Triage, consent and trusting black boxes

Kenneth Boyd

The coronavirus pandemic has brought to public attention a variety of questions long debated in medical ethics, but now given both added urgency and wider publicity. Among these is triage, with its origins in deciding which individual lives are to be saved on a battlefield, but now also concerned with the allocation of scarce resources more generally. On the historical battlefield, decisions about whom to treat first – neither those who would survive without treatment, nor those who would not survive even with treatment, but those who needed treatment to survive – was facilitated by military discipline and the limited effectiveness of treatments available. In the allocation of scarce resources today, by contrast, such decisions are subject to intense public and political scrutiny, and the range of effective treatments available has immeasurably diminished the proportion of 'those who would not survive even with treatment'. If triage decisions are to be made, they now need to be justified in the arena of public opinion by moral arguments which are also politically persuasive.

A number of different aspects of what is required for this endeavour are examined in the first five contributions to this issue of the Journal. In '*Should age matter in COVID-19 triage? A deliberative study*'[1], Kuylen and colleagues report on a deliberative study of public views in the UK, in which participants 'generally accepted the need for triage but strongly rejected 'fair innings' and 'life projects' principles as justifications for age-based allocation,…preferring to maximise the number of lives rather than life years saved'; and concerned that in any resolution 'utilitarian considerations of efficiency should be tempered with a concern for equality and vulnerability'.

A similar concern to temper utilitarian considerations, in this case with an Aristotelian view of the common good as 'the good life for each and every member of the community' is expressed in '*Public health decisions in the COVID-19 pandemic require more than 'follow the science'*' by de Campos-Rudinsky and Undurraga.[2] Public health decisions, they argue, 'always involve layers of complexity, coupled with uncertainty': 'the implication of the incommensurability of basic human goods… is that when tensions

between them arise (such as happened during this pandemic, when preservation of health required the adaptation of how we experience work, education, leisure, family and friendships)), the solution cannot be readily determined by a simple balancing test'. 'Good decision-making in public health policy' they conclude. 'does depend on the availability of reliable data and rigorous analyses, but depends above all on sound ethical reasoning that ascribes value and normative judgement to empirical facts.'

Triage decisions actually made during the pandemic are the subject of '*National health system cuts and triage decisions during the COVID-19 pandemic in Italy and Spain: ethical implications*' by Faggioni and colleagues.[3] Analysing 'the most important documents establishing the criteria for the treatment and exclusion of COVID-19 patients, especially in regard to the giving of respiratory support, in Italy and Spain', they discover 'a tension that stems from limited healthcare resources which are insufficient to save lives that, under normal conditions, could have been saved, or at least could have received the best possible treatment'. In response, they 'set forth a series of concrete ethical proposals with which to face the successive waves of COVID-19 infection, as well as other future pandemics': these include the duty of health authorities 'to plan for foreseeable ethical challenges during a health emergency', and the duty of 'public organisms at the national level, such as national committees on ethics… to prepare the protocols for care and treatment that would help physicians and healthcare workers to manage the predictable uncertainty and distress in healthcare emergencies'.

Turning to a currently pressing international aspect of resource allocation, Jecker and colleagues, in '*Vaccine ethics: an ethical framework for global distribution of COVID-19 vaccines*'[4] marshal an impressive amount of empirical research and ethical theory to argue that 'in order to accelerate development and fair, efficient vaccine allocation…vaccines should be distributed globally, with priority to frontline and essential workers worldwide': 'ethical values to guide vaccine distribution', they conclude, should 'highlight values of helping the neediest,

reducing health disparities, saving lives and keeping society functioning'.

A further important resource often found to be all too scarce during the pandemic was personal protective equipment (PPE). In '*Balancing health worker well-being and duty to care: an ethical approach to staff safety in COVID-19 and beyond*'[5], McDougall and colleagues 'articulate some of the specific ethical challenges around PPE currently being faced by front-line clinicians, and develop an approach to staff safety that involves balancing duty to care and personal well-being'. This includes 'a five-step structured…decision-making framework that facilitates 'ethical reflection and/ or decision-making that is systematic, specific and transparent' and 'guides the decision maker to characterise the degree of risk to staff, articulate feasible options for staff protection in that specific setting and identify the option that ensures any decrease in patient care is proportionate to the increase in staff well-being'.

Because of the pandemic and the fear of health services being overwhelmed by it, research on and treatment of other conditions, no less serious for the individual patient, have lacked resources which urgently require to be restored. Issues in medical ethics not directly related to COVID-19 equally call for renewed attention, not least because analysis of ethical questions raised by the pandemic largely relies on intellectual tools forged in earlier debates on other subjects. Three papers in this issue of the Journal return to subjects often discussed in medical ethics, but with fresh thinking on these, while a fourth examines a question which for many may be genuinely new.

The role and functioning of research ethics committees (RECs) was one of the earliest concerns of twentieth century medical ethics and as these committees grew both in number and in the complexity of their deliberations, they have continued to receive ethical attention. In '*Process of risk assessment by research ethics committees: foundations, shortcomings and open questions*'[6] Rudra observes that 'there is currently no uniform and solid theoretical approach to risk assessment by RECs' and in response develops a detailed 'concept of aggregate risk definition' designed to 'strengthen the coherence of REC

decisions and therefore the trust between researchers and the institution of the REC as such'.

'*Imperfect by design: the problematic ethics of surgical training*'[7] by Das, again addresses a familiar but difficult ethical question: 'How do we ethically validate the current training model for surgeons, in which trainees are often given operative duties that could likely be better handled by a staff physician?' Admitting that the 'deontological responsibilities of individual surgeons are incommensurable with the fundamentally utilitarian nature of the medical system' the author argues that surgeons 'as individuals must be willing to accept that they are knowingly foregoing optimal patient care on a small scale, and navigate the trade-offs which exist at the interface of two (possibly irreconcilable) philosophical system'.

One of the most familiar of all subjects in medical ethics, that of consent, is discussed by Giordano and colleagues in '*Gender dysphoria in adolescents: can adolescents or parents give valid consent to puberty blockers?*'[8] The occasion for this discussion is a recent English judgement suggesting 'that adolescents cannot give valid consent to treatment that temporarily suspends puberty' - a claim which appears to contradict what hitherto was generally considered settled law on adolescent consent to medical treatment. The authors, while not commenting on the specific case in question, carefully examine 'four reasons why consent may be deemed invalid' in cases of this kind: 'the decision is too complex, the decision-makers are too emotionally involved, the decision-makers are on a 'conveyor belt and 'the possibility of detransitioning'.

They argue that 'none of these stand up to scrutiny' and conclude that 'accepting these claims at face value could have serious negative implications, not just for gender diverse youth, but for many other minors and families and in a much broader range of healthcare settings.'

While much has been written on whether patients can trust their doctors, whether doctors can trust their computers has been until recently a less familiar question in medical ethics. This month's Feature Article, '*Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI*'[9] by Durán and Jongsma, together with four critical *Commentaries*, addresses this question with specific reference to the use in medicine of 'black box' algorithms, that is, algorithms whose 'computational processes…do not follow well understood rules' and are 'methodologically opaque to humans'. In order to trust such algorithms, the authors argue, doctors do not necessarily need to understand their computational processes, provided their reliability is supported by 'computational reliabilism', evidence, that is, that the algorithm is 'a reliable process…that yields, most of the time, trustworthy results'. On the other hand, even if the results are trustworthy, the authors warn, that is not sufficient to justify doctors in acting on them: 'clinical findings and evidence need to be interpreted and contextualised, regardless of the methods used for analysis (ie, opaque or not), in order to determine how these should be acted on in clinical practice…even if recommendations provided by the medical AI system are trusted because the algorithm itself is reliable, these should not be followed blindly

without further assessment. Instead, we must keep humans in the loop of decision making by algorithms.'

REFERENCES

1 Kuylen MNI, Kim SY, Ruck Keene A, *et al*. Should age matter in COVID-19 triage? A deliberative study. *J Med Ethics* 2021;75:291–5.
2 de Campos-Rudinsky TC, Undurraga E. Public health decisions in the COVID-19 pandemic require more than 'follow the science'. *J Med Ethics* 2021;47:296–9.
3 Faggioni MP, González-Melado FJ, Di Pietro ML. National health system cuts and triage decisions during the COVID-19 pandemic in Italy and Spain: ethical implications. *J Med Ethics* 2021;47:300–7.
4 Jecker NS, Wightman AG, Diekema DS. Vaccine ethics: an ethical framework for global distribution of COVID-19 vaccines. *J Med Ethics* 2021;47:308–17.
5 McDougall RJ, Gillam L, Ko D. Balancing health worker well-being and duty to care: an ethical approach to staff safety in COVID-19 and beyond. *J Med Ethics* 2021;47:318–23.
6 Rudra P, Lenk C. Process of risk assessment by research ethics committees: foundations, shortcomings and open questions. *J Med Ethics* 2021;47:343–9.
7 Brenna C, Das S. Imperfect by design: the problematic ethics of surgical training. *J Med Ethics* 2021;47:350–3.
8 Giordano S, Garland F, Holm S. Gender dysphoria in adolescents: can adolescents or parents give valid consent to puberty blockers? *J Med Ethics* 2021;47:324–8.
9 Durán JM, Jongsma KR. Who is afraid of black box algorithms? on the epistemological and ethical basis of trust in medical AI. *J Med Ethics* 2021;47:329–35.
10 Lang B. Concerning a seemingly intractable feature of the accountability gap. *J Med Ethics* 2021;47:336.