****

Patient Preference Predictors, apt categorisation and respect for autonomy

Abstract

In this paper, I set out two ethical complications for Rid and Wendler's proposal that a "Patient Preference Predictor" should be used to aid decision-making about incapacitated patients' care. Both of these worries concern how a PPP might categorise patients. In the first section of the paper, I set out some general considerations about the "ethics of apt categorisation" within stratified medicine, and show how these challenge certain PPPs. In the second section, I argue for a more specific – but more contentious claim – that proper respect for the autonomy of incapacitated patients might require us to act on reasons which they could endorse, and show how this claim places important limits on the categories employed by an ethically acceptable PPP. The conclusion shows how these concerns about apt categorisation relate to more familiar worries about Rid and Wendler's proposals.

Key words: Patient Preference Predictor; incapacitated patients; categorisation; stratified medicine; autonomy

Word Count: 3391

Patient Preference Predictors, apt categorisation and respect for autonomy

In this volume, Rid and Wendler argue that family members will often be less accurate at identifying incapacitated patients' treatment wishes than a Patient Preference Predictor (PPP), an algorithm which predicts individuals' choices on the basis of their group memberships (Rid and Wendler, 2013a, 2013b). For example, my parents' prediction of my treatment preferences may well be less accurate than one based on the preferences of other white, 33 year old, middle-class atheists. One important desideratum in making choices about incapacitated patients' care is that we treat them as they would have chosen.[i] If so, it seems we should use a PPP when deliberating about incapacitated patients' care as well as considering family members' predictions. Although forcefully argued, I find Rid and Wendler's proposals disturbing; furthermore, on the basis of an (unscientific) surveying of friends and colleagues, I am not alone. A sense of disquiet is, however, not an argument. In this paper, I set out potential problems for PPPs, related to how they might categorise patients, which explain this sense of unease, and suggest both possible improvements to PPPs and broader conclusions regarding autonomy and personalised medicine.

To clarify my arguments, I shall mention some concerns I will not directly discuss. First, we might find the fact that families are (comparatively) worse at predicting loved ones' choices disquieting, as it undermines a deep assumption that our loved ones "know us best". This might lead us to question the factual accuracy of Rid and Wendler's work. However, although constructing a PPP with high external validity could be difficult, their factual claims about the relative predictive accuracy of families and well-designed PPPs seem plausible. Second, we might object to Rid and Wendler's normative proposals. These objections might be grounded in autonomy-related reasons – for example, patients would want family members to make choices – or non-autonomy related reasons, such as maintaining the "noble lie" that families know best. However, the first proposal faces a

problem Rid and Wendler discuss: individuals' preferences for consulting family members over using a PPP often rest on the false belief that families know best, such that deferring to family members might not reflect patients' autonomous wishes. The second response is even more problematic: even if medical ethics has overstated the value of autonomy, it is hard to argue that we should perpetuate falsehoods.

In this paper, I shall discuss a less familiar worry: that the categories employed by a PPP might be ethically problematic. In the first section I sketch a general account of how moral and epistemic concerns inter-relate in clinical prediction, and show how this "ethics of apt categorisation" challenges some PPPs. In the second section, I use this framework to suggest an autonomy-related objection to use of a PPP: that it does not frame decisions in ways patients could endorse.

1. The Ethics of Apt Categorisation and the PPP: general issues

Despite the hype surrounding "personalised medicine", it is more accurate to say that we are entering an era of "stratified medicine", where interventions are tailored towards patients on their risk-profile, based on their membership of clinically significant sub-groups (Trusheim, Berndt and Douglas, 2007). For example, genetic tests allow us to sub-categorise populations into low or high risk groups, according to genotypic features, and to tailor treatments accordingly. Although the PPP is, despite its name, not a predictive tool (because it aims at retrodiction of what individuals would have chosen), it is part of this trend: by sub-dividing populations into smaller groups, we make choices which are more fitting for individuals. Therefore, assessing the PPP requires considering an under-explored problem for stratified medicine. There are always multiple ways of sub-dividing any population which could generate more accurate risk estimates about individuals thus categorised; how should we choose between these alternative categorisation schemes?

3

The answer to this problem seems simple: (subject to cost considerations), we should always use the most fine-grained categorisation system available, because this will give us the most accurate predictions. However, even if narrower categories are always more predictively accurate – a topic of philosophical controversy (Hajek, 2007) – there are two, inter-related problems with this answer, which I will set out generally before turning to their implications for PPPs specifically. First (less importantly for this paper), given Geoffrey Rose's point that many more cases of disease arise in large "moderate risk" groups than in small "high risk" groups (Rose, 2008, Chap.4), coarse-grained schemes may be better than finer-grained schemes at reducing overall disease burden. When constructing a breast cancer screening programme, a coarse-grained sub-categorisation of the population by age alone will not provide accurate predictions of each individual's chance of cancer, but may be more useful than a fine-grained sub-categorisation involving genetic markers and age as a means of reducing overall population mortality. A PPP does not aim at population benefit, so does not face a version of these worries, but Rose's point is an important reminder that categorisation schemes can have benefits and costs other than those related to predictive accuracy.

Second, even when our categorisation schemes are to be used for making individualised predictions, there can be non-epistemic problems with using certain categorisation schemes. For example, imagine a physician knows that sub-categorising patients according to some genotypic feature will provide a more accurate estimate of their cancer risk than relying on average population risk, but that providing this information will feed into patients' genetic fatalism, leaving them less likely to make behavioural changes in other areas of their lives. If so, categorising patients by genotype conflicts with her duty of non-maleficence. Second, imagine a physician knows that sub-categorising patients by ethnicity will provide more accurate information of their risk of dementia, but, given widespread racist beliefs, this practice will reinforce views about certain races' "feeble-mindedness". If so,

categorising by race increases accuracy but breaches general duties not to perpetuate social injustice.[ii] In both of these cases, the physician might reasonably decide that the benefits of higher accuracy do not outweigh the costs of using more fine-grained categorisations. More could be said about such cases, but my key point is simple: the relatively high accuracy of some way of sub-categorising populations is not a sufficient reason to prefer that system of sub-categorisation, but we must also consider the ethical and political costs of (being seen to be) grouping people in some way. We should consider not only the epistemology but the ethics of apt categorisation.[iii]

Even without developing a full ethics of apt categorisation, we might worry about some PPPs along these lines. For example, we need not claim that it is inherently unfair to categorise individuals on the basis of racial group membership to worry that a PPP which incorporated racial categories might perpetuate societal racism. Alternatively, certain correlations – say between socio-economic class and willingness to continue treatment – often result from unjust background conditions, for example, the costs of accessing healthcare. As such, a PPP based around such correlations (or proxies, such as occupation) may perpetuate background socio-economic injustice.

These considerations do not constitute a knockdown argument against Rid and Wendler's proposals. First, they do not suggest it is wrong to use a PPP *per se*, but only some PPPs. Second, whether or not a PPP would create or perpetuate harm or injustice involves empirical and ethical issues. Third, even if we are confident that using a PPP would be ethically problematic, an overall assessment would have to balance these costs against the putative ethical benefits. Finally, use of surrogates' choices might face analogous challenges – for example, that they are shaped by unjust background conditions – suggesting symmetrical concerns about current practice. Therefore, the comments above suggest reasons for unease at, rather than a fundamental challenge to, Rid and Wendler's proposals.

5

2. The ethics of apt categorisation and respect for autonomy: a specific challenge to the PPP

However, a plausible interpretation of the demand that we respect incapacitated patients' autonomy suggests a more fundamental way in which categorisations might be inapt. Rid and Wendler list six desiderata for any method of determining incapacitated patients' treatment PPP. At least three – we should enable the patient to control how she is treated, we should provide treatment consistent with the patient's considered preferences and values, and we should respect the patient's preferences for how treatment decisions are made – are, at least partly, grounded in the value of "respect for autonomy". Respect for autonomy is a, perhaps the, fundamental ethical concern in medical ethics. However, Rid and Wendler's desiderata fail to capture the full force of the demand that we respect autonomy.

Rid and Wendler's autonomy-related argument in favour of use of a PPP can be rewritten as follows: respect for autonomy demands that surrogates' choices are in-line with what a patient would have chosen on the basis of her considered values; use of a PPP is likely to maximise the number of times we make such choices; therefore, we should use a PPP. However, this argument seems to imply that when treating competent patients, a physician who is certain that she can predict what patients will choose might reasonably choose not to seek patients' actual consent: each time she chooses how to treat a patient, her choice is, in some sense, controlled by what the patient would choose on the basis of his preferences and values; not seeking consent would, presumably, speed up the number of treatments which could be performed; therefore, dropping consent requirements would maximise the number of cases in which autonomy is respected. Even placing to one side epistemological worries about the physician's accuracy, this line-of-reasoning is ethically troubling. One way of capturing this worry is suggested by Buchanan and Brock's useful distinction between "performances" of choice and "evidence" of what patients would choose (Buchanan and Brock, 1989, 116). Ultimately,

respect for autonomy involves respect for the performance of choice, or, more specifically, the performance of choices which meet certain desiderata (they are uncoerced, informed, and so on). Our imaginary physician overlooks this distinction. Clearly, however, incapacitated patients cannot perform choice. How, then, does stressing that respect for autonomy relates, ultimately, to respect for acts of choice, not merely respect for what agents would choose, relate to incapacitated patients? This is a complicated question, which requires far more discussion than space allows, but I will focus on one way of reading Buchanan and Brock's distinction: as stressing that respect for autonomy involves respect for others in their capacity as rational deliberators.[iv]

I propose, then, an intuitively plausible extra desideratum for making choices for incapacitated patients, which captures this concern: the reasons we give for those choices are reasons which those patients would themselves endorse, were they competent.[v] For example, imagine that Jane, who is in a permanent vegetative state, previously expressed horror at the thought of losing her mental capacities. We know that most patients whose first name begins with a "J" would refuse further care under such circumstances. Deciding how to treat Jane by appeal to her past preferences or by appeal to statistical regularities in "J"-named people would lead to the same conclusion. Rid and Wendler agree that in cases where we have clear evidence of past preferences, a PPP should not be used. I suggest, however, that they fail to capture a deep reason for such a preference, well-illustrated in this case: that it is preferable to base our decision on Jane's past preferences, because doing so is framed in terms she could (presumably) endorse, and, hence, respects her capacity for rational deliberation. By contrast, it seems deeply implausible that Jane would base her choice on the first letter of her name; as such, she would not endorse that reason for terminating her treatment. Even if our J-based decision is – in some sense – "right", basing our decision on this fact fails to respect her capacity for rational deliberation. More colloquially: respecting autonomy demands not only that we act on the "right answer", but do so for the "right kind" of reason. (Note that I treat these claims as definitive of

what is involved in respecting autonomy, rather than as an empirically testable prediction about how

agents would themselves prefer to be treated).[vi]

A standard argument for consulting family members rather than physicians about incapacitated

patients' care is that they are more likely to choose as the patient would have chosen. Rid and

Wendler are correct that, given the empirical data, this argument implies we should prefer a PPP to

consulting family members. However, I suggest that respect for autonomy involves more than

choosing as patients would choose, but making decisions using "patient-endorsable reasons". If so

we might offer a second reason for consulting family members, rather than physicians: that they are

more likely to identify reasons which patients would endorse. Rid and Wendler have not shown that

a PPP is preferable to family members on this score. Admittedly, their work suggests reasons to be

sceptical that family members do reliably identify reasons patients endorse: were they to do so, they

would not be so inaccurate! As such, we might have good reasons to worry about current practice,

but my focus here is on comparing current practice to use of a PPP: that family members are often

bad at identifying patient-endorsable reasons does not imply that a PPP is any better. What, then, can

be said about use of a PPP on my broader conception of respect for autonomy?

It might seem that a PPP can never identify reasons which agents would endorse as part of an

autonomous decision-making process, because autonomy essentially involves making choices as an

individual, rather than as a member of a group. However, this is mistaken. As I have suggested

above, a decent account of why autonomy is respect-worthy will view autonomy in terms of agents'

capacity to make choices in-line with their considered values and commitments. Given that agents'

values and commitments are often drawn from their group membership, there is no inherent tension

in thinking of agents as both autonomous and as members of groups; choices not in-line with (at least

certain sorts of) group membership, say religious affiliation, are more likely to be non-autonomous.

Although autonomy and individuality are closely related concepts, a plausible account of autonomy does not suggest that individuals' reasons can never be related to their group membership. Therefore, the expanded account of respect for autonomy does not rule out using a PPP.

However, the expanded account of respect for autonomy does suggest a more complicated problem for PPPs. Any individual will always belong to multiple groups, and only some of those group memberships seem likely to play a role in autonomous decision-making. For example, someone might reason through difficult treatment decisions by thinking about his commitments *qua* Catholic or *qua* father, but seems less likely to think through these choices *qua* tenement-dweller or *qua* middle class person. However, this distinction between reasons-giving and non-reasons-giving-group membership may not reflect the distinction between predictively-useful and predictively-less-useful group memberships. The most predictively accurate PPP might categorise us in ways which do not reflect the group-memberships we treat as relevant when making autonomous choices. For example, Rid and Wendler propose that gender and education might be useful variables for constructing a PPP, but I doubt many people think through treatment choices by viewing themselves *qua* Man or *qua* University-graduate.[vii] This potential "gap" is, I think, part of what makes Rid and Wendler's work unsettling: it seems troubling to think that one's own agonising choices might be predicted by features one never considers when making such choices. On the "endorsable-reasons" account of autonomy, this gap is not only troubling, but normatively relevant: that certain categories are predictively powerful is not a sufficient reason for using those categories if they do not relate to endorsable reasons. Proper respect for autonomy might place serious constraints on which categorisations are "apt" in making treatment choices, but the most predictively successful PPP need not employ only "apt" categorisations.

Rid and Wendler might respond to this problem in at least three ways. First, they could deny that respect for autonomy involves deciding on the basis of endorsable reasons at all. Second, they might concede that deciding on the basis of endorsable reasons is one aspect of respect for autonomy, but claim that respect for autonomy also involves choosing as patients would choose; if so, although a PPP need not capture all of "respect for autonomy", it is preferable to consulting family members. Third, they might argue that there might be PPPs which are both more predictively accurate than family members and which only employ the kinds of categories which (typically) match-up to reasons-giving group memberships; for example, ones restricted to such features as "religiousness", "perceived quality-of-life" or "medical history". Such PPPs would represent an improvement on consulting family members along both dimensions of respect for autonomy. I do not know which of these options (if any) Rid and Wendler would choose, but the third is most interesting, as it would (in principle) be open to empirical confirmation.

Conclusion

At the start of this paper, I expressed disquiet about Rid and Wendler's proposals. There are three broad strategies we might adopt to attack their claims. First, we might deny their empirical findings. Second, we might accept the validity of their empirical findings, but deny that we should use a PPP on non-autonomy related grounds; for example, by appeal to the value of maintaining the "noble lie" that families know best. Third, we might accept the validity of their empirical findings, but deny that proper respect for autonomy implies using a PPP rather than consulting family members; for example, because of how patients would want choices to be made. In this paper, I have developed novel versions of the second and third strategies, related to the ethics of categorisation. First, I argued that even if a PPP would respect autonomy, there might be non-autonomy-related reasons, such as the perpetuation of injustice, not to use a PPP. Second, I argued that a proper understanding of the

nature of respect for autonomy might place limits on which categories a PPP should employ. Neither argument showed that a PPP should not be used in clinical practice; rather, they both counsel care about thinking that the fact that using some way of categorising patients would maximise predictive accuracy is a compelling reason to categorise patients in that way.

References

Buchanan, A. and D. Brock 1989 *Deciding for Others: The Ethics of Surrogate Decision Making*, Cambridge: Cambridge University Press

Hajek*,* A. 2007, "The reference class problem is your problem too", *Synthese*, 156(3): 563-85.

John, S. (forthcoming) "Cancer screening and the ethics of apt categorisation" in Strech, D. *et al*. (eds) *Ethics in Public Health and Health Policy*

Kitcher, P 2001 *Science, Truth and Democracy* Oxford: Oxford University Press

Olsaretti, S 2005 "Endorsement and Freedom in Amartya Sen's *Capability* Approach" *Economics and Philosophy* 21 (1):89-108

Rid and Wendler 2014a "Treatment Decision Making for Incapacitated Patients: Is Development and Use of a *Patient Preference Predictor* Feasible?" *Journal of Medicine and Philosophy XXX*

Rid and Wendler 2014b "Use of a *Patient Preference Predictor* to Help Make Medical Decisions for Incapacitated Patients" *Journal of Medicine and Philosophy XXX*

Rose, G. 2008 *The Strategy of Preventive Medicine* Oxford: Oxford University Press

Scanlon, T. 1998 *What we owe to each other* Cambridge, MA.: Harvard University Press

Trusheim, M.R., E. R. Berndt and F. L. Douglas 2007 "Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers" *Nature Reviews Drug Discovery* 6:287-293

---

[i] Buchanan and Brock (1989) is the canonical justification of such an assumption.

[ii] See Kitcher (2001, Chap.8) for insightful discussion of a related concern about how scientists' epistemic duties and broader social duties might conflict.

[iii] See John (forthcoming) for a more developed account of these issues.

[iv] This is, of course, a very Kantian formulation, which Buchanan and Brock themselves may not accept. However, these claims do not depend on buying the entire Kantian apparatus, but might be grounded in the less metaphysically contentious framework set out by Scanlon (1998).

[v] My thoughts here are influenced by Serena Olsaretti's discussion of a similar issue in the interpretation of Sen's account of capabilities (Olsaretti, 2005).

[vi] I am grateful to Annette Rid for pointing out the importance of clarifying this claim.

[vii] Of course, you might disagree with these claims, and clearly related gender-roles might often structure decisions, but the general point does not depend on this example.