<insert Table 5 here>>

<<insert Table 6 here>>

The final 6-item scale, which we have named ‘The Genetics Outcome Scale’ (GOS) is presented in Figure 5. We re-named the scale in this way because the name ‘The Genetic Counselling Outcome Scale’ could suggest that any health care provider who discusses a genetic risk or a genetic test with a patient is conducting genetic counselling. One aim of developing the short form is to enable patient outcomes to be assessed where genetic tests are done outside the traditional context of clinical genetics services e.g. oncology, paediatrics. To eliminate the confusion over double-negatives, items 17 and 18 were re-worded to change ‘I don’t know’ to ‘I know’. A five-point Likert scale was adopted, and ‘Not Applicable’ was eliminated from the response options following indications from the results in relation to rating scale functioning. Such a response, if present in the middle of the scale as in GCOS-24, especially since in GCOS-24 it is combined with ‘neither agree nor disagree’ compromises the desired ordinal nature of the items and thus the quality / validity of the measurement construct (Smith *et al.,* 2003). Furthermore, item 24 from GCOS-24 was re-worded to be 'I can make decisions about the condition that may change my future or my child(ren)'s future'. This was done to further eliminate the need for ‘not applicable’ responses, and because interview participants found items only applicable to those with children to be problematic.

<<insert Figure 5 here>>

Discussion

This study has developed a short-form (6-item) version of the Genetic Counselling Outcome Scale, potentially suitable for use in clinical audit and clinical evaluations of genetic counselling and testing services. The new scale, 'The Genomics Outcome Scale' or GOS, maintains the ability of GCOS-24 to capture the theoretical construct of empowerment (McAllister *et al.,* 2011a), with the two scales showing a correlation of r=.838 at 99% confidence. Whilst the breadth of the latent trait captured by GCOS-24 has been maintained, sub-dimensional analysis will not be possible with the shorter form, though single-item analysis is a possibility for GOS items representing the previous sub-dimensions. With genetic testing increasingly being performed in contexts outside the traditional models of service provision (Lo *et al.,* 2014, Rahman*,* 2014; Valente *et al.,* 2008) GOS was designed to be applicable both within and outside of clinical genetics services. GOS provides a less burdensome measurement scale for respondents, and produces a significantly reduced number of response permutations ($1.56x10^4$) compared to GCOS-24 ($1.92x10^{20}$).

This study represents the first instance of Item Response Theory (IRT) analysis on GCOS-24, and findings support the call for wider use of IRT methods in PROM development (Embretson, 1996; Hambleton & Jones, 1993; Nguyen, 2014; Reeve, 2002). For despite conferring a number of benefits over the traditional approach of classical test theory (CTT), most notably the ability to examine measurement properties at the item-level, CTT remains largely unquestioned and continues to guide the construction, scoring, refinement, and validation of (PROMs) (Nguyen, 2014), although IRT methods are gaining traction (e.g. Bailey *et al.* (2017). Indeed the importance of IRT analysis in PROM development was recently emphasised by Nguyen *et al.* (2014), who describe how IRT can greatly improve the efficiency and accuracy of PROM measurement, and by Alonso *et al.* (2013), who discuss how the creation of IRT-derived item banks can benefit the outcome measurement community.

Before any instrument designed to capture the impact of specified outcomes can be used in research or clinical practice, COSMIN guidelines state that its measurement properties, i.e. reliability, responsiveness and interpretability, should be assessed and considered adequate (Mokkink *et al.,* 2010). GOS shows the potential for further use, but an important limitation of the current study is that these properties of GOS have not yet been assessed. Further psychometric assessment for test-retest reliability, responsiveness and interpretability (e.g. establishment of the minimum clinically important difference) are needed before widespread use. GOS could now be used in the identification of preference-weights for each descriptive state represented by the items and response levels in the measure, now that the 'Not Applicable' response option has been removed. Preference weights must be elicited for all options to satisfy the calculations, yet assigning a value to the non-ordinal 'Not Applicable' would make the calculations unworkable. Ultimately, further testing and development of GOS would allow the use of this condition-specific measure as the outcome of interest in economic evaluations of genetic and genomic services and tests.

One of the great strengths of GOS is that it was developed from GCOS-24: an internationally recognised PROM of proven validity, reliability and responsiveness which specifically measures outcomes valued by CGS patients. GCOS-24 content validity has previously been demonstrated in the context of CGS (Costal-Tirado *et al.,* 2017; McAllister *et al.,* 2011b), and clinical and research uses have been reported both in the UK and internationally (Inglis *et al.,* 2014; McAllister &Dearing, 2015; Diness *et al.,* 2017; Munoz-Cabello *et al.,* 2017). The substantial qualitative research underpinning GCOS-24, coupled with the further qualitative research in the current study, supports the potential implementation of GOS in CGS and in future research, following important further psychometric assessment.

The inclusion of both quantitative and qualitative methods in the present study design proved to be a valuable decision. Each provided more information on GCOS-24 items than either

method alone, supporting a holistic approach to item selection and strengthening the validity of GOS. An important limitation was the small sample size for the cognitive interviews. Sample homogeneity during the interview stage was a further limitation for this study, with all ten participants being white-British and resident in South-Wales. Cognitive interview findings are therefore limited to the perspective of a small number of families who live in a specific region and speak English. Additionally, only one of the ten interviewees self-reported as 'at-risk' for a genetic condition. Seven were affected, and two were unaffected with an affected child. At-risk individuals are a key demographic of CGS users, and higher representation would have been beneficial. A further limitation was the unknown influence of having 'Not Applicable' included into Option 4 of the GCOS-24 Likert Scale. Incorporating a non-ordinal option into an ordinal scale disrupts the nature of the scale and would have affected quantitative outputs.

Quantitative approaches should be used with caution when analysing subjective topics. In the present study, rather than taking statistical findings at face value, reasonable judgement informed by cognitive interview findings was applied. This is emphasised by the GRM results in this study. Items which at first glance appeared to perform poorly could be viewed alongside the available interview data to reflect an issue with item wording rather than a problem with the underlying outcome. Items asking about children showed a prominent peak for the 'Not Applicable' response option, reflecting the reality that not all CGS users have children; items beginning with 'I don't know' generally performed poorly, reflecting the reality that many people mistakenly interpret the double-negative. In short, both qualitative and quantitative methods offer powerful tools for PROM development.

There is also scope to further develop GOS. The sixth item of GOS ('I can make decisions about the condition that may change my child(ren)'s future / the future of any child(ren) I may have) will be of little relevance to those who do not have and do not plan on having

children; perhaps the respondent is unable to have children. Rewording the item to include a reference to the patient's own future could be beneficial, e.g. 'I can make decisions about the condition that may change my future, my child(ren)'s future / the future of any child(ren) I may have'. Furthermore, elimination of GCOS-24 item 2 ('I can explain what the condition means to people in my family who may need to know') with retention of GCOS-24 item 16 ('I can explain what the condition means to people outside my family who may need to know') resulted in a loss of useful information. This could be addressed by merging the two items to create a new item: 'I can explain what the condition means to people who may need to know'. Such changes, however, were beyond the scope of the present study, as the newly worded items and the resulting new scale would require validation.

In conclusion, GCOS-24 has been shortened to a six-item measure with a five-point Likert scale (GOS). GOS has the potential to be applicable both within and outside the context of clinical genetics, and with only six items will be less burdensome to patients than GCOS-24. Correlation between the two measures (r=.838) suggests that GOS maintains the ability to capture the underlying construct of empowerment. This study could represent the first step in developing a preference-based measure for use in evaluations of genetic counselling and associated genomic testing services.

Acknowledgements

We extend our sincere thanks to all the patient support group members who contributed their time to participate in interviews for this study. We would also like to gratefully acknowledge the significant contribution of the Genetic Alliance UK, who recruited participants for the interview study. Their time, energy and commitment made this work possible. The copyright for GCOS-24 is owned by the Wiley publishing company, and we are grateful for their permission to reproduce GCOS-24 in this article. This was an unfunded Cardiff University MPhil dissertation project, conducted by Peter Grant with supervision from the other authors.

The dataset used for the quantitative study was collected in 2007-10 by Marion McAllister as part of the study to develop and validate GCOS-24, funded by the Medical Research Council.

Figure Titles and Legends

Figure 1: The Genetic Counselling Outcome Scale (GCOS-24) (McAllister *et al.,* 2011b).
Figure 2: 2a (left) & 2b (right): Example GRM item characteristic curves showing items with good (2a) and poor (2b) discriminative properties.
Figure 3: Example GRM item information curve (Emotional Regulation).
Figure 4: Rating Scale characteristic curve
Figure 5: The final scale, the Genomics Outcome Scale (GOS)

Tables

Table 1: Interview Participant Characteristics

Participant	Sex	Condition	Affected, At risk, Unaffected	Has a child?	Received Genetic Counselling?
P1	Male	Nystagmus	Affected	No	No
P2	Male	Ataxia	Affected	No	Yes
P3	Female	Tubular Sclerosis	Unaffected	Yes	Yes
P4	Male	Glaucoma	Affected	No	Yes
P5	Female	Thalassemia Intermedia	Affected	Yes	Yes
P6	Female	Episodic Ataxia	Unaffected	Yes	No
P7	Female	Ehlers Danlos Syndrome	Affected	Yes	No
P8	Female	Dystonia & Ataxia	Affected	Yes	No
P9	Female	Huntington's Disease	At risk	Yes	Yes
P10	Male	Leber's Hereditary Optic Neuropathy	Affected	No	No

Table 2: Empowerment sub-dimensions, with definitions, showing which GCOS-24 items were designed to capture each dimension, and which of those items were valued by cognitive interview participants in this study.

Empowerment Dimension	**Definition (The belief that one...)**	**Corresponding GCOS-24 Items**	**Highly Valued Items***
Cognitive Control	...has sufficient information about the condition, including risks to oneself and one's relatives, and any treatment, prevention and support available.	1, 3, 12, 14, 18, 23	12 or 18
Decisional Control	...can make important life decisions in an informed way.	10, 13, 24	24
Behavioural Control	...can make effective use of the health and social care systems	2, 5, 7, 9, 15, 16, 17, 22	2 or 16; 5 or 15; 17; 9
Emotional Regulation	...can manage their feelings about having a genetic condition in the family	4, 11, 21	4
Hope	...can look to the future having hope for a fulfilling family life, for oneself, one's family, and/or one's future descendants	6, 8, 19, 20	8, 20

*Highly valued items with a similar meaning are separated by 'or'.

McAllister, M., Dunn, G., Todd, C. (2011a) Empowerment: qualitative underpinning of a new clinical genetics-specific patient-reported outcome. *Eur J Hum Gen* **19(2)**, 125-130.

Table 3: Items remaining in consideration following principle 2 of item selection.

Dimension of Empowerment	**Item Under Consideration**
Cognitive Control	(12) I don't know if this condition could affect my other relatives (brothers, sisters, aunts, uncles, cousins).
	(18) I don't know who else in my family might be at risk for this condition.
Decisional Control	(24) I can make decisions about the condition that may change my child(ren)'s future / the future of any child(ren) I may have.
Behavioural Control	(9) I am able to cope with having this condition in my family.
	(15) I know how to get the non-medical help I/my family needs (e.g. educational, financial, social support)
	(16) I can explain what the condition means to people outside my family who may need to know'
	(17) I don't know what I can do to change how this condition affects me/my children.
Emotional Regulation	(4) When I think about the condition in my family, I get upset.
Hope	(8) I feel positive about the future.
	(20) I am able to make plans for the future

Table 4: Rasch Analysis and iterative removal of items. Bolded row corresponds to optimal item combination.

	Person		Item				
	Separation	Reliability	Separation	Reliability	Mean Infit MNSQ (SD)	Mean Outfit MNSQ (SD)	Problematic items (infit>1.3)
All 10 selected items	1.42	0.67	9.25	0.99	1.01 (0.26)	1.05 (0.29)	Item 12 (infit:1.49) Item 18 (1.37)
9 items (no 18)	1.39	0.66	9.18	0.99	1.01 (0.3)	1.05 (0.34)	Item 12 (1.73)
8 items (no 18, 9)	1.21	0.59	9.41	0.99	1 (0.26)	1.05 (0.3)	Item 12 (1.62)
7 items (no 8,9,18)	1.03	0.51	9.78	0.99	1 (0.23)	1.05 (0.25)	Item 12 (1.47)
9 items (no 12)	1.5	0.69	9.96	0.99	1.01 (0.27)	1.05 (0.31)	Item 18 (1.62)
7 items (no 12, 8, 9)	1.15	0.57	10.56	0.99	1.01 (0.21)	1.04 (0.23)	Item 18 (1.41)
6 (no 12, 8, 9, 15)	**1.14**	**0.56**	**11.52**	**0.99**	**1.01 (0.19)**	**1.04 (0.2)**	**Item 18 (1.32)**
6 (no 12, 8, 9, 16)	1.06	0.53	10.45	0.99	1.02 (0.22)	1.05 (0.24)	Item 18 (1.39)
6 (no 12, 8, 9, 17)	0.96	0.48	10.54	0.99	1.01 (0.19)	1.04 (0.23)	Item 18 (1.38)

Table 4. **Person Separation:** how efficiently a set of items can separate persons measured. **Item Separation**: how well a sample of people is able to separate each item. **Person reliability:** does the test discriminate the sample? Depends chiefly on sample ability variance, length of test, and number of categories per item. **Item reliability:** is the sample large enough to precisely locate items within the latent variable? Depends chiefly on item difficulty and person sample size.

Infit (inlier-sensitive fit): Sensitive to response patterns to items targeted at the person. **Outfit (outlier-sensitive fit)**: Sensitive to response patterns to items targeted away from the person e.g. underfit for lucky guesses or mistakes. **Mean-squares (MNSQ)**: the size of the randomness i.e. amount of distortion. 1.0 expected value; <1 indicates predictability and redundancy; >1 indicate unpredictability.

Table 5: Rasch results for optimal item combination (GOS).

Item name	Total Raw Score	Observed Count	Measure	SE	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
GCOS4P	1261	395	0.46	0.04	1.12	1.8	1.22	2.9
GCOS16	1948	395	-0.40	0.04	0.97	-0.4	1.02	0.4
GCOS17P	1342	395	0.36	0.04	0.70	-5.3	0.71	-4.8
GCOS18P	1291	395	0.42	0.04	1.32	4.5	1.32	4.3
GCOS20	2144	395	-0.70	0.04	1.06	0.9	0.97	-0.4
GCOS24	1751	395	-0.14	0.04	0.91	-1.5	1.02	0.4
		Mean:	0.00	0.04	1.01	0	1.04	0.5
		SD:	0.44	0.00	0.19	3.0	0.2	2.9
PERSON: REAL SEPARATION.: 1.14 RELIABILITY.: 0.56								
ITEM: REAL SEPARATION: 11.52 RELIABILITY: 0.99								

Table 6: Rasch Rating scale results.

	Person		Item			
	Separation	Reliability	Separation	Reliability	Mean Infit MNSQ (SD)	Mean Outfit MNSQ (SD)
Short-6 recoded 1234567 → 1233345	1.06	0.53	10.72	0.99	0.99 (0.18)	1.01 (0.18)
1234567 →1223445	1.05	0.53	11.63	0.99	1 (0.18)	1.02 (0.18)

Supplemental data:

GCOS-24 Items: Item Characteristic Curves and Item Information Curves

References

Alonso, J., Bartlett, S.J., Rose, M., Aaronson, N.K., Chaplin, J.E., Efficace, F., Leplège, A., Lu, A., Tulsky, D.S., Raat, H., Ravens-Sieberer, U., Revicki, D., Terwee, C.B., Valderas, J.M., Cella, D., Forrest, C.B.; PROMIS International Group. (2013) The case for an international patient-reported outcomes measurement information system (PROMIS®) initiative. Health Qual Life Outcomes. Dec 20;11:210. doi: 10.1186/1477-7525-11-210.

Andrich, D. (1978) A rating formulation for ordered response categories. *Psychometrika*, **43**, 561-73

Bailey, C., Tully, M.P., Pampaka, M., Cooke, J. (2017). Rasch analysis of the Antimicrobial Self-Assessment Toolkit for National Health Service (NHS) Trusts (ASAT v17). *Journal of Antimicrobial Chemotherapy, 72*(2), 604-613. doi:10.1093/jac/dkw434

Barr, P.J., Scholl, I., Bravo, P., Faber, M., Elwyn, G., McAllister, M. (2015) Assessment of patient empowerment: A systematic review of measures. *PLOS ONE* May 13;10(5):e0126553

Bernhardt, B., Biesecker, B., Mastromarino, C. (2000) Goals, benefits and outcomes of genetic counselling: Client and genetic counsellor assessment. *American Journal of Med Gen* **94**, 189-197.

Bjorner, J., Kosinski, M., Ware, J. (2003) Using item response theory to calibrate the head impact test (HIT) to the metric of traditional headache scales. *Qual Life Res* 12(8): 981-1002

Bravo, P., Edwards, A., Barr, P.J., Scholl, I., Elwyn, G, McAllister, M. (2015) Conceptualising patient empowerment: a mixed methods study. *BMC Health Serv Res* 2015; 15:252 DOI: 10.1186/s12913-015-0907-z.

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J.F., Bruce, B., Rose, M; PROMIS Cooperative Group. The Patient-Reported Outcomes

Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care. 2007 May;45(5 Suppl 1):S3-S11.

Clarke, A., Parsons, E., Williams, A. (1996) Outcomes and process in genetic counselling. *Clinical Genetics* **50,** 462-469.

Costal-Tirado, A., McDermott, A., Thomas, C., Ferrick, D., Harris, J., Edwards, A., McAllister, M. (2017) Using patient-reported outcome measures for quality improvement in clinical genetics: an exploratory study. *Journal of Genetic Counseling*, 1-12.

Diness, B., Overbeck, G., Hjortshoj, T., Hammer, T., Timshel, S., Sorensen, E., McAllister, M. (2017) Translation and Adaptation of the Genetic Counselling Outcome Scale (GCOS-24) for use in Denmark. *J Genet Couns*

Embretson, S. (1996) The new rules of measurement. *Psychol Assess* **8(4)**, 341-349.

Ericsson, K., Simon, H. (1980) Verbal reports as data. *Psychological Review* **87,** 215-250.

Haley, S., Coster, W., Andres, P., Kosinski, M., Ni, P. (2004) Score comparability of short forms and computerised adaptive testing. *APMR* **85(4)**: 661-666.

Hambleton, R., Jones, R. (1993) Comparison of classical test theory and item response theory and their applications to test development. *Instructional Topics in Educational Measurement*, 38-47.

Harniss, M., Amtmann, D., Cook, D., Johnson, K. (2007) Considerations for developing interfaces for collecting patient-reported outcomes that allow the inclusion of individuals with disabilities. *Med Care* **45**:48-54.

Inglis, A., Koehn, D., McGillivray, B., Stewart, S. (2015) Evaluating a unique, specialist psychiatric genetic counselling clinic: uptake and impact. *Clinical Genetics* **87(3)**: 218-224.

Irwin, D., Varni, J., Yeatts, K., DeWalt, D. (2009) Cognitive interviewing methodology in the development of a pediatric item bank. *Health Qual Life Outcomes* **7(3)**.

Linacre, J. (2018) Winsteps® Rasch measurement computer program. Beaverton, Oregon.

Lo, C., Martindale, J., Hadjivassiliou, M., Martin, P., Dalton, A., Bandmann, O. (2014) The Documentation of Consent and Disclosure of Neurogenetic Testing Outside Clinical Genetics. *Neurogenetics* **15(1)**: 19-21.

Macleod, R., Craufurd, D., Booth, K. (2002) Patients' Perceptions of what makes Genetic Counselling effective: An interpretative phenomenological analysis. *J Health Psy* **7(2)**: 145-156.

Madlensky L, Trepanier AM, Cragun D, Lerner B, Shannon KM, Zierhut H. (2017) A Rapid Systematic Review of Outcomes Studies in Genetic Counseling. J Genet Couns. Jun;26(3):361-378. doi: 10.1007/s10897-017-0067-x. Epub 2017 Feb 6.

McAllister, M., Payne, K., Macleod, R., Nicholls, S., Donnai, D., Davies, L. (2008) Patient empowerment in clinical genetics services. *J Health Psychology* **13(7)**, 895-905.

McAllister, M., Dunn, G., Todd, C. (2011a) Empowerment: qualitative underpinning of a new clinical genetics-specific patient-reported outcome. *Eur J Hum Gen* **19(2)**, 125-130.

McAllister, M., Wood, A., Dunn, G., Shiloh, S., Todd, C. (2011b) The Genetic Counselling Outcome Scale: a new patient-reported outcome measure for clinical genetics services. *Clinical Genetics* **79**: 413-424.

McAllister, M., Dearing, A. (2015) Patient reported outcomes and patient empowerment in clinical genetic services. *Clin Genet* **88**, 114-121.

McAllister, M. (2016) Genomics and patient empowerment. Kumar & Chadwick (eds) *Genomics and Society.* Philadelphia: Elsevier.

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P. *et al.* (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of

health status measurement instruments: an international Delphi study. *Qual Life Res* **19(4)**: 539-549.

Munoz-Cabello, P., Garcia-Minaur, S., Espinal-Vallejo, M. *et al.* (2018) Translation and Cross-Cultural Adaptation with Preliminary Validation of GCOS-24 for Use in Spain. *J Genet Couns* **27(3)**:732-743.

Nguyen, T., Han, H., Kim, M., Chan, K. (2014) An introduction to Item Response Theory for Patient-reported Outcome Measurement. *Patient*, **7(1)**, 23-35

Payne, K., Nicholls, S. G., McAllister, M., *et al*. (2007) Outcome measures for clinical genetics services: a comparison of genetics healthcare professionals and patients' views. *Health Policy* **84(1),** 112-122.

Payne, K., Nicholls, S., McAllister, M., MacLeod, R., Donnai, D., Davies, L. (2008) Outcome Measurement in Clinical Genetics Services: A systematic review of validated measures. *Value in Health* **11(3)**: 497-508.

Rahman, N. (2014) Realizing the promise of cancer predisposition genes. *Nature* **505(7483)**: 302-308.

Ritchie, J., Spencer, L. (1994) Qualitative data analysis for applied policy research. *Analysing qualitative data*, 173-194.

Rizopoulos, D. (2006) ltm: An R package for Latent Variable Modeling and Item Response Theory Analyses. *Journal of Statistical Software*, **17(5)**.

Samejima, F. (1969) Estimation of latent ability using a response pattern of graded scores. *Psychom Monogr* **34**, 386-415.

Sinnott, P., Joyce, V., Barnett, P. (2007) Preference Measurement in Economic Analysis. Guidebook. Health Economics Resource Center CA.

Skirton, H. (2001) The Client's Perspective of Genetic Counselling – A Grounded Theory Study. *J Genet Couns* **10(4)**: 311-329.

Smith, J., Wakely, M., De Kruif, R., Swartz, C. (2003) Optimising Rating Scales for self efficacy (and other)research. *Educ and Psy Meas* **63 (3):** 369-391.

Streiner, D. (2003) Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess* **80(1)**: 99-103.

Terwee, C., Mokkink, L., Knol, D., Ostelo, R., Bouter, L., de Vet, H. (2012) Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* **21(4)**: 651-657.

Valente, E., Ferraris, A., Dallapiccola, B. (2008) Genetic Testing for Paediatric Neurological Disorders. *Lancet Neurol* **7(12)**: 1113-26.

Wang, C., Gonzalez, R., Merajver, S. (2004) Assessment of genetic testing and related counselling services: current research and future directions. *Soc Sci Med* **58**: 1427-42.