BMJ Open Respiratory Research

Novel method to select meaningful outcomes for evaluation in clinical trials

Charlie McLeod,[1,2] Richard Norman,[3] Jamie Wood,[4,5] Siobhain Mulrennan,[6,7] Sue Morey,[6] André Schultz,[8,9] Mitch Messer,[10] Kate Spaapen,[11] Matthew Stoneham,[12] Yue Wu,[13] Alan Smyth,[14] Christopher Blyth,[1,10] Steve Webb,[15,16] Steven Mascaro,[17] Owen Woodberry,[17] Tom Snelling[13,18]

For numbered affiliations see end of article.

Correspondence to
Dr Charlie McLeod;
charlie.mcleod@health.wa.gov.au

ABSTRACT

Background A standardised framework for selecting outcomes for evaluation in trials has been proposed by the Core Outcome Measures in Effectiveness Trials working group. However, this method does not specify how to ensure that the outcomes that are selected are causally related to the disease and the health intervention being studied. Causal network diagrams may help researchers identify outcomes that are both clinically meaningful and likely to be causally dependent on the intervention, and endpoints that are, in turn, causally dependent on those outcomes. We aimed to (1) develop a generalisable method for selecting outcomes and endpoints in trials and (2) apply this method to select outcomes for evaluation in a trial investigating treatment strategies for pulmonary exacerbations of cystic fibrosis (CF).

Methods We conducted a series of online surveys and workshops among people affected by CF. We used a modified Delphi approach to develop a consensus list of important outcomes. A workshop involving domain experts elicited how these outcomes were causally related to the underlying pathophysiological processes. Meaningful outcomes were prioritised based on the extent to which each outcome captured separate rather than common aspects of the underlying pathophysiological process.

Results The 10 prioritised outcomes were: breathing difficulty/pain, sputum production/clearance, fatigue, appetite, pain (not related to breathing), motivation/demoralisation, fevers/night sweats, treatment burden, inability to meet personal goals and avoidance of gastrointestinal symptoms.

Conclusions This proposed method for selecting meaningful outcomes for evaluation in clinical trials may improve the value of research as a basis for clinical decisions.

Key messages

⇒ Which outcomes are both meaningful to people affected by cystic fibrosis (CF) and causally related to the underlying pathophysiological processes of disease and the mechanism of action of the intervention in question.

⇒ This study identified 10 outcomes that are both meaningful and likely to be causally affected by treatment(s) for pulmonary exacerbations in CF. This is the first step towards the development of weighted outcome measures for use when evaluating the effect of treatment interventions for pulmonary exacerbations in adults and children.

⇒ We present a rational approach for selecting outcomes that are meaningful, causally related to disease processes, and likely to be impacted by interventions under study; application of this method could improve the quality of clinical research.

INTRODUCTION

Selecting appropriate outcomes for evaluation in clinical studies is critical for ensuring the value of that research.[1] The Core Outcome Measures in Effectiveness Trials (COMET) initiative was established in 2010, and provides a framework for the development of core outcome sets (COSs). COSs are a minimum set of meaningful outcomes that should be measured and reported in all trials of a specific disease or study population.[2] COMET specifies that the outcomes included should arise from exposure to a causal factor or a health intervention.[2] The rationale for this is that outcomes that are not causally related to the pathophysiology of the disease of interest and dependent on the mechanism of the intervention being studied may produce misleading information.[3] Causal network diagrams may help researchers identify outcomes that are both clinically meaningful and likely to be causally dependent on the intervention, and endpoints that are, in turn, causally dependent on these outcomes; it may also help to identify endpoints that are likely to capture very similar information about the overall experience, because they are causally related to the same outcome, or to outcomes that are causally closely related.[4]

Outcomes can be defined as patient characteristics or biological processes targeted for improvement by an intervention (eg, lung function), and endpoints as the specific measurable parameter(s) corresponding to those outcomes (eg, change in the percentage

predicted forced expiratory volume in 1 s (FEV_1) from baseline to day 10) (McLeod *et al*, unpub). Outcomes in clinical studies should be patient-centred and meaningful; that is, they should capture either directly, or indirectly (as a valid surrogate), how a person feels, functions or survives.[5] To this end, outcomes should be considered important to those affected by the disease in question. It is increasingly recognised as important to include consumers in the selection of outcomes in clinical studies, but this has only occasionally been done.[5]

A causal network diagram (or directed acyclic graph) can help to represent the causal relationships between relevant factors (whether measurable or not), including the various relevant outcomes of an intervention for a given disease. Causal network diagrams use unidirectional arrows, or 'arcs', to connect factors in a pairwise fashion in which the direction of the arc represents the causal direction of the relationship, that is, from cause to effect.[6] They can be used to explicitly represent and explore our understanding of the causal mechanisms underlying a problem domain, or for clinical problems, the pathophysiological processes which give rise to various symptoms, functional outcomes and disease states (including death), and how and where in the process various treatments are thought to act.[7]

Causal network diagrams may also help to identify factors that are not related causally to outcomes of clinical importance, and which are therefore unlikely to be useful or reliable as surrogate outcomes, even if correlated with the outcome.[4 8] For example, the number of cigarettes smoked per day may correlate strongly with lung function in people with emphysema; if evaluating the impact of a quit smoking intervention, measuring any reduction in smoking may be a useful surrogate for (eventual) improved lung function, but measuring reduction in smoking would not be reliable as a surrogate if evaluating the effect of inhaled steroid therapy on lung function in emphysema. Inspection of a causal network diagram (see figure 1) quickly reveals that smoking intensity lies as a mediator on the causal pathway between the quit smoking intervention and lung function, as it is both an effect of the intervention and the cause of the change in lung function. But smoking intensity does not lie as a mediator on the causal pathway between steroid therapy and lung function, even if smoking and lung function remain strongly correlated. Finally, causal network diagrams may help to identify any causal relationships between outcomes, particularly whether one outcome is causally dependent on another, or whether they share another outcome as a common cause. In general, sets of outcomes should be selected in a way that ensures they provide maximal information about the overall outcome for the patient. For example, if measuring the impact of a quit smoking intervention using two outcomes, its impact on dyspnoea and angina may be more informative than its impact on dyspnoea and cough (which share airway inflammation as a common cause), or its impact on angina and myocardial infarction (which share coronary atherosclerosis as a common cause).

Here, we present a novel approach for selecting meaningful patient-centred outcomes when evaluating pulmonary exacerbations in people with cystic fibrosis (CF). We aimed to (1) develop a generalisable method for selecting outcomes and endpoints for evaluation in clinical studies and (2) apply this method to select outcomes for use when evaluating interventions for pulmonary exacerbation in CF. Our ultimate goal is to include these outcomes in outcome measure instruments which could be used to compare interventions by observing disparate endpoints among trial participants.
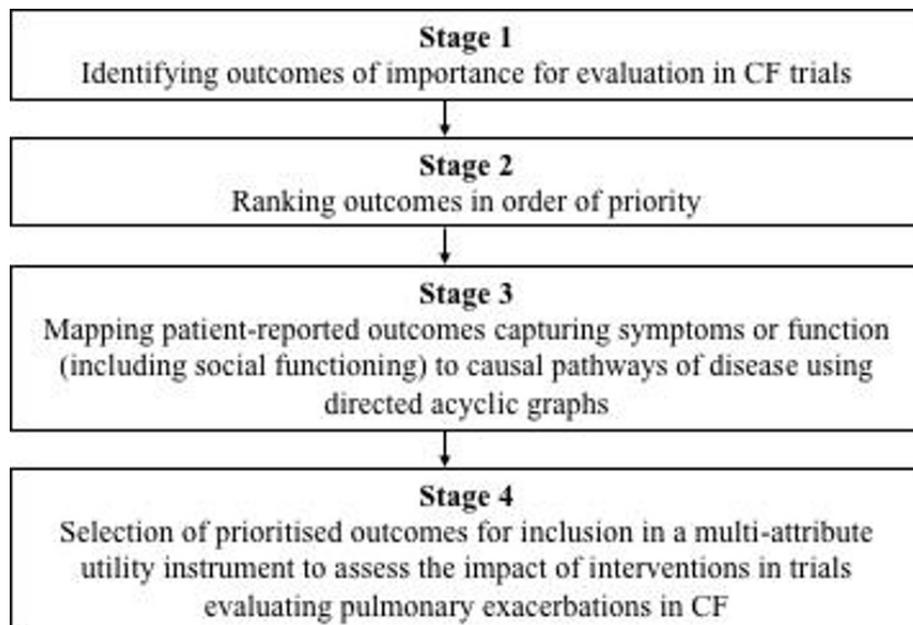
METHODS
Overview
This project comprised four stages conducted between October 2018 and August 2019 (figure 2). Written or
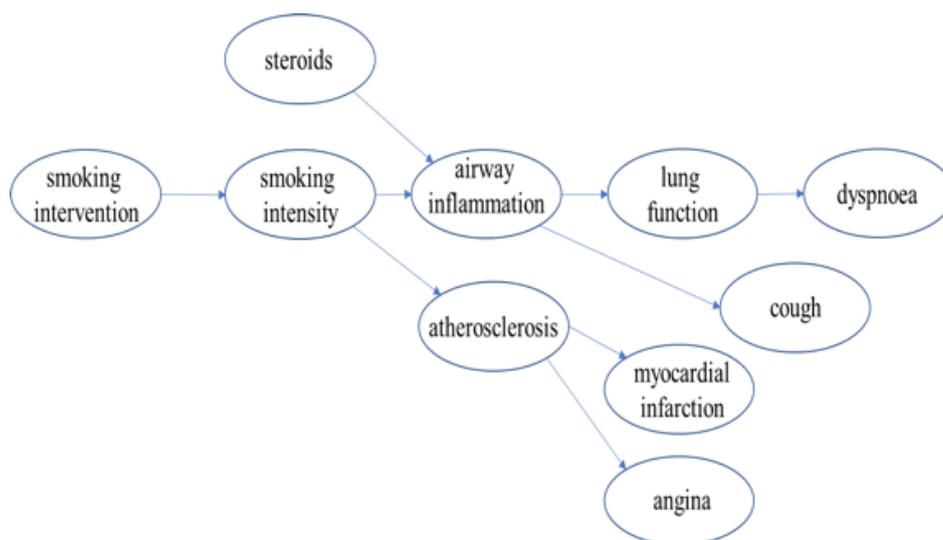
Figure 1 Causal diagram: smoking and myocardial infarction.

Figure 2 Stages of research.

online consent was obtained for participation in all stages. Participants were not paid to participate, however, those who attended workshops in-person received small compensation for incurred costs.

Patient and public involvement

People affected by CF were involved in the design and conduct of this research. During the feasibility stage, formulation of the specific research questions, methods of recruitment and wording for questionnaires were informed by workshop discussions with people affected by CF. Two consumers (MM and KS) were also included as investigators. This research area has also been identified as a priority for research by consumers through a previous James Lind consumer priority setting exercise.[9] Once published, participants will be informed of the results through a dedicated website (https://adaptive healthintelligence.org.au).

Stage 1

The literature was reviewed to identify a comprehensive range of outcomes and endpoints reported in CF clinical studies, and two 2-hour workshops were held using a modified Delphi approach[10] to elicit meaningful health outcomes among (1) people >13 years with CF and (2) carers of children or adults with CF. We aimed to elicit answers to two questions: (1) 'What CF-related health outcomes are important to people affected by CF?' and (2) 'What adverse effects of treatment are important to people affected by CF?' Attendees were asked to suggest their own outcomes, and when this was exhausted, they were asked to consider the importance of any other outcomes identified in the literature review. The workshop focused explicitly on clinically meaningful

outcomes, that is outcomes that aim to capture how a person feels or functions, rather than mechanistic outcomes (eg, spirometry, radiographic changes and sputum or blood inflammatory biomarkers). Workshops for carers were conducted in person, while those for people with CF were conducted via videoconference owing to the infection control risk of face-to-face contact. An online survey comprising the same two questions was conducted over 5 weeks (see online supplemental S1); participation was voluntary and advertised via local and national CF consumer groups and research networks, and the main paediatric and adult CF treatment centres in Perth, Australia.

Stage 2

Two further workshops were conducted for similarly comprised groups and under the same conditions described for stage 1. The aim of these workshops was to obtain agreement on a prioritisation of the outcomes identified in stage 1, capturing how people feel and function from the perspective of (1) people with CF ≥13 years old and (2) carers. The prioritisation exercise (see online supplemental S2) was also conducted as an online survey over 3 weeks.

Stage 3

An expert knowledge elicitation workshop (facilitated by SM) was conducted to try to derive an agreed causal network diagram linking the outcomes identified in stages 1 and 2 to underlying pathophysiological processes, and to each other. A group of clinical domain experts in CF pulmonary exacerbations, people with CF ≥13 years old and carers attended. Experts were provided a brief explanation of causal networks and then asked

to consider and either agree with, or revise, a baseline causal framework in which airway infection gives rise to inflammation (treatable by various agents) which in turn affects functional outcomes and gives rise to a range of symptoms. On agreement that this basic framework was appropriate and sufficient for the task, emphasis shifted to specific symptoms and functional outcomes. Prior to the workshop, these were partitioned into potentially discrete domains or subnetworks. Experts were provided the opportunity to review and revise these subnetworks, and were also allowed to adjust them as the workshop proceeded if the need became apparent.

Experts were asked to consider each such subnetwork in turn, and in isolation from other subnetworks. Initially, subnetworks consisted of 2–10 factors that were fully disconnected. Experts were prompted to identify the most causally important factors and, in particular, key common causes that might strongly influence many other factors within the subnetwork. They were also advised to keep the subnetworks simple, omitting weak causal influences, and to avoid the creation of cycles (in which a factor is described as causally dependent on itself) by specifying only the most immediate and dominant direction of causal influences. As experts suggested causal relationships between factors, the facilitator added them to the network, and the group considered them in terms of their strength, redundancy or potential conflict with other relationships. Based on the group consensus, causal relationships were then either retained or removed. Any factors that remained without any significant causal relationships to other factors at the end of the process were reconsidered and either retained, moved to another subnetwork or removed entirely if no longer considered worth keeping.

After each subnetwork was considered in isolation, the connections between subnetworks were considered in much the same way, with the exception that experts could suggest connections between either the subnetworks themselves or between any individual factors contained within them.

After the workshop, subnetworks were converted into Bayesian network submodels (structure only) and reviewed a second time by a smaller group of CF pulmonary exacerbation domain experts for any remaining issues and inconsistencies.

Stage 4

A single agreed list of 10 priority outcomes across both people with CF ≥13 years old and carers was produced by representatives from the group above. Priority outcomes were selected based on two considerations; (1) the importance of the outcome as assessed by both groups and (2) the extent to which each outcome captured separate rather than common aspects of the underlying pathophysiological process and the overall outcome experience.

RESULTS
Stage 1 and 2

Thirty-six people participated in stage 1: 6 people with CF ≥13 years old (one was <25 years old) and 8 carers participated via workshop; 7 people with CF (two aged <25 years old) and 15 carers participated via the online survey.

Fifty-five people participated in stage 2: 12 people with CF (one aged <25 years old) and 15 carers attended the prioritisation workshops; and 12 people with CF (none aged <25 years old) and 16 carers participated via the online survey.

A condensed list of identified outcomes based on results from stages 1 and 2 are presented in table 1. A more extensive list is provided in online supplemental S3.

Stage 3

Ten domain experts in CF pulmonary exacerbation participated in stage 3, including two paediatric respiratory physicians, two adult respiratory physicians, two paediatric infectious diseases physicians, three adults with CF and one carer. Based on the outcomes identified in stages 1 and 2, participants settled on eight broad pathophysiological domains which became the focus of separate causal subnetworks: respiratory, gut, sinus, endocrine, hearing, mental health, general and functional. General outcomes included those that captured an aspect of an individuals' overall health, such as fatigue or appetite. Functional outcomes comprised those that related to an individuals' ability to perform activities of daily living or realise their own aspirations.

Figure 3 illustrates the relationship between outcome domains and figure 4 illustrates consensus causal subnetwork models for the respiratory domain. Causal subnetwork models for general symptoms, the gastrointestinal system, mental health and functional outcomes are included in online supplemental S4–S7. Lighter arrows were used to indicate domains that were causally related to all other nodes.

Stage 4

Endocrine, hearing and sinus-related outcomes were not as highly prioritised by people with CF or carers compared with other domains (see table 1); individual outcomes within these domains were consequently not further considered for inclusion in the final list of priority outcomes.

Excessive sputum production/poor clearance was identified as an important respiratory outcome, and was found to be causally related to all other outcomes within the respiratory domain including severe cough, coughing 'spasms', wheezing, chest tightness and haemoptysis. Breathing difficulty (described as shortness of breath or consciousness or awareness of breathing) was found to be important to people with CF and carers alike, but being a relatively uncommon symptom in children may be less applicable in studies of pulmonary exacerbations compared with alternatives such as sputum production/

Table 1 Patient-reported outcomes ranked by perceived importance by people >13 years affected by CF and carers of children

Rank	People with CF ≥13 years	Carers of children 0–18 years
1	Coughing up blood	Coughing up blood
2	Shortness of breath/difficulty breathing	Shortness of breath/difficulty breathing
3	Feeling fatigued/deconditioned	Presence of pain
4	Anxiety/worry	Feeling anxious/worried
5	Sputum production (presence of, or worsening from baseline)	Feeling sad/depressed
6	Sadness/depressed mood	High treatment burden
7	Tiredness	Gastrointestinal (abdominal pain, diarrhoea, flatulence)
8	High treatment burden	Impaired hearing
9	Inability to cough/clear up sputum	Sputum production (presence of, or worsening from baseline)
10	Inability to meet personal/school/work goals	Poor appetite/eating difficulties
11	Feeling unwell	Chest tightness
12	Poor exercise tolerance	Lack of energy
13	Nausea/vomiting	Not feeling well
14	Coughing	Reduced exercise tolerance
15	Reduction of usual activities	Fever
16	Nausea/vomiting	Inability of child to meet personal/school/work goals
17	Difficulty sleeping	Poor weight
18	Presence of pain	Difficulty sleeping
19	Presence and severity of wheezing	Reduced ability to concentrate/think clearly
20	Being worried about your CF	Wheeze
21	Reduced ability to concentrate/thing clearly	Presence of sweats/chills
22	Chest tightness	Reduction of usual activities
23	Coughing	Amount of school/work missed by child
24	Gastrointestinal (diarrhoea/bloating/flatulence)	Presence of headaches
25	Fever	Feeling unwell
26	Amount of school/work missed	
27	Poor weight gain or weight loss	
28	Presence of severity of chills/sweats	
29	Irritable/feeling cranky	

CF, cystic fibrosis.

poor clearance. Unlike breathing difficulty, sputum production/poor clearance is also causally related to faecal and urinary incontinence via its relationship with coughing spasms, and so may capture additional useful information compared with breathing difficulty alone.

In the general symptoms domain, fevers/night sweats and poor appetite were identified as priority outcomes because they were meaningful and important to both people with CF and carers, and were relatively independent of the other outcomes already identified.

Pain was a highly prioritised as an outcome by carers and was considered moderately important to people living with CF; this outcome featured as important across a number of domains, including the general symptoms, respiratory and gastrointestinal domains. Pain emerged

as a causally dominant outcome within the gastrointestinal domain. Other gastrointestinal symptoms (such as flatulence, diarrhoea and steatorrhoea) were also identified as priorities, as they were meaningful outcomes among both people with CF and to carers, and were relatively causally independent of gastrointestinal pain.

In the mental health domain, a person's overall feeling of amotivation/demoralisation was identified as a priority outcome, as it was itself causally influenced by a range of other important outcomes, including anxiety/worry, sadness or depression, irritability and alcohol and drug dependence.

In the function domain, treatment burden and an inability to meet personal, school or work goals were identified as priorities; these outcomes were also found
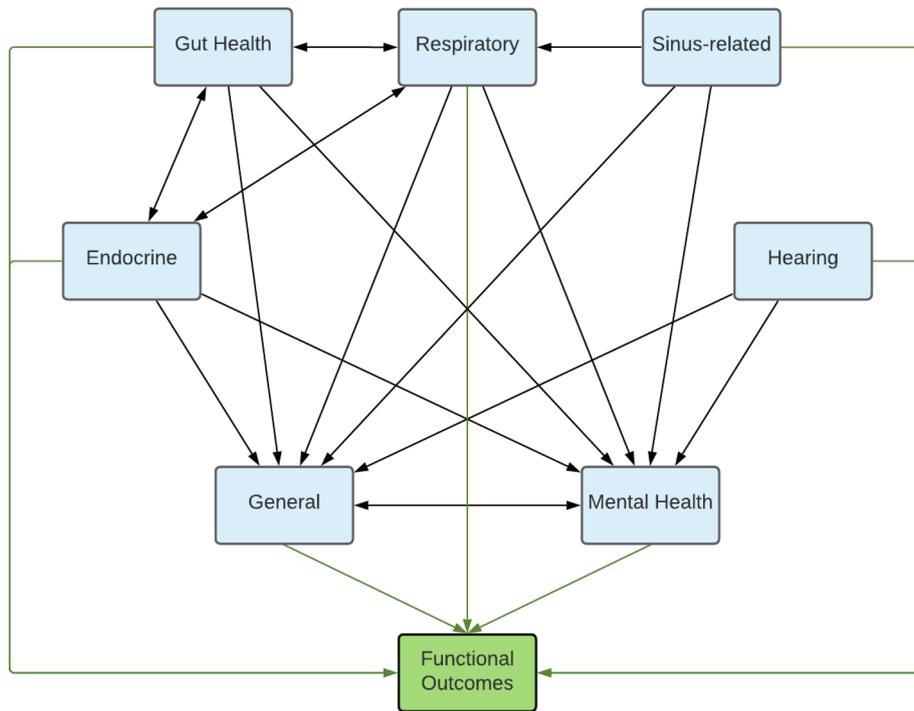
Figure 3 Causal subnetwork model demonstrating the relationships between outcome domains.

to be causally related to most other important outcomes in the functional subnetwork, either as a common cause, or as a common effect of those outcomes. While not a functional outcome per se, it is worth noting that hospitalisation was found to be important given its impact on quality of life; it was also a factor that impacted on all outcomes in the function subnetwork.
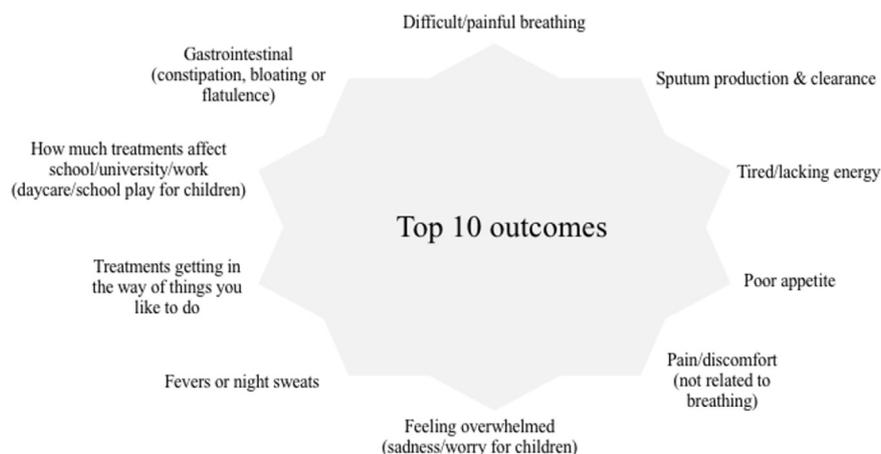
The 10 outcomes prioritised by people with CF and carers that independently map to causal disease processes are presented in figure 5.

DISCUSSION

This is the first attempt to prioritise outcomes for reporting in studies of CF pulmonary exacerbations, as nominated by people with CF and their carers. Selection of these patient-centred outcomes was achieved using a novel approach. We used causal network diagrams to select a subset of 10 meaningful outcomes that collectively capture as much of the overall outcome experience as possible. The next step will be to work with people with CF and carers to quantify the relative importance that they place on each of these outcomes. This weighting will inform the development of separate weighted outcome measurement instruments for use in children and adults with CF, as a single summary measure of the overall
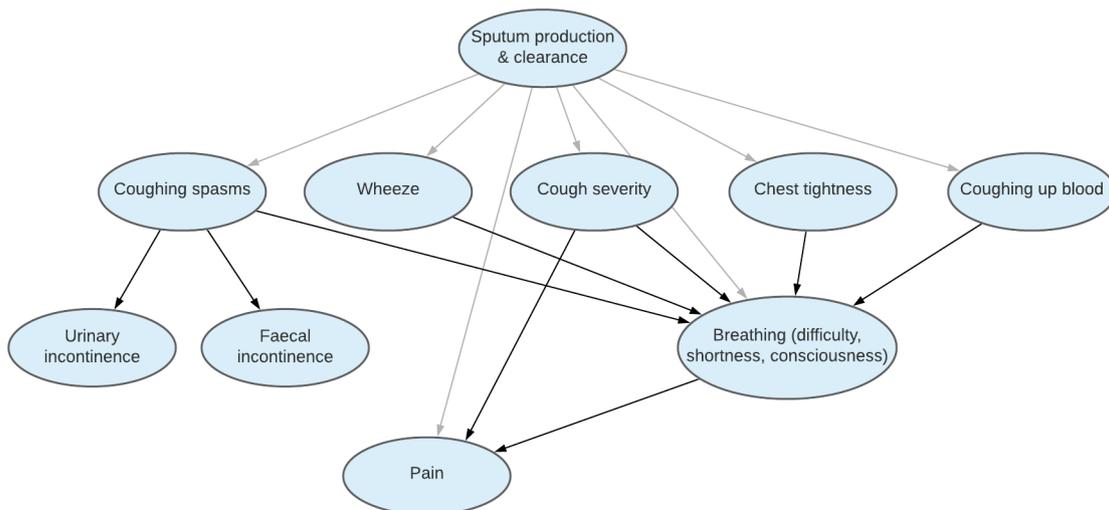
Figure 4 Respiratory subnetwork model illustrating causal relationships.

Figure 5 Ten priority outcomes identified by people with CF and carers. CF, cystic fibrosis.

outcome of pulmonary exacerbations of CF, and which could therefore be used to evaluate interventions.

Research priorities for CF from the perspective of more than 1000 consumers from 23 countries have been reported,[11] as well as an assessment of the extent to which current CF studies match these priorities.[12] However, there is not yet consensus on a COS for universal adoption when evaluating interventions in studies of pulmonary exacerbations in CF.

People with CF and carers alike prioritised breathing difficulty, excessive sputum production/poor clearance, fatigue, pain, amotivation/demoralisation, high treatment burden and inability to meet personal/school/work goals as important and meaningful. Parents/carers gave higher priority than people with CF to hearing impairment, reduced appetite and gastrointestinal symptoms; conversely people with CF gave higher priority than carers to tiredness, sleeping difficulty and 'feeling unwell'. Failure to separately derive priority outcomes for these two groups could be a limitation, but we decided on this pragmatic approach (1) because of the considerable overlap in outcomes between the groups and the critical importance of carers, in particular, for young children and (2) because the intention was to use a causal understanding of the same underlying disease process, and people with CF, carers and other domain experts all have valid expertise to contribute.

Limitations of this work included the relatively small number of participants in the consumer workshops; however, almost all outcomes identified by review of the literature were also independently identified and considered by the participants, so we are confident that we have not overlooked important outcomes. Participants were largely from Australia where people have access to universal healthcare; this may limit the generalisability of our results to populations who may have poorer access to healthcare. While we assume that the views of participants in this study represent a broader population affected by CF, the age and disease severity of individual participants with CF is likely to materially impact on the ascertainment and perceived importance of each outcome. Young

people with CF (aged <25 years old) were not well represented in either the workshops or the online surveys, and we did not try to separately ascertain priorities from children, or how their perspectives differ from that of their carer(s), which means we were unable to draw meaningful conclusions about the differences in outcome selection between different age groups. It is not entirely clear why it was difficult to recruit adolescents to this study. Possible contributing factors include the competing demands on their time or a lack of understanding about the relevance or importance of the study for them.

The framework we present here for selecting outcomes for use in clinical studies benefits from direct involvement of patients and families, and is arguably more rigorous than alternative methods which ignore how such outcomes causally relate to the underlying disease process and with each other. We expect that standardised reporting of these outcomes in clinical studies evaluating treatments for pulmonary exacerbations in CF would improve the value of those studies. We believe that this approach has the potential for broader application to the selection of outcomes in studies across various clinical problem domains, especially where the underlying pathophysiology is complex and the potential range of outcomes is broad.

Further work is required to engage children and adolescents living with CF to ascertain what health outcomes they prioritise, and how these preferences compare to those of their carer(s). It will also be necessary to quantify the relative importance of each of the outcomes identified here. A discrete choice experiment will be the focus of subsequent research; participants affected by CF will be presented with hypothetical choice sets and asked to choose between treatment options to ascertain how they value different outcomes resulting from treatment. Weighted outcomes will then be incorporated into a multiattribute utility instrument designed to capture the impact of trial interventions as a single score. This tool will require validation as an outcome assessment tool and will be compared with traditional outcomes such as FEV_1 before being considered for use more broadly.

Author affiliations
[1]Infectious Diseases, Perth Children's Hospital, Nedlands, Western Australia, Australia
[2]Infectious Diseases Implementation Research, Telethon Kids Institute, Nedlands, Western Australia, Australia
[3]School of Population Health, Curtin University Bentley Campus, Bentley, Western Australia, Australia
[4]Abilities Research Center, Department of Rehabilitation and Human Performance, Icahn School of Medicine at Mount Sinai, New York, New York, USA
[5]Department of Physiotherapy, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia
[6]Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia
[7]The Faculty of Health and Medical Sciences, The University of Western Australia, Perth, Western Australia, Australia
[8]Wal-yan Respiratory Research Centre, Telethon Kids Institute, Nedlands, Western Australia, Australia
[9]Department of Respiratory Medicine, Perth Children's Hospital, Nedlands, Western Australia, Australia
[10]Wesfarmers Centre for Vaccines and Infectious Diseases, Telethon Kids Institute, Nedlands, Western Australia, Australia
[11]Consumer advocate, Perth, Western Australia, Australia
[12]Consumer advocate, Melbourne, Victoria, Australia
[13]School of Public Health, The University of Sydney, Sydney, New South Wales, Australia
[14]Evidence Based Child Health Group, School of Medicine, University of Nottingham, Nottingham, UK
[15]School of Population Health and Preventive Medicine, Monash University, Clayton, Victoria, Australia
[16]Department of Intensive Care Medicine, St John of God Health Care, West Perth, Western Australia, Australia
[17]Bayesian Intelligence, Upwey, Victoria, Australia
[18]Menzies School of Health Research, Casuarina, Northern Territory, Australia

REFERENCES
1 Moher D, Glasziou P, Chalmers I, *et al*. Increasing value and reducing waste in biomedical research: who's listening? *The Lancet* 2016;387:1573–86.
2 Prinsen CAC, Vohra S, Rose MR, *et al*. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. *Trials* 2014;15:247.
3 Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605.
4 Lederer DJ, Bell SC, Branson RD, *et al*. Control of confounding and reporting of results in causal inference studies. guidance for authors from editors of respiratory, sleep, and critical care journals. *Ann Am Thorac Soc* 2019;16:22–8.
5 US Department of Health and Human Services CfDEaRCaCfBEaRC. *Multiple endpoints for clinical trials: guidance for industry. In: services UDoHaH, editor*, 2017.
6 Pearl J. An introduction to causal inference. *Int J Biostat* 2010;6:7.
7 McLachlan S, Dube K, Hitman GA, *et al*. Bayesian networks in healthcare: distribution by medical condition. *Artif Intell Med* 2020;107:101912.
8 Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012;31:2973–84.
9 Rowbotham NJ, Smith S, McPhee M, *et al*. EPS1.9 question CF: a James Lind alliance priority setting partnership in cystic fibrosis. *Journal of Cystic Fibrosis* 2017;16:S38-S:S38.
10 Custer RL, Scarcella JA, Stewart BR. The Modified Delphi Technique - A Rotational Modification. *J Career Tech Educ* 1999;15.
11 Rowbotham NJ, Smith S, Leighton PA, *et al*. The top 10 research priorities in cystic fibrosis developed by a partnership between people with CF and healthcare providers. *Thorax* 2018;73:388–90.
12 Kalaitzis IS, Rowbotham NJ, Smith SJ, *et al*. P251 how does the current clinical trials landscape reflect the James Lind alliance top ten research priorities for CF? *J Cyst Fibros* 2018;17.S130.