







Respecting formerly autonomous persons: clarifying the role of the Personalised Patient Preference Predictor (P4) in substituted judgement

Brian D Earp ^{1,2} Sebastian Porsdam Mann ^{1,3,4}
Tessa van Veenendaal ² Jemima Winifred Allen,^{2,5}
Sabine Salloch ⁶ Karin Rolanda Jongsma ⁷ Matthias Braun ⁸,
Walter Sinnott-Armstrong,⁹ Julian Savulescu,^{1,2} Annette Rid,¹⁰
David Wendler¹⁰

INTRODUCTION

In a recent paper,¹ we proposed a Personalised Patient Preference Predictor (P4), building on earlier work by Rid and Wendler.² The P4 is a hypothetical computer program that would, in the context of surrogate decision-making (eg, following a substituted judgement standard), use generative artificial intelligence (AI) models to infer a patient's underlying values and preferences and, on that basis, predict which treatment option they would choose in the current situation. Such AI models, we suggested, could be 'fine-tuned' on various pre-existing data or materials produced by, or otherwise pertaining to (eg, including information about), the patient considered as a unique individual.^{1,3}

Annoni has offered a thoughtful critique of this proposal, challenging what he (implicitly) characterises as assumptions underlying our work.⁴ In his critique, in addition to advancing

the negative claim that the P4 would not, as we had suggested,¹ allow surrogate decision-makers to more effectively show respect for formerly competent patients, Annoni makes the positive claim that the P4 *could* help surrogates better honour the patient's unique identity. We appreciate Annoni's contribution to this important ethical conversation and welcome the opportunity to clarify our position.

We agree with Annoni that the P4 serves ethical purposes beyond showing respect for autonomy, and that this may include honouring their identity.⁵ We also agree, contrary to what Annoni implies about our position, that respect for autonomy is not simply reducible to mere preference satisfaction. However, we do not agree with Annoni's claim that the P4 does not, or could not, promote the value of autonomy at all. We argue that Annoni's view of respect for autonomy is too narrow, insofar as it seems to require occurrent explicit consent, which is impossible for incapacitated patients. We advocate for a spectrum approach to autonomy: When direct choice or consent is possible, autonomy will usually be best respected by deferring to the person's competent decision. When it is not possible, however, honouring a person's competently held/endorsed values and preferences (even if only inferred from prior data) is more autonomy-respecting than certain alternatives, such as relying solely on others' best interests judgements.

ANNONI'S RECONSTRUCTION OF OUR (?) ARGUMENT

Annoni begins his critique by citing and quoting from our recent article,¹ while also using our distinctive terminology (ie, referring to the 'P4'). However, he then

states that he will 'abstract from' our work and focus instead on the following argument, which he does not explicitly attribute to us (or anybody else):

1. Substitute judgement enables respect for the autonomy of incapacitated patients.
2. Respecting autonomy entails choosing what incapacitated patients would choose.
3. Choosing what patients would choose requires identifying their treatment preferences.
4. The P4 enables the identification of patients' treatment preferences.
5. Therefore, using the P4 enables respect for the autonomy of incapacitated patients.

Annoni⁴ states that '[m]ost critiques so far have targeted the ability of P4 models' to accurately predict patients' preferences (ie, premise 4), although he does not cite any such critiques, so the basis for this claim is unclear. Annoni then states that he will focus, not on the technical ability of a P4 to accurately identify the treatment a patient would choose if competent, but rather on two assumptions which he believes are implied by the first two premises of the above argument: first, the view that *respect for autonomy can be equated with the satisfaction of individual treatment preferences* (an assumption Annoni associates with premise 2; however, this arguably conflates satisfaction of treatment preferences with what a person would choose—the language actually used in premise 2); and, second, the view that *incapacitated patients have a residual autonomy to be respected* (an assumption that Annoni associates with premise 1).

To clarify, in our original paper,¹ we neither argue for nor assume either of these premises or the associated claims. We do describe the position that 'merely' choosing the treatment an incapacitated patient would have chosen is one way—but by no means the only way—for proxy decision-makers to show respect for the patient's (former) autonomy. But we did not thereby endorse this position, and immediately went on to discuss various critiques of it. We also did not refer to any sort of 'residual' autonomy that is somehow possessed by a person who is now deemed incapacitated; instead, throughout the article (and in our abstract), we refer to patients' *former* autonomy as the relevant concept. With those clarifications in mind, we can now turn to the substance of Annoni's critique.

¹Centre for Biomedical Ethics, National University of Singapore, Singapore

²Uehiro Oxford Institute, University of Oxford, Oxford, UK

³Center for Advanced Studies in Bioscience Innovation Law (CeBIL), Faculty of Law, University of Copenhagen, Copenhagen, Denmark

⁴Faculty of Law, University of Oxford, Oxford, UK

⁵Department of Paediatrics, Monash University Faculty of Medicine Nursing and Health Sciences, Melbourne, Victoria, Australia

⁶Institute for Ethics, History and Philosophy of Medicine, Hannover Medical School, Hannover, Germany

⁷Julius Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

⁸Chair of Social Ethics, University of Bonn, Bonn, Germany

⁹Department of Philosophy and Kenan Institute for Ethics, Duke University, Durham, North Carolina, USA

¹⁰Clinical Center Department of Bioethics, National Institutes of Health, Bethesda, Maryland, USA

Correspondence to A/Prof Brian D Earp; bdearp@nus.edu.sg

ANNONI'S CRITIQUE: AUTONOMY AND SUBSTITUTED JUDGEMENT

Annoni's critique centres on challenging two 'dominant assumptions in the current ethical debate over the use of algorithmic Personalized Patient Preference Predictors (P4) in substitute judgment for incapacitated patients'. These two allegedly dominant assumptions, as summarised in the abstract of his article, are, first, that 'the autonomy of a patient who no longer has decision-making capacity can be meaningfully respected through a P4-empowered substitute judgement', and, second, 'that respect for autonomy can be reduced to merely satisfying a patient's individual treatment preferences'. We agree that the second claim, already mentioned above, is implausible and we are not aware of anyone who seriously argues for it. So, we will focus on the first claim in what follows.

What would it mean for a P4-empowered substituted judgement to 'meaningfully respect' the autonomy of a patient who no longer has decision-making capacity? By definition, it cannot mean respecting the current autonomy of someone who is not currently autonomous. We did not argue for such a notion in our paper. Instead, we discussed how the P4 could help surrogates to infer the underlying values and preferences held by someone who is now non-autonomous, *when they were last autonomous*. And we suggested that choosing the treatment that best reflects those competently held or endorsed values and preferences, as predicted by the P4, could be one way of showing respect for the person by virtue of the autonomy they once had.

While we did not explicitly invoke the so-called 'endorsed life' view (developed by Phillips and Wendler⁶ and summarised by Annoni), our proposal shares with it the core idea that respecting a formerly autonomous person can *include* (but does not necessarily *require*) making decisions aligned with their previously expressed values and preferences. It is this concept of respecting autonomy through substituted judgement that Annoni specifically critiques—using the endorsed life approach as an example of how a substituted judgement standard might be specified—arguing that it extends the notion of autonomy too far. He suggests that respecting autonomy requires direct, contemporaneous choice, and claims that any weaker

form of 'autonomy respect' would lead to troubling consequences.

ADDRESSING ANNONI'S CRITIQUE

Annoni's argument against substituted judgement as a means of respecting autonomy can be summarised as follows:

- A. If this 'weaker' sense of respecting autonomy applies to incapacitated patients, then it must also apply to patients deemed competent; and
- B. If it applies to competent patients, it would justify deceptive (or inappropriately paternalistic) practices when those practices align with patients' presumed values (eg, to use Annoni's case, administering a placebo to a competent patient who does not consent to this, and without their knowledge, based on a reasonable assumption that this is what they would prefer).

However, even if A were true, B would not follow. The flaw in Annoni's reasoning, as we see it, is that he overlooks how respecting autonomy involves different types of consideration depending on the patient's circumstances. For competent patients, their current choices have clear priority—and deceptive practices directly undermine these choices, regardless of how well they might align with the patient's broader values, whether actual or assumed.

Informed consent is the classic mechanism for respecting patient autonomy in competent individuals: it allows patients to make healthcare decisions based on adequate information about risks and benefits. Yet clearly, this mechanism cannot function for incapacitated patients. The impossibility of obtaining informed consent in such cases, however, does not logically entail that an incapacitated patient's autonomy cannot be respected in any meaningful way. For incapacitated patients, when direct autonomous choice is no longer possible, we must look to other expressions of their autonomy: first to specific advance directives when available, and only then to their more general values and preferences (however these may be evidenced or inferred).

This approach does not create a double standard but rather acknowledges that respecting autonomy requires different strategies in different contexts. The P4 aims to improve our ability to honour patients' values and preferences when direct autonomous choice is no longer an option, without suggesting that the same approach should override the choices of those who can still decide for themselves

(see a related discussion by some of us here).⁷

A MORE NUANCED UNDERSTANDING OF AUTONOMY

Annoni's restrictive view of autonomy is also counterintuitive and inconsistent with decades of bioethical thinking and practice. If respect for autonomy required only 'explicit, first-person, actual and specific consent' as Annoni suggests, then many widely accepted medical practices would be deemed autonomy-violating.

Consider advance directives themselves—these documents represent an extension of autonomous choice beyond the point of incapacity, yet cannot possibly specify every future medical situation. Similarly, broad consent for biobanking and longitudinal research, emergency treatment for unconscious patients and even prescription renewals for chronic conditions all operate without meeting Annoni's stringent standard of contemporaneous explicit consent. Yet bioethicists generally consider these practices consistent with respecting autonomy when properly implemented.

Rather than being all-or-nothing, respect for autonomy in healthcare exists on a spectrum. When a person lacks current decision-making capacity, ethical practice has long recognised that substituted decisions informed by the patient's competently held/endorsed values and preferences—while imperfect—remain more autonomy-respecting than decisions based solely on others' judgements of the patient's best interests.

The P4 aligns with this established understanding by potentially offering more accurate connections to the patient's authentic prior values than current surrogate decision-making processes. While traditional surrogates often misidentify patient preferences,^{8,9} the P4 draws directly from the patient's own recorded expressions and behaviours. This approach does not claim to replicate full autonomous decision-making, but rather to strengthen the connection between current medical decisions and the person's underlying values and preferences.

WHAT THE P4 PROPOSAL ACTUALLY ENTAILS

The proposed P4 aims to infer deeper values and preferences—not mere surface choices—using materials primarily produced by (or otherwise including information about) individuals themselves. This approach attempts to capture

the unique identity and agency of patients when direct autonomous choice is no longer possible. Crucially, we positioned P4 as a supplement within broader care structures, not as an autonomous decision-maker fully replacing human judgement. Our paper emphasised integration with advance care planning, surrogate involvement and clinical oversight.

CONCLUSION

When properly understood, the P4 represents a promising approach to enhancing substituted decision-making by more faithfully connecting medical decisions to patients' competently held/endorsed values and preferences. Whether conceptualised as respecting autonomy, honouring identity (or the individual's 'true self'),^{10 11} or reducing surrogate burden, the P4 aims to serve multiple ethical goods simultaneously.

We appreciate Annoni's contribution to clarifying the conceptual landscape around autonomy and substituted judgement. However, a more accurate understanding of our proposal reveals that P4 remains ethically desirable across multiple foundations, including not only honouring a patient's unique identity, but also respecting their former autonomy.¹²

X Brian D Earp @briandavidearp

Acknowledgements This research is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG3-GV-2023-012). WS-A is grateful for support from OpenAI and Duke University. This work was also supported in part by the Clinical Center Department of Bioethics, which is in the Intramural Program of the US National Institutes of Health. During the drafting of this paper, Claude 3.7 Sonnet and GPT-4.5 were used to edit, shorten and/or summarise text supplied by the authors. All ideas and arguments are the authors' own. Each author has made a substantial contribution to the work, which has been thoroughly vetted for accuracy and assumes responsibility for the integrity of their contributions (see ref 12).

Contributors BDE is the first and corresponding author. The co-first author is SPM. The co-senior authors are AR and DW. The middle authors are TvV, JWA, SS, KRJ, MB, WS-A and JS (co-corresponding author). After

initial drafting by BDE and SPM, all authors participated substantively in multiple rounds of revision and agreed to the final version.

Funding SPM's research was supported by a Novo Nordisk Foundation Grant for a scientifically independent International Collaborative Bioscience Innovation & Law Programme (Inter-CeBIL programme - grant no. NNF23SA0087056). BDE and JS are supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG3-GV-2023-012). WSA was funded by OpenAI and Duke University. AR and DW are funded by the Clinical Center Department of Bioethics.

Disclaimer The views expressed here are those of the authors and do not represent the position or policies of the National Institutes of Health or the US Department of Health and Human Services.

Competing interests We have an 'intellectual' competing interest in that we are responding to a critique of our idea.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; internally peer reviewed.



OPEN ACCESS

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

BDE and SPM are joint first authors.

AR and DW are joint senior authors.



To cite Earp BD, Porsdam Mann S, van Veenendaal T, et al. *J Med Ethics* 2025;**51**:462–464.

Received 10 January 2025

Accepted 6 April 2025

Published Online First 13 May 2025

J Med Ethics 2025;**51**:462–464.

doi:10.1136/jme-2025-110729

ORCID iDs

Brian D Earp <http://orcid.org/0000-0001-9691-2888>
Sebastian Porsdam Mann <http://orcid.org/0000-0002-1867-2097>

Tessa van Veenendaal <http://orcid.org/0009-0000-2386-3763>

Sabine Salloch <http://orcid.org/0000-0002-2987-2684>

Karin Rolanda Jongsma <http://orcid.org/0000-0001-8135-6786>

Matthias Braun <http://orcid.org/0000-0002-6687-6027>

REFERENCES

- 1 Earp BD, Porsdam Mann S, Allen J, et al. A Personalized Patient Preference Predictor for Substituted Judgments in Healthcare: Technically Feasible and Ethically Desirable. *Am J Bioeth* 2024;**24**:13–26.
- 2 Rid A, Wendler D. Use of a patient preference predictor to help make medical decisions for incapacitated patients. *J Med Philos* 2014;**39**:104–29.
- 3 Earp BD, Veenendaal T, Porsdam Mann S, et al. Digital psychological twins in medicine: addressing risks to human relationships. In: Chandra Y, Fan R, eds. *Artificial intelligence and the future of human relations: eastern and western perspectives*. Singapore: Springer Nature Singapore, 2025. Available: <https://www.academia.edu/126496115/>
- 4 Annoni M. It is not about autonomy: realigning the ethical debate on substitute judgement and AI preference predictors in healthcare. *J Med Ethics* 2025;**51**:450–5.
- 5 Earp BD, Hannikainen IR, Dale S, et al. Experimental philosophical bioethics, advance directives, and the true self in dementia. In: Hens K, De Block A, eds. *Advances in experimental philosophy of medicine*. Bloomsbury, 2023: 259–89.
- 6 Phillips J, Wendler D. Clarifying substituted judgement: the endorsed life approach. *J Med Ethics* 2015;**41**:723–30.
- 7 Earp BD, Demaree-Cotton J, Savulescu J. Against Externalism in Capacity Assessment-Why Apparently Harmful Treatment Refusals Should Not Be Decisive for Finding Patients Incompetent. *Am J Bioeth* 2022;**22**:65–70.
- 8 Shalowitz DI, Garrett-Mayer E, Wendler D. The accuracy of surrogate decision makers: a systematic review. *Arch Intern Med* 2006;**166**:493–7.
- 9 Ciroldi M, Cariou A, Adrie C, et al. Ability of family members to predict patient's consent to critical care research. *Intensive Care Med* 2007;**33**:807–13.
- 10 Toomey J, Lewis J, Hannikainen IR, et al. Advance Medical Decision-Making Differs Across First- and Third-Person Perspectives. *AJOB Empir Bioeth* 2024;**15**:237–45.
- 11 Toomey J, Lewis J, Hannikainen IR, et al. The True Self and Decision-Making Capacity. *Am J Bioeth* 2024;**24**:86–8.
- 12 Porsdam Mann S, Vazirani AA, Aboy M, et al. Guidelines for ethical use and acknowledgement of large language models in academic writing. *Nat Mach Intell* 2024;**6**:1272–4.