

Non-maleficence and the ethics of consent to cancer screening

Lotte Elton 

History and Philosophy and Science, University of Cambridge, Cambridge, UK

Correspondence to

Dr Lotte Elton, Department of History and Philosophy of Science, University of Cambridge, Cambridge CB2 1TN, UK; lotteelton@gmail.com

Received 6 February 2020

Revised 14 July 2020

Accepted 12 August 2020

Published Online First

21 September 2020

ABSTRACT

Cancer screening programmes cause harm to individuals via overdiagnosis and overtreatment, even where they confer population-level benefit. Screening thus appears to violate the principle of non-maleficence, since it entails medically unnecessary harm to individuals. Can consent to screening programmes negate the moral significance of this harm? In therapeutic medical contexts, consent is used as a means of rendering medical harm morally permissible. However, in this paper, I argue that it is unclear that the model of consent used within therapeutic medicine can be applied unproblematically to preventive medicine. Invitation to screening changes the pragmatic norms and expectations of the patient–doctor encounter such that two key principles of consent may be violated. First, the pragmatics of a medical invitation are such that patients may fail to be adequately informed, since patients appear to assume medical invitations are made with their best interests in mind, even where information to the contrary is outlined. Second, screening invitations may place pressure on patients; in the context of a medical encounter, to make an invitation to screening may constitute an inducement to accept. In order to be sure that a patient’s consent to a screening invitation is valid, we must make clear to patients that their decision to accept screening may be shaped not only by how information about screening is presented, but by the pragmatic form of the invitation itself.

INTRODUCTION

Cancer screening is often promoted for ‘saving lives’, the implication being that all patients undergoing screening stand to benefit. However, even where there is an aggregate benefit within a population, one of cancer screening’s inherent limitations is that some will benefit at the cost to harm to others.¹ These screening-related harms are often justified by policy-makers on the basis that screening programmes are effective—in other words, that they lead to a net reduction in morbidity and mortality.² Yet such a justification appears to violate the ethical principle of non-maleficence, since individuals are harmed by the intervention of medical professionals in the pursuit of a population-level benefit. In therapeutic medical contexts, non-maleficence is often overridden by the principle of autonomy, in the guise of consent. Can consent do the same moral work in preventive contexts like screening?

In this paper, I will explore whether an appeal to consent might justify overriding non-maleficence in cancer screening. I will first demonstrate why the principle of non-maleficence is relevant to screening programmes, even where they provide a net benefit. Next, I will assess the argument that harms incurred during screening can be made morally irrelevant

where patients have given their informed consent to participate. I will argue that the paradigm of consent as it applies to therapeutic medicine is not necessarily easily applicable to preventive medicine, since there exists a pragmatic difference between the two contexts. To resolve this issue, I extend Chwang’s³ discussion of framing effects to argue that patients must be made aware of this pragmatic difference in order for their consent to screening to be valid. Where this is the case, a patient’s consent may carry sufficient moral weight to render permissible any harm that they incur through screening.

For simplicity, I refer throughout to ‘doctors’ as the agents of screening. In practice, this is often not the case: in the UK, for example, the National Health Service (NHS) Screening Programme makes most invitations to screening programmes. Nonetheless, it seems reasonable to conflate the two as being, in patients’ eyes, ‘medical agents’.

SCREENING AND NON-MALEFICENCE

Even where screening programmes are effective, it remains the case that a net aggregate benefit is compatible with harm to many specific individuals. It is worth noting here that my concern is with *uncompensated* harms. By this, I mean harms that are incurred by individuals who gain no benefit from screening. In other words, while an individual who is diagnosed with cancer through their participation in a screening programme may undergo painful cancer treatment, this harm is morally outweighed if the early detection of their cancer prolonged their life. For other individuals, however, screening leads to overdiagnosis: the detection of abnormalities which would never have progressed to cause harm. This results in many patients receiving treatment that is medically unnecessary, such as cervical excision to treat precancerous cells which might have regressed spontaneously. Overtreatment has its own consequences: women undergoing cervical excision, for example, are more likely to experience preterm labour.⁴

That screening causes harm appears to violate non-maleficence, one of the core principles of medical ethics as outlined by Beauchamp and Childress.⁵ Non-maleficence stems from the axiom *primum non nocere*—‘first, do no harm’. Yet it is evident that medical professionals often *do* harm their patients in order to achieve some greater benefit: a surgeon may, for example, amputate a patient’s gangrenous leg in order to save their life. How can we distinguish harms which are ‘acceptable’ from those that are ‘unacceptable’? I would argue that we can morally justify the harm of amputating a patient’s gangrenous leg on the grounds



© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Elton L. *J Med Ethics* 2021;47:510–513.

that it is medically necessary to save their life, but that we cannot justify amputating *both* of their legs if amputation of only one would bring about the same benefit. In other words, harm is medically necessary—and thus morally acceptable—if it is the least harmful means of generating a greater benefit for a given patient. I therefore define non-maleficence as a *prima facie* obligation not to cause harm which is medically unnecessary. The distinction between medically necessary and unnecessary harm is useful in that it makes clear why we might deem it unacceptable for a patient's healthy leg to be amputated in order to save another patient's life. Amputating a patient's leg to save their own life carries the same net benefit as amputating a patient's leg to save another patient's life, but the latter case involves medically unnecessary harm.

Imagine, though, that in order to save a patient's life we must infect another patient with a cold. Perhaps, the relative balance of harm and benefit in this case might lead us to believe that the infliction of a medically unnecessary cold is morally acceptable. However, given that harms of screening include unnecessary mastectomies,⁶ I would suggest that the harm inherent in screening is more akin to leg amputation than to a cold. As a result, we cannot easily override non-maleficence in screening whenever there is an overall benefit, since it is tricky to mandate serious and medically unnecessary harm on some individuals even where it brings about a net population benefit. How, then, might we resolve the issue of non-maleficence in screening? One possible answer is to suggest that screening-related harm can only be justified where expected benefit to a *particular individual* outweighs expected harm to that same individual. This, however, is an impossible standard: for any particular individual screened, we cannot know, at the outset, whether the benefit to them outweighs the harm. More than this, I would argue that we cannot know this even after screening has taken place, since to assess the balance of benefit and harm requires counterfactual knowledge of what would have happened had we not intervened. This is difficult, if not impossible, to obtain.

Resolving the balance of population-level beneficence and individual-level non-maleficence is, evidently, a tricky endeavour. Might there be an easier way to justify harm in screening? Although screening necessarily causes medically unnecessary harm, any given individual participating in screening could, in principle, benefit: screening-related harm results from risks imposed on an individual in order to benefit either that individual or others. In this way, screening can be viewed as a sort of lottery. It is much more likely that a lottery participant will be harmed by losing the money spent on the lottery ticket than that they will benefit by winning the jackpot. The potential benefit, however, is great, which may make the harm of wasting money seem insignificant in comparison. In buying a lottery ticket, a participant therefore takes on a risk which will result in benefit either to themselves or to others. If we are to accept that both lotteries and screening confer great benefit on at least some individuals, it might be argued that a rational agent may wish to accept the invitation to screening—despite the harms it might entail—on the basis that they stand to potentially benefit. If an individual wishes to accept the invitation, can their consent to the offer make morally irrelevant any harm they might suffer?

THE PROBLEM OF CONSENT IN PREVENTIVE MEDICINE

Let us first consider the moral status of consent as it applies to therapeutic medicine. Consent does moral work by 'justifying action that would otherwise violate important norms, standards or expectations'.⁷ Where a patient gives their consent to the

amputation of their gangrenous leg in order to save their life, they render morally permissible an action which would usually violate important norms. In order for their consent to be valid, it must be both informed and autonomous; in other words, they must have sufficient information about the consequences of the action to which they give their consent, and be free of pressuring influences.¹ These principles of consent—well-established in therapeutic medical contexts—have historically been applied unproblematically to screening. However, several authors have hinted at the idea that there is an ethical distinction between 'everyday' medical practice and preventive interventions such as screening.^{8–9} Those who make this argument suggest that preventive medicine is ethically different to therapeutic medicine, although few have outlined exactly why. In this section, I aim to demonstrate that there is a pragmatic difference between the patient-initiated requests of therapeutic medicine and the doctor-initiated invitations of preventive medicine. This changes the ethical responsibilities which must be borne by the doctor, and the terms by which consent can be expected to do moral work.

In therapeutic medicine, the patient presents to a doctor with a *concern*: most often, a symptom or set of symptoms. A patient makes a request for this concern to be addressed by the doctor, thereby entering into a transaction in which the 'special social and cognitive authority' of the doctor is acknowledged.¹⁰ In screening, the transaction is different. Here, a doctor approaches an individual who has not sought out a medical encounter, and *invites* her to participate in screening. If we make a distinction between the contexts of patient-initiated requests and doctor-initiated invitations, how might this lead us to conclude that consent to the latter cannot meet the transactional demands of consent? Approaching healthy individuals with the invitation to participate in a medical intervention disrupts the normal framework of the therapeutic doctor–patient encounter in which a patient makes a request of the doctor. If, as Rebecca Kukla argues, the 'rituals of the clinic' are 'essential to giving the transaction [of consent] the pragmatic form it has',¹⁰ then we cannot expect the same formulation of consent to automatically apply in the case of preventive medicine. In the rest of this section, I will argue that invitations to screening present a problem on two grounds: first, that the pragmatics of making a medical invitation are such that a patient is liable to be misinformed; and second, that to make a medical invitation at all may constitute an inducement to accept.

Informed consent

It is well documented that patients may be misinformed about screening—for example, that they are not always clear about the risks and potential outcomes of participating in a screening programme.¹¹ This is perhaps unsurprising, given that overdiagnosis is rarely quantified in screening trials¹² and seldom included in patient information leaflets.¹³ Even where risk information is provided to patients, the way in which risk is presented may affect the decisions patients make: for example, a patient may consent to an intervention which is described as conferring a 90% chance of survival, but refuse the same intervention when it is described as having a 10% chance of mortality. This is termed the 'framing effect'.³ A wide range of public health literature has sought to address framing effects, detailing tools such as decision aids and risk communication primers which explain, for example, that 90% survival and 10% mortality mean the same thing.^{14–15} I do not dismiss these endeavours: it is valuable to consider how patients can be provided with the most clear and relevant information, and how framing effects might be minimised in order

to make patients' consent more morally robust. Here, though, I present a further concern: that even leaving aside issues of risk heuristics and communication, the pragmatic form of a medical invitation is such that the 'informed' condition of consent may be difficult to meet, even were we to optimise the communication of risk to patients.

Doctors are trusted.¹⁶ When patients make requests of their doctors to attend to their concerns, they assume their doctor will not suggest medical solutions which are 'useless, unprofessional, too risky or illegal'.⁷ The options given in response to a patient request are thus curtailed by normative constraints even before a patient's consent is sought. But doctor invitations have a different pragmatic meaning to patient requests. In these cases, patients expect—and do so reasonably—that a health professional would not invite them to participate in an intervention unless the health professional expected that intervention would be of benefit to the recipient. This may lead a patient to downplay adverse consequences of screening such as overdiagnosis. We can see this clearly in the case of therapeutic misconception, where patients invited to participate in clinical research often report benefit to themselves and other research subjects as the primary aim.¹⁷ This goes beyond lack of disclosure of adequate information: such beliefs persist even where patients are clearly told that the study is not intended to benefit them but future patients, and even where they demonstrate an understanding of randomisation and the use of placebos.¹⁸ In my view, therapeutic misconception points to a much deeper issue: patients appear to assume that medical invitations are made with the intention of benefiting them personally.

When a doctor makes an offer of screening, the pragmatic implication understood by patients is that this is an offer *in the patient's best interests*. This is the case even where the risks of participation are outlined, and even where it is made clear that the programme has population level rather than individual benefit. By making the invitation to screening, doctors are suggesting to healthy people that via their participation in the screening programme they will be even healthier, or live a healthy life for longer.¹⁹ Here, we can see how a conduit/container model of communication outlined by Manson and O'Neill—whereby information is transferred between agents who 'originate' and 'receive' messages—fails to adequately capture the complexity of the consent transaction.⁷ Even where communication of a message is intelligible and relevant, it cannot be assumed that both parties come to share the same understanding. Patients expect that their doctors will behave both beneficently and non-maleficently towards them. As a result, they may fail to understand when doctors are unable to uphold these duties, even when this is clearly explained.

Patient choice

Let us turn to my second consideration: that invitations to screening are problematic from the point of view of patient choice. Kukla discusses the pragmatic implications of invitations, outlining how invitations 'leave the invitee neither obligated nor with a neutral free choice'.²⁰ An individual is free to decline the invitation to screening. But an invitation—even where it may be declined without insult—must be *welcoming*: it must be delivered in such a way that it conveys encouragement rather than indifference. In Derrida's words, it should not imply 'if you don't come, never mind, it doesn't matter'.²⁰ Do offers of screening constitute 'invitations', in this pragmatic sense? I believe so. To take an example, the letter inviting women to cervical screening begins 'the NHS offers cervical screening to save lives from cervical cancer'. This is hardly 'never mind, it doesn't matter'.

General Practice surgeries are advised to add a personal message to letters on the grounds that GP endorsement encourages patients to accept the invitation of screening.^{21 22} Even though a caveat is added that the decision to accept cervical screening is 'your choice', this is at odds with the pragmatic implication of the rest of the message.

What if were to remove 'nudges',²³ like personal messages written by the GP, from screening invitations? This would certainly go some way to lessening the pressure on a patient to accept. However, I have a further concern. If we are to understand fully the pragmatic implications of an invitation, we must consider the context in which invitations are made, and the relationship of power between inviter and invitee. In recent decades, there has been a concerted effort to move to a model of healthcare which emphasises a sharing of power between doctor and patient.²⁴ Nonetheless, doctors enjoy a high degree of authority within the medical environment²⁵; as a result, in the context of a medical encounter, to make the invitation to screening *at all* constitutes at least a partial inducement to accept. It could be argued that the process of obtaining consent in any medical situation is necessarily moulded by dynamics of power and authority. But there is a distinction between patients making a request and doctors making an invitation. Consider the difference between approaching a lawyer for legal advice, and receiving a letter from a lawyer suggesting you bring a lawsuit. Lawyers, like doctors, occupy positions of authority. Where a patient has made a request of a doctor, or has sought the advice of a lawyer, she has chosen to participate in a transaction structured by relations of power and authority. Where a patient is invited to take part in screening, or to bring a lawsuit, she has not sought out this authority, nor does she have a concern she wishes to be addressed. Medical authority has thus been imposed on her; in such a situation, we must consider carefully whether her consent can be taken to be valid.

IMPROVING CONSENT

Are we to conclude that patients are irreconcilably misinformed about screening? Are invitations to screening always inherently coercive? Not necessarily. As I have outlined, the pragmatic form of an invitation to screening may mean that patients are pressured into accepting, or misinformed about its implications. Perhaps, though, we could see the pragmatic form of a screening invitation as a kind of framing effect, leading patients to perceive screening differently on account of how the invitation is made and who it is made by. How might we tackle the issue if we were to view it as a framing effect? Chwang³ argues that a patient's consent may be invalidated where framing effects are present, but that this consent can be revalidated by 'debiasing': alerting patients to the framing effects to which they are subject. Thus, rather than simply presenting patients with information (eg, that an intervention has 10% mortality and/or 90% survival), we can instead present *metainformation*, which highlights framing effects: for example, by explaining to patients that they liable to make a different decision about accepting or rejecting the intervention depending on whether they are told that it has 10% mortality or 90% survival. We can also apply a similar information/metainformation distinction to my earlier example of therapeutic misconception. Research participants are given information: for example, 'the aim of this research study is to confirm that Drug X is effective in treating cancer'. Despite this, participants still appear to believe that trials are conducted for their own benefit. Perhaps, though, therapeutic misconception could be addressed by providing metainformation: for

example, ‘research participants may be given the impression that the research trial is intended to treat their illness, even though this is not its main aim’.

The provision of metainformation is, in my view, the most compelling means to address the issues I have raised in this paper about the pragmatics of medical invitations. What might the provision of metainformation about screening look like? Screening information literature could make clearer the fact that many patients (reasonably) believe that an invitation to screening is made with their best interests in mind, even though it is inevitable that some individuals will be harmed by screening. Similarly, invitations to screening could stress not only that accepting the invitation is a patient’s choice, but also explain that patients may well feel pressured to take up invitations made by their doctor. Such measures would go some way to debiasing patients, since they highlight to patients that their consent to screening is shaped by the fact that they are invited to participate by doctors. If, through debiasing, we are able to eliminate the framing effect brought about by the pragmatic form of a medical invitation, this would negate my concern that such invitations may pressure or misinform patients. In turn, this would allow patients’ consent to screening invitations to do moral work in justifying the harm which patients may incur if they undergo screening.

CONCLUSION

Non-maleficence in screening is not a problem that can easily be cast aside, even where a screening programme has population-level benefit. In therapeutic medical contexts, non-maleficence is sometimes overridden by an appeal to consent. However, it is not clear that the paradigm of consent as it applies to therapeutic medicine can be applied unproblematically to preventive medicine cases. The invitation to screening changes the pragmatic norms and expectations of the encounter between patient and doctor, with the result that consent may be unable to do the same moral work as in therapeutic encounters initiated by a patient. One means by which we could address this problem is through the provision of metainformation, which makes clear to patients how the pragmatics of a screening invitation may affect their decision to accept an invitation to screening. This has relevance beyond cancer screening. Where individual harm may result from other doctor-initiated invitations—for example, prenatal screening for genetic abnormalities, or cardiovascular health checks—providing information to patients about the risks and benefits of the intervention may be insufficient. In order for a patient’s consent to a doctor-initiated invitation to be valid, we must make clear to the patient that they are subject to framing effects, which relate not only to risk information, but to the pragmatic form of the invitations itself.

Acknowledgements Dr Stephen John critically reviewed two earlier drafts of this paper in his capacity as academic supervisor.

Contributors LE drafted the manuscript and made all revisions.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. No datasets were generated or analysed in producing this manuscript.

ORCID iD

Lotte Elton <http://orcid.org/0000-0003-4621-9131>

REFERENCES

- Jepson RG, Hewison J, Thompson AGH, *et al.* How should we measure informed choice? the case of cancer screening. *J Med Ethics* 2005;31(4):192–6.
- Plutynski A. Ethical issues in cancer screening and prevention. *J Med Philos* 2012;37(3):310–23.
- Chwang E. Consent’s been framed: when framing effects invalidate consent and how to validate it again. *J Appl Philos* 2016;33(3):270–85.
- Wuntakal R, Castanon A, Landy R, *et al.* How many preterm births in England are due to excision of the cervical transformation zone? nested case control study. *BMJ Pregnancy Childbirth* 2015;15(1):1–7.
- Beauchamp T, Childress J. *Principles of biomedical ethics*. 5th ed. Oxford University Press, 2001.
- Raffle A, Gray M. *Screening: evidence and practice*. Oxford, UK: Oxford University Press, 2007.
- Manson NC, O’Neill O. *Rethinking Informed Consent in Bioethics*. Cambridge, UK: Cambridge University Press, 2007.
- Cochrane AL, Holland WW. Validation of screening procedures. *Br Med Bull* 1971;27(1):3–8.
- Charlton BG. Screening, ethics, and the law. *BMJ* 1992;305(6852):521.
- Kukla R. Communicating consent. *Hastings Center Report* 2009;39(3):45–7.
- Hersch J, Jansen J, Barratt A, *et al.* Women’s views on overdiagnosis in breast cancer screening: a qualitative study. *BMJ* 2013;346(1):f158–13.
- Heleno B, Thomsen MF, Rodrigues DS, *et al.* Quantification of harms in cancer screening trials: literature review. *BMJ* 2013;347:f5334–9.
- Dreier M, Borutta B, Seidel G, *et al.* Communicating the benefits and harms of colorectal cancer screening needed for an informed choice: a systematic evaluation of leaflets and booklets. *PLoS One* 2014;9:e107575.
- McCaffery KJ, Jansen J, Scherer LD, *et al.* Walking the tightrope: communicating overdiagnosis in modern healthcare. *BMJ* 2016;352.
- Hersch J, Barratt A, Jansen J, *et al.* Use of a decision aid including information on overdiagnosis to support informed choice about breast cancer screening: a randomised controlled trial. *Lancet* 2015;385(9978):1642–52.
- Oliver D. A matter of trust - doctors, the NHS, patients, and the public. *BMJ Opin*, 2018. Available: <https://blogs.bmj.com/bmj/2018/02/20/david-oliver-a-matter-of-trust-doctors-the-nhs-patients-and-the-public/> [Accessed 21 Nov 2019].
- Appelbaum PS, Roth LH. The structure of informed consent in psychiatric research. *Behav Sci Law* 1983;1(4):9–19.
- Lidz CW, Albert K, Appelbaum P, *et al.* Why is therapeutic misconception so prevalent? *Camb Q Healthc Ethics* 2015;24(2):231–41.
- Marshall KG. Prevention. how much harm? how much benefit? 4. The ethics of informed consent for preventive screening programs. *CMAJ* 1996;155(4):377–83.
- Kukla R. That’s what she said: The language of sexual negotiation. *Ethics* 2018;129(1):70–97.
- Benton SC, Butler P, Allen K, *et al.* GP participation in increasing uptake in a national bowel cancer screening programme: the pearl project. *Br J Cancer* 2017;116(12):1551–7.
- England PH. Cervical screening: ideas for improving access and uptake. *Cerv Screen Cover data* 2019.
- Thaler R, Sunstein C. *Nudge: improving decisions about health, wealth and happiness*. London: Penguin, 2009.
- Goodyear-Smith F, Buetow S. Power issues in the doctor-patient relationship. *Health Care Anal* 2001;9(4):449–62.
- Koeck C. Imbalance of power between patients and doctors: the price of ignoring the shadow. *BMJ* 2014;349:1–2.