Risk classifications interfere with preference-sensitive, decision support

Mette Kjer KALTOFT[a], Jack DOWIE[ab1],

[a]*University of Southern Denmark*
[ab]*London School of Hygiene and Tropical Medicine*

Abstract. Many guidelines for prevention and treatment still locate persons in risk classes (e.g. low, moderate, high) on the basis of thresholds placed on a continuous metric for a single criterion (e.g. risk of developing x). These 'traffic light' signals can lead to inferior decisions through their mono-criterial focus and lack of preference-sensitivity to the multiple criteria relevant to the person. It is arguably unethical to communicate to someone that they are at low, moderate, or high risk of x solely on the basis of the unpublished and often unknown preferences of the group that has set the classification thresholds. Any prior classification and labelling will interfere with the individual's balanced processing of information on the performance of all treatment options on their multiple relevant criteria - including treatment side effects and burdens as well as main benefit - and jeopardise meeting the requirements for fully informed and preference-based consent to any subsequent action. Personalised decision support tools based on Multi-Criteria Decision Analysis can help fulfil these objectives, with apomediative (at home) e-decision support especially appealing because of its empowering and resource-saving potential. The individual's absolute risk score is required in these tools since any threshold-based risk classification will interfere with the coherence of the analysis across the multiple criteria.

Keywords: risk classification, guidelines, Multi-Criteria Decision Analysis, decision support, preferences

Introduction

Guidelines for the prevention and treatment of many conditions locate persons in risk classes (stereotypically the 'traffic lights' of low, moderate, and high) on the basis of cut-offs placed on a continuous scale. Sometime the original source of these thresholds is lost in time and cannot now be identified; they have simply become normalised and are used because data based on them exist and can be analysed.

Many thresholds have no discernible analytical basis and are presented as the consensus of a group of experts, usually exclusively or mainly clinicians from the topic area. For example, the NICE guidelines for breast cancer triage the chance of a 40-50

[1] Corresponding author: jack.dowie@lshtm.ac.uk

year old developing breast cancer into >3% = 'near population level risk'; 3%-8% = 'moderate risk'; >8% = 'high risk' [1]. No explanation is published for adopting these particular thresholds, though frequently NICE refers to its thresholds as resulting from the Guidance Development Group's 'taking various considerations into account'. In prostate cancer, total PSA levels of 4.0 ng/ml and 10.0 ng/ml are typically used as triaging thresholds, again without any explicit analytical basis [2]. Some thresholds have a purely statistical basis, such as parameter values on a single-criterion distribution. Bone fragility guidelines invariably follow this practice, natural since osteoporosis is statistically-defined. A Bone Mineral Density of -2.5 standard deviations below the mean of a young US female constitutes osteoporosis. Other thresholds are based in whole or part on single-criterion cost-effectiveness analysis. For example, the National Osteoporosis Guidelines Group (NOGG) guidelines (https://www.sheffield.ac.uk/NOGG/downloads.html) use Frax-based age-specific thresholds derived from cost-effectiveness analyses [3].

The number of guidelines which eschew formal risk classification and promote the use of the individual's unclassified absolute risk is increasing. The 2015 update of the Joint British Societies guidelines on primary prevention of cardiovascular disease [4] is notable in its replacement of the traffic lights of 2005 [5] with the 'continuum of risk' concept. However, they still have difficulty abstaining from referring to patients as 'high risk' 'elevated risk' or 'increased risk' and JBS3 still refers to an intervention threshold. The provenance of these thresholds is not, however, the focus here. It is their use in person-centred care, including and especially their communication to the person. Their use in research and population-level management is an entirely separate question.

The heart disease example we draw on later relates to cholesterol and statins:

> *LDL (low-density lipoprotein cholesterol, also called "bad" cholesterol) … can increase your chances of getting heart disease…The lower your LDL cholesterol number, the lower your risk. If your LDL is 190 or more, it is considered very high. Your doctor will most likely recommend a statin in addition to making healthy lifestyle choices…You may also need to take a statin even though your LDL level is lower than 190. After figuring your 10-year risk, your doctor will recommend a percentage by which you should try to lower your LDL level through diet, exercise, and medication if necessary.* [https://www.webmd.com/cholesterol-management/guide/understanding-numbers#1]

In summary, looking across conditions, we can find no *analytical* basis for any particular thresholds (e.g. 10%, 15%, 20%) and resulting risk characterisations other than population-level effectiveness and cost-effectiveness analyses. Evidence-based in any other sense, most are not. While possibly appropriate at the policy level, this makes them inappropriate in person-centred care or clinical practice guidelines.

'Traffic light' risk classifications can cause considerable harm by their mono-criterial focus and lack of preference-sensitivity (completely appropriate as preference-insensitivity is on the road). The communication of risk class to the individual person

- interferes with their decisional autonomy, and the possibility of unbiased personalised, preference-sensitive, multi-criterial decision making, by pre-emptive psychological framing
- hinders the giving and obtaining of the person's legally required informed and preference-based consent to any provider action, since the required unbiased personalised assessment of all harms and benefits has not been undertaken.

To give their informed and preference-based consent to any test or treatment, the person must be informed about the harms and benefits of all the relevant options, with the magnitudes of those harms and benefits being assessed on the basis of their personal importance weights at or near the point of decision. While this requirement is rarely fully met today (except in surgery) it will be a prominent feature of the future we envisage [6]. A key implication is that it is not acceptable to focus on the single outcome proposed as the main criterion, e.g. CVD in the heart case, fracture in the bone health case, Cancer in oncology. The decision process must address the other considerations that matter to the person - equally as seriously and equally as analytically. When combined with the requirement for the individual to be able to weight those criteria explicitly and transparently, one is driven towards some form of multi-criteria analysis personalised decision support tool [7]. The use of a mono-criterion, threshold-based guideline, based on average patients, becomes suspect.

Among the many obstacles to the provision of absolute risks are mistaken provider and professional beliefs that risk classifications aid risk communication by simplification. But traffic light risk communication invariably goes beyond straightforward simplification of the risk assessment, to imply that a particular action – stratified risk management or mitigation - should follow. Since any decision requires value-based preference judgments to be combined with risk assessment, communicating a risk class based on a mono-criterial risk assessment will inevitably contaminate the multi-criterial preference elicitation required for risk management.

It is well known that labelling, such as being at 'high risk' can have stigmatising and/or disabling effects. But what if the person actually seeks a definitive label to eliminate unwanted and troubling uncertainty about their health? Unfortunately, in many situations, providing a diagnosis, as supposed to a probabilistic differential diagnosis, represents ethically suspect mis-information. The prognoses will almost always be uncertain, and the person is therefore presented with a confusing mix of an apparently certain diagnosis and set of uncertain prognoses under alternative therapeutic options. The focus should be on the latter in person-centred care.

Methods

Thresholds are, by definition, based on particular *value judgments,* so that clinicians following threshold-based guidelines are unwittingly imposing the *preferences* of some group at the time they were set or confirmed. The implied trade-offs between the screening/test and treatment errors at guideline thresholds are most unlikely to match those elicited from any individual person. For informed and preference-based decision making and consent, the person requires the best estimates available for the sensitivity,

specificity, and population-relevant Predictive Values Positive and Negative *at their specific result*. In principle, they require them for All-cause mortality, for Condition-specific mortality, and for Condition-specific development.

Since providing all this data to the person is likely to be more overpowering than empowering, the good news is that thresholds and labels are no longer needed. The person can receive – or should be able to receive – personalised decision support based on their specific individual absolute result. A personalised decision support tool (PDST) based on Multi-Criteria Decision Analysis (MCDA) will hold the above data for all possible test results and can produce scores for all the management options - without provider censoring - given a specific test result.

The type of MCDA most compatible with ethical person-centred decision making and most able to ensure informed and preference-based consent is the value-based, compensatory model. This takes the form of a 'weighted-sum' calculation, which multiplies the personalised numerical ratings for the performance of each option on each criterion by the relative weight assigned to the criterion by the person, and then sums these weighted scores to get an overall preference-sensitive score for each option. The performance ratings for all options on all criteria must be on the same continuous 0-1 (0 to 100%) scale and be personalised to the absolute risk of the individual concerned. Any segmented classification of an absolute risk for any criterion will undermine cross-criterial comparability and hence the coherence of the analysis.

Inter-mediative Personalised Decision Support Tools (PDSTs) - often called Shared Decision Making Support Tools or Patient Decision Aids - are designed to *help the clinician and patient decide together, in their encounter, what is best for the patient. Apomediative* PDSTs are a 'direct-to-citizen' e-health resource designed to *help the person decide what is best for themselves*, including whether to seek a health service consultation and/or to prepare for, and engage in, a hybrid apo-intermediative consultation [8]. These apomediative tools are analogous to the comparison websites for goods and services now proliferating on the internet, but with individual criterion weighting. Insofar as they contain an uncensored set of options and commit to the personalised assessment of benefits and harms of all these, apomediative support tools help ensure that the key requirements of self-produced health are met, along with legally informed and preference-based consent to any subsequent provider action. The clinician will very often become involved in the making of the supported decision subsequently, not only because they have greater relevant knowledge (e.g. ability to identify some contraindications), but also because only they have the power to order a test, prescribe a medication or refer for, or undertake, surgery.

Result

A brief illustration of an apomediative PDST is provided here for the statin decision: Should I go, or not go, to my general practitioner to discuss taking statins? To engage with the tool, go to https://goo.gl/H7P51r. Figure 1 is an illustrative screen capture for a 60 year old male smoker with a systolic blood pressure of 160 and a total cholesterol

of 5. This example is hypothetical and has no empirical claims to be a properly developed and validated decision support tool.

Figure 1. Final screen capture for Statin decision, with pre-entered and supplied Ratings and Weightings

The PDST involves the person: completing an online instrument to obtain an estimate of their personalised risks of All-Cause and Cardiovascular Mortality in the next ten years; self-assessing their blood pressure and total cholesterol level, which are the two inputs required, along with age, sex and smoking status, to complete the EuroSCORE-based instrument; self-rating the treatment burden of statins (in respect of burden ratings there may also be inputs from significant others, such as family and other carers, or caregivers if in a cared-for facility); and assigning relative importance weights to the four criteria (10 year mortalities, statin side effects and statin burden).

The PDST includes All-Cause Mortality as well as CVD mortality, because the authors of the underlying paper rightly stress that exclusive use of a condition-specific cause of mortality (such as CVD), gives the misleading impression that the person is immortal if they do not die from that condition [9]. Most published papers which report the effect of alternative options on multiple criteria can be the basis of an MCDA-based PDST, making them an excellent vehicle for the translation of research findings into 'bedside' decision making. (Numerous examples are provided at http//:www.cafeannalisa.org.uk.)

Discussion

It should be noted that person-centred care does not mean that the person simply gets what is optimal for them in a personalised decision support tool. Even if they agree with the *opinion* of the tool, the person is also a citizen – one person among many other

persons – and cost-effectiveness is therefore a valid principle. However, it must not be disguised within risk characterisations, as it is when individuals are said to be at 'high risk' only when it is cost-effective to treat them (e.g. in relation to bone fragility). The person is respected as a person and as a citizen only when they are given the honest - and justifiable - reason for any restriction on option availability.
The aim of this paper is strictly limited. It seeks only to provide proof of method that online decision support tools based on the technique of Multi-Criteria Decision Analysis (MCDA) can meet both the ethical requirements for person-centred care and the legal requirements for consent. And to argue that this can and should be done without reference to thresholds or classifications imposed on continuous risk measures for a single criterion.

Conclusion

Risk classifications may be helpful, even essential, to researchers developing the probabilistic prognoses for treatment options that are required inputs for an MCDA-based PDST. However, these labels are not needed in person-centred care and mono-criterion-based risk classification labelling may do more harm than good.

Acknowledgements

We are grateful to Jesper Bo Nielsen for comments on an earlier version.

Funding

The software used was installed at https:// ale.rsyd.dk as part of a nationally funded project to develop decision support tools. SATS J.nr. 1-1010/116/27

Conflict of Interest

Jack Dowie has a financial interest in Annalisa but does not benefit from its non-commercial use.

References

[1] NICE, Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer, *Clinical guideline* 164, Last updated: March 2017, p37.

[2] N. Mottet, J. Bellmunt, M. Bolla, et al., Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent, *European Urology* **71** (2017) 618–629..

[3] J.A.Kanis, E.V. McCloskey, H. Johansson, O. Strom, F. Borgstrom, A. Oden, et al. Case finding for the management of osteoporosis with FRAX® - Assessment and intervention thresholds for the UK. *Osteoporos International* **19** (2008), 395–408.

[4] JBS3 Board, Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3), *Heart* **100** (2014), Supplement 2.

[5] JBS2 Board, Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice (JBS2), *Heart* **91** (2005), Supplement 5.

[6] J. Dowie, M.K. Kaltoft, The future of health is self-production and co-creation based on apomediative decision support, *Medical Sci*ences **6** (2018), 66.

[7] J. Dowie, M.K. Kaltoft, G. Salkeld, M. Cunich, Towards generic online multicriteria decision support in patient-centred health care, *Health Expectations* **18** (2013), 689–702.

[8] G. Eysenbach, From intermediation to disintermediation and apomediation: new models for consumers to access and assess the credibility of health information in the age of Web2.0, *Studies in Health Technology and Informatics* **129** (2007), 162–6.

[9] H. Støvring, C. G. Harmsen, T. Wisløff, et al., A competing risk approach for the European Heart SCORE model based on cause-specific and all-cause mortality, *European Journal of Preventive Cardiology* **20** (2013) 827–836.