

## A Generic Rapid Evaluation Support Tool (GREST) for Clinical and Commissioning Decisions

Jack Dowie<sup>ab</sup>, Vije Rajput<sup>c</sup>, Mette Kjer Kaltoft<sup>b</sup>

<sup>a</sup>London School of Hygiene and Tropical Medicine, London, UK

<sup>b</sup>University of Southern Denmark, Odense, Denmark

<sup>c</sup>Stonydelph Health Centre, Tamworth, UK

### Abstract

*A fast and frugal generic tool can provide decision support to those making decisions about individual cases, particularly clinicians and clinical commissioners operating within the budget and time constraints of their practices. The multinational Generic Rapid Evaluation Support Tool (GREST) is a standard preference-sensitive Multi-Criteria Decision Analysis-based tool, but innovative insofar as an equity criterion is introduced as one of six. Equity impact reflects the number of population QALYs lost or gained in moving from Old (current intervention) to New (contemplated intervention). In the exemplar UK implementation Claxton's NHS Willingness to Pay per QALY is the numeraire. Any weight from 0 to 100% may be assigned to the equity criterion but its presence affirms that it is persons-as-citizens who experience any opportunity harms or benefits arising from actions within the health service commons. A fully-operational but demonstration-only version is available on open access, as proof of concept and method.*

### Keywords:

Clinical, Decision Support Systems, Health Equity

### Introduction

Clinicians and clinical commissioning groups are routinely making decisions about the use of new and/or expensive interventions for individual cases, including patients with multiple morbidities. They may be facing an increasing number of such decisions in the UK, if the recent High Court decision regarding off-label/licence drugs is upheld

(<https://www.bbc.co.uk/news/health-45588983>).

We perceive a major gap in the support available for such decisions. NICE-type evaluations can cover only a tiny minority of interventions and their remit is currently restricted to on-label use. Decision aids developed according to normative standards such as IPDA<sup>Si</sup> [1] are very limited in coverage, constitute information aids rather than decision support tools that produce an opinion, and rarely introduce budget considerations. Guideline recommendations, such as those produced by GRADE, cannot, by definition, reflect individual preferences, only group or sub-group averages, and leave the decision maker analytically unsupported in the task of processing the extensive summary of evidence at the point of decision.

Whenever a proposed NEW intervention is to be comparatively evaluated against a current OLD one within time and resource constraints, a flexible, rapid, generic and inexpensive decision support tool is needed.

The required tool must be practical and useful. It must meet the SMART criteria - Specific, Measurable, Achievable, Relevant, and Timely - to the extent each of these are reasonable in the given decision context. More simply, the decision support must be 'fast and frugal'. It must be pitched at the most appropriate point (trade-off) on the 'rigour-relevance' continuum. It will therefore be a long way from a highly analytical NICE-type evaluation in one direction, but also far removed from an expert deliberation-based guideline in the other. While endorsing long term efforts to develop normatively superior methods of linking evidence and clinical practice [2-6] or dealing with the complexities of value-based care [7-8] we have the simple but limited ambition of providing a rapid and practical method of improving on the present decision making process, whatever it is. The tool will not seek to replace or deter the development of superior tools, but reflect the belief that the *normatively best*, or even *normatively better*, may be the enemy of the *empirically better* – a point well-accepted in relation to drugs and devices, not yet for decisions. Evaluations should reflect this and use the actual process they would replace as comparator.

### Method

Multi-Criteria Decision Analysis (MCDA) can provide the basis for such fast and frugal decision support tools. Given their basis in this technique they can provide an opinion on the worth of each Option (one of the actions that can be taken) by combining its Ratings (how well it performs on relevant Criteria) with the criteria Weightings (how important each criterion is relevant to the others), making the opinion a preference-sensitive one. Introductory materials and numerous examples of the implementations of MCDA in the decision support context are available at <http://cafeannalisa.org.uk> including the short video 'Powtoons'. In the space available and to avoid duplication we leave further details of Method to the following Results section and the online tool introduced there.

In Multi-Criteria Decision Analysis-based tools, it is entirely feasible to include an equity outcome criterion and our proposed tool, implemented in the Annalisa template [9], does this. However, we emphasize that this innovation occurs in the context of a more extended, multicriterial analysis, otherwise

the tool will not be appropriate in person/citizen-centred healthcare decision making.

**Result**

The Generic Rapid Evaluation Support Tool (GREST) is presented here as proof of concept and method, based on internal testing, and as the basis for feedback in relation to future development and implementation.

The two options in GREST are OLD (the current intervention, such as usual/standard/current care) and NEW (the contemplated replacement). The six criteria in the default prototype are:

- A condition/decision-specific Biomarker (e.g. Bone Mineral Density)
- A condition/decision-specific Function(al) Index (e.g. Six Minute Walk Test)
- Option (test/treatment) Side effects
- Option (test/treatment) Burden (e.g. arising from frequency and mode of delivery of medication)
- Health-Related Quality of Life (HRQOL)
- Equity (reflecting the Harm/Foregone Benefits to Others in ‘North-East’ cases and Benefit/Foregone Harm to Others in ‘South-West’ ones – see below)

It will be noted that Life Expectancy is not a criterion in the current GREST, so if there is any effect on this from the new intervention, it would need to be discussed separately in the light of the GREST opinion. It will also be noted that monetary cost is not a separate criterion, being introduced only through the sixth, equity, criterion.

To engage with the tool and understand the method underlying it, go to <https://ale.rsyd.dk> (enter 1513 as survey ID). This version is in English and uses the EQ-5D-5L tariff for 11 countries to establish Health-Related Quality of Life (HRQOL) values used in deriving the ratings for both the HRQOL and Equity criteria. Nine tariffs are from <https://euroqol.org/eq-5d-instruments/eq-5d-about/valuation-standard-value-sets/crosswalk-index-value-calculator>[10]; others tariffs included are for Poland [11] and China [12],

This is a demo version provided as a proof of concept and method on open access. Only anonymous, non-confidential data should be entered. No responsibility is taken for data security.

The performance ratings of OLD and NEW on the first five criteria in GREST are elicited in whatever way is compatible with the clinician’s or clinical commissioning group’s practice timescale and resources. The Ratings for both OLD and NEW should be the BEANs (Best Estimates Available Now) and hence as evidence-based as is possible, and as expertise-based as is necessary, within the actual resource and time constraints of the decision makers. The person/patient is to be regarded as the expert on Treatment Burden. The derivation of the ratings for the equity criterion is explained in a separate sub-section below.

The relative importance Weighting of each criterion, including equity, is elicited on a 0-100% scale, where 0 indicates of no importance and 100% of extreme importance. The six responses are summed and percentage to give and display the set of provisional criterion weights that add to 100%. In this interactive tool they may be changed by cursor on inspection

of their graphical display. More sophisticated weight elicitation procedures, such as swing weights or discrete choice experiments, may have greater normative appeal than Visual Analog Scales, but lack either individual applicability or practicality in the typical time and resource scale contemplated.

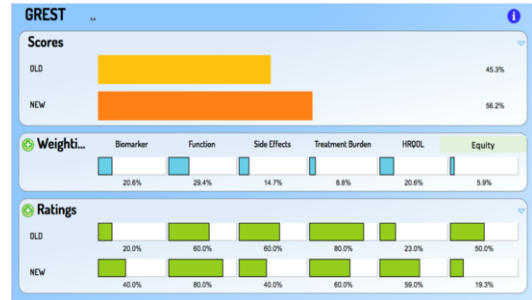


Figure 1: North-East Case with 5.9% Equity Weight

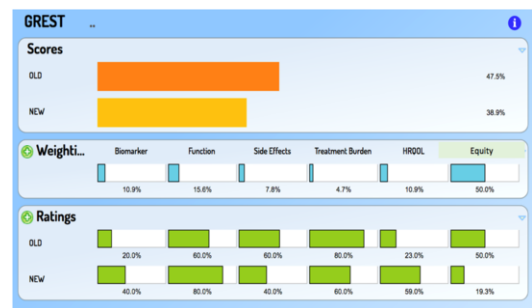


Figure 2: North-East Case with 50% Equity Weight

The GREST output is in the form of a ‘deciographic’ - a single screen showing all Ratings, Weightings, and the evaluation Scores for OLD and NEW. As in all standard MCDA applications, the Score combines the relative importance criterion Weightings with the evidence- and expertise-informed option performance Ratings, by the Expected Value algorithm.

Giving 50% weight to equity in Figure 2 (bottom) rather than the 5.9% in Figure 1 (top) flips the opinion to OLD.

The ratings in the above example are purely illustrative. We imagine NEW to be superior on Biomarker, Function, and HRQOL, but inferior on Side Effects and Treatment Burden – and, by the definition of a ‘North-East’ GREST, it produces net ‘Harm to Others’. The 50% Equity rating for OLD reflects complete uncertainty about its value. In the Ratings panel a longer bar always means better and in the Weightings panel a longer bar always means more important.

**The Equity Criterion**

Person-centred care – and value-based healthcare - is not all about the rights of the person. Persons are also citizens who have responsibilities and duties within a resource-constrained public health service. Only by applying a generic outcome measure as a criterion within a personalised multi-criterial decision analysis can we move to the coherent involvement of equity in decision making about clinical cases. In GREST, the equity outcome criterion is defined and measured as the Harms or Benefits to Others created in moving from OLD to NEW in a resource-constrained service. In measuring these,

any *condition-specific* outcome such as *cancer* mortality/morbidity, as contrasted with *all-cause* mortality/morbidity, is ruled out as equity-irrelevant.

We have chosen Health-Related Quality of Life (HRQOL) as our generic equity outcome measure, the current GREST explicitly eschewing evaluations where NEW alters life expectancy. We use EQ-5D-5L as the HRQOL metric, though the method is not tied to any specific instrument.

Where the NEW intervention is *more expensive* than the OLD, but also *more effective*, we are in the North-East quadrant of the Cost-Effectiveness plane. There will be ‘opportunity harms’ (*foregone benefits*) to others, usually anonymous and unidentifiable, as a result of substituting NEW for OLD.

Where a NEW intervention is *less expensive* than the OLD, but also *less effective*, we are in the South-West quadrant of the Cost-Effectiveness plane, the only one recognized in NICE, where lower effectiveness is disallowed as a source of improved cost-effectiveness. This restriction of cost-effectiveness to *incremental* cost-effectiveness constitutes political interference with the implementation of a neutral technique [13]. In this quadrant there will be ‘opportunity benefits’ (*foregone harms*) to others, usually anonymous and unidentifiable, as a result of substituting NEW for OLD.

In measuring the foregone benefits and harms in the UK context, we use Karl Claxton’s estimate of the NHS’s revealed willingness to pay (WTP) for a Quality-Adjusted Life Year (QALY) of £12,000. ‘The central or mean estimate of £12,936 is likely, if anything, to be an overestimate’, hence our use of £12,000. [14] (p4).

Equity impact is calculated as the number of QALYs that would be moved from Others to the recipient of the NEW intervention. So, if the NEW intervention (assuming no life expectancy effect) costs £24,000 per year compared with £1,600 for the current OLD treatment (Extra cost = £22,400), and it improves HRQOL from 0.23 to 0.59 (QALY gain = 0.36), the Equity impact (Harm to Others) created by adopting NEW is calculated as:

$$\frac{£22,400}{0.36} = £62,222 \text{ per QALY}$$

$$£62,222 / £12,000 = 5.2.$$

The generation of the 1 QALY involved in substituting NEW for OLD for a single (identifiable) person creates a QALY loss of 5.2 to anonymous others elsewhere in the service. At the population level there is a net loss of 4.2 QALYs from the shift. This loss may all be borne by one anonymous person, or be distributed as lost quality-adjusted weeks or days across small or large numbers of people. The fact is that this distribution is unknown. (The issue of ‘statistical compassion’ is raised in the final section.)

If we reverse the OLD and NEW data and move into the South-West quadrant of *decremental* cost-effectiveness, the arithmetic is the same but the effect the opposite. Substituting NEW for OLD now creates a QALY gain of 5.2 to anonymous others elsewhere in the service. At the population level there is a net gain of 4.2 QALYs.

To map the QALY movement on to the required 0-1 ratio scale for the MCDA-based tool, we take its reciprocal, subtracting this from 1 in the South-West case. Moving 5.2 QALYs becomes 0.193 in the NE case (Figures 1 and 2) and 0.807 in the SW (Figures 3 and 4).

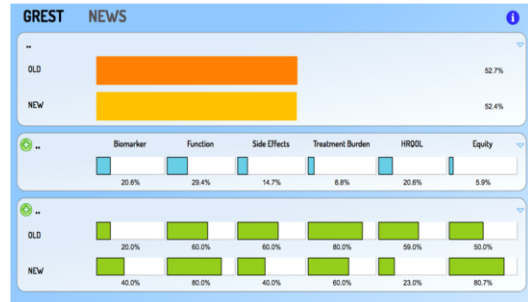


Figure 3: South-West case with 5.9% equity weight

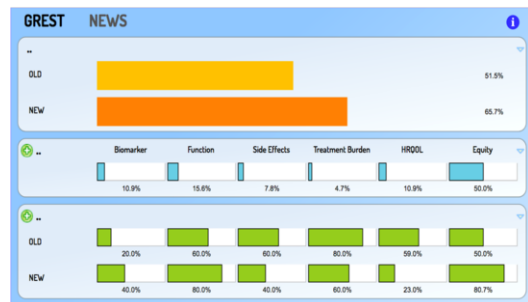


Figure 4: South-West case with 50% equity weight

The above screens result from reversing the data in the earlier example, making NEW a South-West intervention, where it produces lower HRQOL but at greatly reduced cost, in contrast to the North-East situation where it produces the same amount of higher HRQOL but at greatly increased cost.

Giving about 6% weight to equity in Figure 3 (top) is insufficient to shift the verdict in favour of NEW, but assigning 50% weight in Figure 4 (bottom) clearly does so.

What can we conclude about the change in equity in these two situations? The most reasonable assumption is that the anonymous others who are either gainers or losers from NEW will be randomly distributed. Under this assumption and assuming an equity rating for OLD of 50% (reflecting complete uncertainty), the net effect of NEW will be to *increase inequity* if it is a North-East intervention (replacing the less costly and less effective OLD) and *reduce inequity* if it is a South-West one (where it is less costly and less effective than OLD).

The magnitude of the effects will be a function of the numbers in the particular clinical case. But we can note that at a service level, the potentially very large numbers of opportunity harms from NE interventions may add up to a massive increase in inequity while the failure to consider SW interventions can add up to a huge potential failure to reduce inequity. The same applies in the individual screening context.

## Discussion

Claxton makes our basic point in relation to NICE decisions about new drugs, where the threshold being used is far above the average observed WTP for a QALY of £12,000, but it is equally pertinent at the clinical/clinical commissioning level.

“The evidence suggests that more harm than good is being done, but it is the unidentified and unrepresented NHS patients who bear the true (health) opportunity costs. Although finding reasons to approve new drugs is undoubtedly politically expedient, this cannot be ethically literate, because the interests of NHS patients, whether they are identifiable or not, are just as real and equally deserving of the type of care and compassion that can be offered by a collectively funded health care system. It is to be hoped that NICE will begin to place the unidentified NHS patients who bear the real opportunity costs at the heart of its deliberative process.” [15] (p6).

The GREST tool does not mandate any amount of concern with the impact on others, but it does mandate a statement of the weight to be assigned to it, in the 0 to 100% range, in generating the opinion of the decision support tool. As persons-as-citizens, we incur both a right and responsibility to be informed of the impact on others of our use of the health service commons in self-producing and co-creating health [16].

Objections to introducing a societal element into clinical decisions can be expected from clinicians as well as patients, but rejected on the ground that equity is a social matter for all citizens, and it is not for a profession to decide as they should to be able to ‘protect their patients’ from concern with it. In any case, a healthcare budget-holder cannot escape this wider responsibility for all those within their aegis.

Treating any individual person as a means to achieve some population policy end— to reduce the incidence of this, increase the uptake of that — has to be rejected even if it is in their *perceived* individual best interests. (Infectious diseases are an exception where the law rightly takes away the right of an individual to directly jeopardise the health of others.) But neither is it ethical for the person as citizen to be left in — or allowed to opt into — complete ignorance as to the impact on others, in whatever direction. An empirical resolution of the clash between individual and community [17] is brought no nearer by simple repeated noting of its existence, or by attempts to solve it solely at the collective level, leaving the individual uninformed in any direct way.

The person is not only a citizen, but usually a member of a family or other group of significant others. Nothing said here is meant to imply that the person will not wish to take the feelings and consequences for these near others into their decision making. Nor to imply that the benefits of a ‘family focus’ in provider care should be ignored or underestimated. We merely note that if relatives are driving demands to ‘do everything possible’ against the perceived implicit, or explicitly expressed, wishes of the patient, this transparent support will be helpful in empowering the patient.

Both the ethical requirements of ‘patient-centred’ care and the legal requirements of ‘reasonable patient’ care suggest that the introduction of ‘equity’ into clinical practice guidelines, recently advocated [18-21], is not an appropriate way to go. The GRADE subgroup propose that guidelines panels recommend that clinical practitioners ‘consider’ the ways in which an individual patient may be affected through being a member of a ‘disadvantaged’ group. However, this seems to be ethically suspect, insofar as it could potentially distort the personalized option ratings for, and personal criterion weightings of, the specific individual. The case for a personalized decision analytic approach using individualised ratings and weightings is well made by Wasfy and colleagues [22] and not satisfactorily refuted by Lightner [23]. GREST is on offer whether or not guidelines panels were to abdicate

from making a recommendation in the face of uncertainty [24].

Everyone seeking to participate in the equity debate is confronted by the sociopsychological phenomenon, emphasised by Williams and Cookson, among others, of the difficulty individuals have in relating to the thousands of anonymous others who have equal rights in the health service commons.

“For many people [the] notion of ‘statistical compassion’ seems to create both intellectual and psychological difficulties. It is as if personal empathy with one or two individuals is possible, but, paradoxically, if many individuals are involved, this capacity to empathise diminishes. This difference between focusing on groups and focusing on individuals also distinguishes economists (and managers) from clinicians and others dealing with people at an individual level. The latter often claim that they are under an ethical duty to do everything possible for the person in front of them no matter what the consequences might be for everybody else. If this assertion is taken at its face value, it would imply that clinicians should ignore their responsibilities for the welfare of their other patients except when that patient is in front of them. It seems most unlikely that any clinician would actually behave in that way, so perhaps the statement should not be taken at its face value, but regarded instead as part of the rhetoric of medical practice, designed to bolster the doctor-patient relationship. But whatever may be the role of such statements, it is clear that in a public policy context, where distributive justice is an explicit objective, it is clearly not ethical for a clinician to ignore the consequences of his or her actions concerning the treatment of one patient for the health of other patients for whom the system is also responsible.” [25] (p1866).

## Conclusion

The individual can only self-produce their health, and co-create it with healthcare professionals and significant others, within the wider environment and socioeconomic constraints in which they live and work. Most of the recent advances in key health indicators (e.g. life expectancy) are attributable to sectors other than healthcare, so that the creation of health and health equity needs to be approached within a much wider and comprehensive framework [26]. Equitably, cost-effective, cross-sectoral public health policies are an essential complement to any changes in clinical decision making [27].

However, the transparent introduction of equity considerations into decision making in relation to individuals can play an important role in affirming the social nature of a public health service. A public health service is not a private health service. It is ultimately the person-as-citizen who is experiencing any opportunity harm or benefit that arises from an action within the health service commons [28]. All health service decision makers should therefore be showing the ‘statistical compassion’ appropriate to their level of responsibility. GREST provides a simple way to allow the opportunity implications of a decision to be highlighted, but then set at whatever level is desired — including zero.

**Author Contributions:** JD developed the generic GREST tool in response to a challenging clinical decision, incorporating the insights and suggestions of VR and MKK. He wrote the paper, incorporating their observations, revisions and additions.

**Funding:** This research received no external funding. The software used for GREST was installed at <https://ale.rsyd.dk> as part of a nationally funded project to develop decision support tools. (Danish National Board of Health/SATS J.nr. 1-1010/116/27.)

**Conflicts of Interest:** Jack Dowie has a financial interest in the Annalisa software implementation of MCDA, but does not benefit from its non-commercial use.

## References

- [1] G. Elwyn, A.M. O'Connor, C. Bennett, R.G. Newcombe, M. Politi, M-A. Durand, et al., Assessing the quality of decision support technologies using the International Patient Decision Aid Standards instrument (IPDASI), *PLoS One* **4** (2009), e4705.
- [2] C.F. Manski, Reasonable patient care under uncertainty, *Health Econ* **27** (2018), 1397–1421.
- [3] A. Basu, Manski's views on patient care under uncertainty: Comment. *Health Econ* **27** (2018), 1422–4.
- [4] K. Claxton, Positive tails and normative dogs: Comment on Manski 2018, *Health Econ* **27** (2018), 1425–7.
- [5] E. McIntosh, Decentralized decision making through adaptive minimax regret — Complex yet intuitively appealing: Comment on Manski 2018, *Health Econ* **27** (2018), 1428–30.
- [6] C.F. Manski, Response to commentaries on "Reasonable patient care under uncertainty", *Health Econ* **27** (2018), 1431–4.
- [7] L.P. Garrison, M.V. Pauly, R.J. Willke, and P.J. Neumann, An Overview of Value, Perspective, and Decision Context—A Health Economics Approach: An ISPOR Special Task Force Report 2, *Value Heal* **21**, 2018, 124–30.
- [8] P.J. Neumann, R.J. Willke, and L.P. Garrison, A Health Economics Approach to US Value Assessment Frameworks—Introduction: An ISPOR Special Task Force Report 1, *Value Heal* **21**, 2018, 119–23.
- [9] J. Dowie, M.K. Kalsoft, G. Salkeld, and M. Cunich, Towards generic online multicriteria decision support in patient-centred health care, *Health Expect* **18** (2013), 689–702.
- [10] B. van Hout, M.F. Janssen, Y.S. Feng, T. Kohlmann, J. Busschbach, D. Golicki, et al., Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets, *Value in Health* **15** (2012), 708–715.
- [11] D. Golicki, M. Niewada, B. van Hout, M.F. Janssen, and A.S. Pickard, Interim EQ-5D-5L Value Set for Poland: First Crosswalk Value Set in Central and Eastern Europe, *Value Heal Reg Issues* **4** (2014), 19–23.
- [12] N. Luo, G. Liu, M. Li, H. Guan, X. Jin, and K. Rand-Hendriksen, Estimating an EQ-5D-5L Value Set for China, *Value Heal* **20** (2017), 662–9.
- [13] J. Dowie, M.K. Kalsoft, J.B. Nielsen, and G. Salkeld, *Caveat emptor* NICE: biased use of cost-effectiveness is inefficient and inequitable. *F1000Res* **4** (2015), 1078.
- [14] K. Claxton, S. Martin, M. Soares, N. Rice, E. Spackman, S. Hinde, et al., Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold, *Health Technol Assess* **19** (2015), 1–504.
- [15] K. Claxton, M. Sculpher, S. Palmer, and A.J. Culyer, Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? *Health Econ* **7** (2015), 1–7.
- [16] J. Dowie and M.K. Kalsoft, The future of health is self-production and co-creation based on apomediative decision support, *Med Sci* **6** (2018), 66.
- [17] A. Granata and A. Hillman, Competing practice guidelines: using cost-effectiveness analysis to make optimal decisions, *Ann Intern Med* **128** (1998), 56–63.
- [18] V.A. Welch, E.A. Akl, G. Guyatt, K. Pottie, J. Eslava-Schmalbach, M.T. Ansari, et al. GRADE equity guidelines 1: considering health equity in GRADE guideline development: introduction and rationale, *J Clin Epidemiol* **90** (2017), 59–67.
- [19] E.A. Akl, V. Welch, K. Pottie, J. Eslava-Schmalbach, A. Darzi, I. Sola, et al. GRADE equity guidelines 2: considering health equity in GRADE guideline development: equity extension of the guideline development checklist, *J Clin Epidemiol* **90** (2017), 68–75.
- [20] V.A. Welch, E.A. Akl, K. Pottie, M.T. Ansari, M. Briel, R. Christensen, et al., GRADE equity guidelines 3: considering health equity in GRADE guideline development: rating the certainty of synthesized evidence. *J Clin Epidemiol* **90** (2017), 76–83.
- [21] K. Pottie, V. Welch, R. Morton, E.A. Akl, J.H. Eslava-Schmalbach, V. Katikireddi, et al., GRADE equity guidelines 4: considering health equity in GRADE guideline development: evidence to decision process, *J Clin Epidemiol* **90** (2017), 84–91.
- [22] J.H. Wasfy, K. Armstrong, C.E. Milford, and T.M. Sundt, Bicuspid aortic disease and decision making under uncertainty - The limitations of clinical guidelines, *Int J Cardiol.* **181** (2015), 169–71.
- [23] D.J. Lightner, Re: Wasfy et al., Guidelines versus decision analysis is the wrong conversation, *Int J Cardiol* **186** (2015), 60.
- [24] M. Bretthauer and M. Kalager, When no guideline recommendation is the best recommendation, *Lancet* **392** (2018), 898–9.
- [25] A. Williams and R. Cookson, Equity in health, *Handbook of Health Economics* **1B** (2000), 1863–910.
- [26] J. Bircher and S. Kuruvilla, Defining health by addressing individual, social, and environmental determinants: New opportunities for health care and public health, *J Public Health Policy* **35** (2014), 363–86.
- [27] G. Mooney, *The Health of Nations: Towards a New Political Economy*, Zed Books, London, 2012.
- [28] K. Hassanally, Overgrazing in general practice: the new *Tragedy of the Commons*, *Br J Gen Pract* **65** (2015), 81.

## Address for correspondence

Jack Dowie. Email: [jack.dowie@lshtm.ac.uk](mailto:jack.dowie@lshtm.ac.uk).

## Queries about survey on <https://ale.rsyd.dk>

Mette Kjer Kalsoft. Email: [mkaltoft@health.sdu.dk](mailto:mkaltoft@health.sdu.dk)