

Autonomy-based criticisms of the patient preference predictor

EJ Jardas , David Wasserman, David Wendler

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

Correspondence to

Dr David Wendler, Department
of Bioethics, National Institutes
of Health Clinical Center,
Bethesda, Maryland, USA;
dwendler@cc.nih.gov

Received 26 May 2021
Accepted 29 November 2021
Published Online First
17 December 2021

ABSTRACT

The patient preference predictor (PPP) is a proposed computer-based algorithm that would predict the treatment preferences of decisionally incapacitated patients. Incorporation of a PPP into the decision-making process has the potential to improve implementation of the substituted judgement standard by providing more accurate predictions of patients' treatment preferences than reliance on surrogates alone. Yet, critics argue that methods for making treatment decisions for incapacitated patients should be judged on a number of factors beyond simply providing them with the treatments they would have chosen for themselves. These factors include the extent to which the decision-making process recognises patients' freedom to choose and relies on evidence the patient themselves would take into account when making treatment decisions. These critics conclude that use of a PPP should be rejected on the grounds that it is inconsistent with these factors, especially as they relate to proper respect for patient autonomy. In this paper, we review and evaluate these criticisms. We argue that they do not provide reason to reject use of a PPP, thus supporting efforts to develop a full-scale PPP and to evaluate it in practice.

INTRODUCTION

Many patients, including up to 50% of older adults, have decisional incapacity in the sense that they are unable to make their own treatment decisions.¹⁻³ Patient appointed or legally designated surrogates, often a family member, work with the medical team to make treatment decisions for these patients. For example, it is estimated that up to 70% of treatment decisions near the end of life are made by the patient's surrogate.⁴

Surrogates are instructed to try to make the treatment decision the patient would have made if they had been able to make decisions. To implement this 'substituted judgement' standard, surrogates can sometimes appeal to preferences the patient documented in an advance directive or disclosed in conversation. Otherwise, as is true in the majority of cases, surrogates must rely on what they know about the patient to try to determine which decision they would have made in the circumstances.

Reliance on those who best know the patient to implement the substituted judgement standard is intended to maximise the chances that decisionally incapacitated patients receive the treatments they want and avoid the treatments they do not want. This approach also keeps the patient's family and loved ones involved in the process of making decisions regarding their care. Unfortunately, families and loved ones frequently do not know which treatment the patient would have chosen for themselves.

Empirical studies find that, when it is unclear which option is clinically preferable, patient appointed and next of kin surrogates accurately predict patients' treatment preferences only slightly more often than random guessing.⁵ Given the difficulties making these decisions, and the fact that they often involve life and death situations, it is not surprising that many surrogates experience significant distress as a result of making treatment decisions for incapacitated patients.^{6,7}

Clinicians increasingly rely on machine learning and artificial intelligence to address challenges in medical decision making. Consistent with this trend, several commentators have proposed to try to address the challenges with surrogate decision making by supplementing current practice with statistical algorithms which provide a prediction of the patient's preferred treatment. Several algorithms have been proposed in the literature, including the autonomy algorithm, an artificial intelligence based resuscitation algorithm and a patient preference predictor (PPP).⁸⁻¹⁰ These algorithms vary in a number of respects, including the specific information on which they base their predictions. The autonomy algorithm would train to make decisions based on electronic medical records. A resuscitation decision algorithm would rely on information gleaned from advance care planning discussions. The PPP would make predictions based on the sociodemographic characteristics of the patient and the extent to which they are correlated with specific treatment preferences. Despite these differences, these approaches use new technology, including artificial intelligence, to increase the chances that patients are treated consistent with their preferences. These approaches are thus susceptible to criticisms regarding respect for patient autonomy and how patients want treatment decisions to be made for them. The goal of the present paper is to assess the most common criticisms made in this regard. Because most of the criticisms take the PPP as their target, we too focus on the PPP, although most of the arguments apply to the other proposals as well.

The PPP is a proposed computer-based algorithm which would attempt to leverage the fact that patients' treatment preferences are correlated with many of their characteristics, including age, gender, religiousness, level of risk-taking and past treatment decisions.¹⁰ Determining precisely which characteristics are predictive of which treatment preferences will require additional research. The results of this research would then be used to develop a PPP, which would take the individual patient's characteristics as inputs and generate a prediction of which treatment the patient would want in the circumstances. Previous research found that a very



© Author(s) (or their
employer(s)) 2022. No
commercial re-use. See rights
and permissions. Published
by BMJ.

To cite: Jardas EJ,
Wasserman D, Wendler D.
J Med Ethics
2022;**48**:304-310.

preliminary PPP, one that incorporates only a few characteristics of the patient, predicts their treatment preferences as accurately as their surrogates.⁵ This finding suggests that a more refined PPP, one that incorporates a greater number of predictive characteristics, might yield more accurate predictions than surrogates alone. If it does, incorporating a PPP into the decision-making process could increase the chances that decisionally incapacitated patients receive the treatments they want and avoid the treatments they do not want.¹ Moreover, given that not knowing which treatment the patient would want is a significant source of stress for surrogates, this approach might also decrease their decisional burden. These considerations provide support for conducting the research necessary to identify predictive characteristics, using these data to develop a full-scale PPP, and testing it in the clinical setting to see whether it is accurate and feasible.¹¹

Some critics argue that use of a PPP is unlikely to increase surrogates' predictive accuracy and decrease surrogates' decisional burden. Others have argued that, even if use of a PPP achieves these goals, its use would still be problematic on other grounds, especially on the grounds that it fails to respect patient autonomy. This paper describes and then evaluates these objections. We conclude that they do not succeed and, therefore, do not provide reason to reject the use of a PPP or similar approaches for improving decision-making for decisionally incapacitated patients.

CRITICISMS OF THE PPP

Patients oppose the PPP

Some commentators argue that people do not support incorporating a PPP into the process of making treatment decisions for them in the event of decisional incapacity. For example, John characterises the use of a PPP as 'disturbing' (Rid, p170)¹⁰ and says many of his friends and colleagues agree.

Existing data suggest this is likely a minority view. In the largest survey conducted to date, nearly 80% of patients would want the PPP to be used to help make treatment decisions for them if its use increases surrogates' predictive accuracy.¹² Specifically, 54% of respondents would want their families to receive the PPP prediction, 18% would want to be treated according to the PPP prediction unless their family objects, and 7% would want to be treated according to the PPP prediction even if their family objects. Only 15% of respondents did not want the PPP to be used at all. Similarly, in a qualitative study, most physicians were open to the use of an artificial intelligence based resuscitation decision algorithm.⁸ These data suggest that patients' and doctors' preferences provide support for rather than an objection to the use of a PPP.

The family should decide, not the PPP

Some commentators argue that families have a right to make treatment decisions on behalf of their decisionally incapacitated loved ones. This right applies even when families are unable to accurately predict the patient's treatment preferences. Brock argues that families operate as an independent moral unit with

decision-making authority.¹³ He elaborates, 'The family's role in times of illness is not only to provide care and to help make decisions, but also to anchor the self and counter the alienation from self and body that serious illness can bring (Brock, p168).¹³ On this account, what matters most is not whether the specific treatments patients receive are consistent with their preferences, but whether, in times of dire need, their care is guided by loved ones, with whom they share the bond of family membership.

Some commentators support this approach by arguing that what is best for a patient cannot be determined independent of what is best for their loved ones. These arguments are based on accounts of autonomy¹⁴ or personhood¹⁵ which emphasise the personal significance of selflessness and family identity, and put less normative weight on independence: 'when one identifies with the interests of others and acts accordingly, one is also acting in one's self interest (Mappes, p32).¹⁶

One way to understand these arguments is in terms of the view that the family has a claim to make treatment decisions for decisionally incapacitated patients regardless of what the patient wanted. This version constitutes a rejection of the substituted judgement standard and would require a complete revision of current ethical and legal standards for surrogate decision making. Moreover, understood strictly, this view seems implausible. Imagine a 30-year-old patient who is temporarily unconscious due to a car accident, but is expected to live 50 more years with a good quality of life if they receive short-term treatment. We would not allow the patient's family to decline treatment for them based on their own preferences. At the very least, we would require compelling evidence that the patient would not have wanted the treatment, or that the patient endorsed whatever treatment decisions the family made.

A second way to understand these arguments maintains that patients care about more than simply which treatments they receive. They also care about the process by which treatment decisions are made for them. Specifically, many patients want their family to be involved in making treatment decisions if they develop decisional incapacity. While this is true for many patients, it is not true for all. In the survey cited previously, 7% of patients did not want their family involved in making treatment decisions for them.^{12 17} Other surveys find that up to one in four patients want their doctors, not their family, to make treatment decisions for them.¹⁸

Moreover, patients who want their families involved in treatment decision-making frequently explain this preference by citing the belief that their families know which treatments they want. Only a small proportion (21.8%) prioritise keeping their family involved when they don't know the patient's treatment preferences. There is also evidence that even this preference is influenced unconsciously by a belief that the family really does know the patient's treatment preferences.¹⁹ These data undermine claims that most patients inherently value their family's involvement when the family does not know the patient's treatment preferences.

Finally, inclusion of families and surrogates and use of algorithmic-based approaches are not mutually exclusive. Instead, these approaches could be incorporated into the decision-making process in ways that supplement rather than supplant surrogates. For example, the surrogate could be asked if they would like to know the PPP's prediction, the PPP prediction could be routinely provided to the surrogate, the patient could be treated based on the PPP's prediction unless the family objects (soft default), or the patient could be treated based on the PPP's prediction even if the family objects (hard default). At most, the argument that families should be involved in the

¹A full-scale PPP also might be used by individuals with decisional capacity. For example, an individual might explore what predictions the PPP would make for them and, in cases where the prediction diverges from their treatment preferences, note this fact in their advance directive. More generally, patients who are unsure of which treatment they want in a given situation might consider which treatment choice the PPP would predict for them. We thank an anonymous reviewer for encouraging us to include this point.

decision-making process is an objection to a hard default. It is not an objection to the first three ways in which the PPP might be implemented. Patient could also be given the opportunity to specify in advance whether and how they want the PPP or other algorithmic approaches used for them.

The PPP will not be more accurate than surrogates

Patient support for a PPP as indicated in prior surveys is based on the possibility that incorporating it into the decision-making process will increase the chances that they are treated consistent with their preferences. Some argue that use of a PPP or other algorithmic supplements are unlikely to achieve this goal for the simple reason that predicting treatment preferences is difficult. These individuals suggest that families may have reached the upper limit on how accurately we can predict the treatment preferences of decisionally incapacitated patients.²⁰

This criticism might be right, but it asks an empirical question. As far as we can tell, the best way to answer it would be to develop a full-scale PPP and test its accuracy in practice. During development, it will be important to monitor, assess, and update key performance indicators to ensure that a PPP is being used appropriately, including ensuring that predictive accuracy is being measured in a meaningful way. If the critics are right and use of a PPP does not increase predictive accuracy, that would provide compelling reason not to use it. But what if the PPP (or other algorithmic-based approach), *is* found to improve predictive accuracy? Are there other reasons not to use it? Some commentators answer in the affirmative.

The PPP assumes we lack the freedom to choose

Some commentators argue that the PPP assumes our personal characteristics determine the choices we make. This assumption fails to respect our autonomy by implying that we as agents are no more than the sum of our demographic characteristics, that we must choose as others in our demographic group choose, or that we lack the capacity to choose otherwise. For example, Sharadin²¹ contends that:

It is not true that agents' preferences are *determined* in this way by their demographic membership: although there are strong correlations between demographic characteristics and preferences, we are reluctant to say that the reason *why* we prefer (say) not to receive palliative care is *because* we are (say) straight, white, unmarried, college-educated 30-somethings....it can appear to be a violation of our capacity to have, form, and revise these preferences on the basis of reasons to treat our preferences as if they are determined by merely statistical information (Sharadin, p860).²¹

Sharadin faults the PPP for assuming that patients' choices are fully determined by their demographic factors, thereby precluding freedom of choice. This complaint is mistaken in two ways. The PPP does not assume that the demographic factors it employs, such as age, gender, education and socioeconomic status are (1) causes at all, let alone that they (2) fully determine the patient's choice. With respect to (1), the purpose of the PPP is not to identify the causes of patients' treatment preferences, but rather to predict, as accurately as possible, which treatment the patient prefers. Reliance on a demographic factor for predictive purposes does not assume that the characteristic causes a choice, only that it helps to predict that choice. For example, being born in a Midwestern state in the US vs a Northeastern one may be predictive of not wanting aggressive treatment at the end of life. But, this appeal to the patient's region of birth does not assume that it is a causal antecedent of their treatment preferences.

The PPP does not assume, for example, that the distance one was born from the ocean, or the severity of the winters one experienced when young, has a causal influence on one's preferences regarding end-of-life treatment. Instead, the PPP is based on the claim that the statistical association between place of birth and treatment preferences makes the former useful for predicting the latter. It may be that region of origin is merely correlated with other variables that are themselves causally relevant, but difficult to measure, such as a comparative reluctance to be a burden on others, which traces to diffuse cultural norms more prevalent in some areas of the country than others.

With respect to (2), the PPP does not assume that the causal factors underlying the associations it exploits fully determine the patient's choice. To continue with our example, knowing that a person prefers not to burden others does not mean that we know for certain which treatment decision they will make. Indeed, it is extremely unlikely that the PPP will be able to predict patients' treatment preferences with 100% accuracy, providing compelling reason to think the factors it inputs do not fully determine patients' treatment preferences.

The PPP relies on naked statistical evidence

Although the PPP does not assume that the variables it employs determine patients' treatment preferences, it has been argued that the PPP fails to respect autonomy because it predicts those preferences as if they were fully determined by the variable(s) it employs—it takes account of no other evidence in making its predictions. As Sharadin²¹ notes, this is one of the objections raised in the legal context to the use of 'naked statistical evidence' to decide a case—a reliance on nothing but the probability that a group to which the individual belongs acts or decides in a particular way.ⁱⁱ

Taking a standard example from the legal literature, Sharadin asks whether the fact that 99% of the blue-coloured buses in town are owned by the Blue Bus Company is enough to determine that the Company is legally liable for the recent death of a man who was run over by a blue-coloured bus. According to many legal scholars, that fact cannot be the only piece of evidence used to find the Company negligent.²² Even if the Blue Bus Company demonstrably caused 99% of all the bus accidents in town, that fact could not be the only piece of evidence used to find them negligent. In other words, liability should not be determined on the basis of a statistical piece of evidence alone, even when its probability, 99% in the present case, far exceeds the 51% probability generally thought to satisfy the 'preponderance of the evidence' standard in civil cases. Rather, the statistical evidence must be considered in light of additional, more individualised forms of evidence.

Exclusive reliance on the fact that the Blue Bus Company owns 99% of the buses is objectionable because it infers the defendant's responsibility for the negligent act from the proportional size of its bus fleet. It thereby fails to consider the possibility that the Company took steps to increase the safety of its buses, such as providing extended safety training, hiring safer drivers, or equipping its buses with more effective brakes, compared with the other bus companies in the city. Failing to take these factors into account fails to respect the autonomy of the Blue Bus Company. In effect, it denies the company's freedom to act better

ⁱⁱA similar objection is raised when the evidence is of the individual's own past 'base rate' of acting or deciding in a particular way, rather than the base rate of a group to which he belongs. Sharadin does not discuss the issue presented by individual base rates, nor will we.

(or worse) than the group average defined by the percentage of blue buses it owns.

There is debate over whether this line of reasoning constitutes a valid objection to the use of naked statistical evidence in the legal context. For present purposes, we need not enter that debate. Even if this is a valid objection in that context, there are several reasons to think it does not constitute a valid objection to the PPP.

First, the PPP is designed to input all available predictors regarding the patient's treatment preferences. It does not rely on a single statistic, no matter how strong that predictor might be, when other evidence is available, including more individualised evidence.

Second, and related to this, there are, as noted previously, at least four ways to implement the prediction issued by the PPP. Only the hard default approach directs clinicians to treat the patient based on the PPP's prediction independent of any other evidence, including evidence provided by the family. On the other approaches, evidence held by the family which suggests the patient differs from their demographic group is taken into account. Hence, arguments against relying on naked statistics provide reason to reject, at most, the hard default, not reason to reject the PPP itself.

Third, an important reason why statistical evidence is more objectionable when it is offered against the defendant is that it is being used to challenge the default outcome. In criminal cases, the default is a presumption of innocence, a presumption that should be set aside only when the evidence is sufficiently compelling. Even in civil cases, where less is at stake, the default is a presumption of the defendant's non-liability, a presumption that can be overcome only by a 'preponderance of evidence.'

The PPP is intended to be used in cases where it is not clear what the patient would have chosen and it is also not clear what is medically best for the patient. As recently as 50 years ago, clinical practice assumed a default of prolonging life in these cases. But, in the years since Quinlan and other 'right to refuse treatment' cases, this has ceased to be the default in all cases. Although some treatments, like cardiopulmonary resuscitation, are still considered the default unless there is clear reason to think the patient opposed them, most of the cases for which the PPP would be used are ones for which there is no preferred course of treatment. Consider a case in which a patient with moderate Alzheimer disease needs intubation and mechanical ventilation to survive an episode of pneumonia. In these cases, there is no presumption to be overcome and correspondingly no reason to restrict the kinds of evidence that can be 'admitted'. Instead, treatment decisions should be based on the patient's specific circumstances, as well as all the available evidence regarding their preferences and values.

This last point is underscored by the fact that the PPP may be most valuable, and its use least controversial, when there is no individualised evidence regarding the patient's treatment preferences. Most importantly, clinicians increasingly face the dilemma of determining how to treat decisionally incapacitated patients who left no evidence regarding their treatment preferences, and who have no friends or relatives to implement the substituted judgement standard on their behalf.^{23 24} It is in application to these cases of 'unrepresented' patients that the PPP would be the most 'naked', incorporating no individualised evidence beyond what is available from examination of the patient. Yet, this seems precisely the type of case where use of the PPP would be most valuable and least objectionable.²⁵ In sum, the PPP typically will not be used as a 'naked' statistic but, when it is, its use does not seem problematic.

The PPP uses non-endorsed reasons

The goal of the PPP is to maximise the chances that decisionally incapacitated patients receive the treatments they want and avoid the treatments they don't want. A final criticism argues that this focus on how the patient is treated misses an important element of respect for patient autonomy. Respecting patients' autonomy is not simply a matter of treating them in the ways they prefer to be treated. It is also important to make decisions for the right reasons, reasons the patients themselves endorse.^{26 27}

John illustrates the distinction between endorsable and non-endorsable reasons with the case of a woman, Jane, who is elderly and spent her life as an observant Catholic. Assuming both age and religion are predictive of patients' treatment preferences, the PPP would take both into account when predicting which treatment Jane would want if she loses decisional capacity. Jane, however, may regard her religion but not her age as a reason to choose a particular course of treatment. In contrast, a nominal Catholic who is active in American Association of Retired Persons might regard her age, but not her religious background, as a reason to choose a particular treatment.

The claim here is not simply that the individuals would not themselves take certain factors into account when making medical decisions. Instead, the point is that they oppose regarding certain characteristics as reasons for deciding which treatments they do and don't receive. As John puts it, these characteristics do not 'mesh' (John, p864)²⁷ with the patient's self-understanding. Hence, appealing to them when making treatment decisions for an incapacitated patient fails to respect their autonomy.

This criticism would be compelling if the goal of the PPP were to reason as the patient would have reasoned or to identify the best reasons for or against a particular treatment. However, the PPP does not reason and its goal is not to identify the reasons for or against a particular treatment. Taking the example mentioned previously, the PPP does not regard the fact that an individual was born in the Midwest as a reason not to provide them with aggressive medical care. Instead, the PPP regards the fact that an individual was born in the Midwest as evidence that they would not want to receive aggressive treatment.

John might respond that, in order to fully respect patient autonomy, the process of making decisions for them during periods of decisional incapacity should rely on only those *predictors* which the patient would endorse as reasons for preferring a particular treatment. This version of the criticism requires an argument that, in making decisions for decisionally incapacitated patients, surrogates should not simply try to maximise the chances that they receive the treatments they want and avoid the treatments they don't want. To fully respect patient autonomy, surrogates also should not employ predictors that the patient themselves would not regard as reasons for making treatment decisions. To assess this claim, it will be helpful to consider what exactly is involved in respecting individual autonomy.

A prominent view maintains that the capacity to act autonomously is important simply because it enables individuals to make their own decisions. On this view, respect for individual autonomy (liberty, self-determination, sovereignty) is important because it allows individuals to determine the course of their lives for themselves. However, use of the PPP is intended for patients who lack the capacity to determine the course of their lives. As a result, use of the PPP cannot fail to respect patient autonomy in this sense.

Granting that use of a PPP does not conflict with respect for autonomy in this narrow sense, some commentators endorse a broader conception of autonomy. They argue that the importance

of autonomy, and respect for it, are not limited to allowing individuals to actively make their own decisions. In addition, respect for individual autonomy is important because it increases the chances that the course of an individual's life is consistent with the preferences and values they endorse, thereby promoting what some have called the value of authenticity.^{28 29} Put in terms of the metaphor of the narrative of one's life, respect for autonomy not only permits individuals to write their own life narrative, it also increases the chances that the resulting narrative is consistent with their preferences and values. Importantly, respect for autonomy in this sense is relevant to making decisions for patients who have lost decisional capacity. While these patients cannot make their own decisions, it is possible to make decisions for them which are consistent (or not) with their own preferences and values regarding how their lives go.

To assess whether use of the PPP might conflict with respect for autonomy in this sense, we first need to determine which things qualify as aspects of the patient's life in the relevant sense. The metaphor of the narrative of a life can be misleading here because it seems to suggest an overly broad understanding. To take an extreme example, a truly comprehensive narrative of the lives of individuals who are born in Belgium today would include the fact that they are born in a country which was devastated during World War I. But that fact is not relevant to respecting these individuals' autonomy. Why not?

An obvious explanation is that promoting the authenticity of individuals' lives involves respecting their preferences and values for how their lives go once they have started—in this case, well after their birth country's devastation in World War I. While this account is more plausible, it is still too broad. In particular, individuals frequently have preferences regarding things that occur while they are alive, but which have nothing to do with their own lives, including preferences regarding the lives of others. This is the point of a frequently cited passage by Derek Parfit³⁰:

Suppose that I meet some stranger on a train. She describes her life's ambitions, and the hopes and fears with which she views her chances of success. By the end of our journey, my sympathy is aroused, and I strongly want this stranger to succeed. I have this strong desire even though I know that we shall never meet again (Parfit, p151).³⁰

Now imagine that a friend hears of this story and works hard to help the stranger succeed. Doing that might be nice for the stranger, and it might even show a kind of respect for Parfit's wishes. But, helping the stranger would not be a case of promoting or respecting Parfit's autonomy, even though he very much wants the stranger to succeed. Respecting the autonomy of others is not a matter of trying to bring about whatever states of affairs they endorse. It is a matter of allowing individuals to determine how their lives go based on their own preferences and values.

This example illustrates that simply caring about something, even caring about it deeply, does not imply that it concerns our lives in the sense that is relevant to respect for our autonomy. What, then, determines which things concern our lives in this sense? To begin to answer this question, consider how we would have to alter Parfit's story to make the stranger's success relevant to Parfit's life in the relevant sense. There are at least two options. Imagine that, in addition to caring about the stranger's success, Parfit actively works to help the stranger succeed. Or, imagine that, once the stranger succeeds, she uses her success to help Parfit. In these cases, the success of the stranger is relevant to how Parfit's life goes. This suggests, very generally, that there

are two aspects to the course of our lives: the impact that the world and others have on us and the impact we have on others and the world.³¹ The things we care about, then, become part of the course of our lives when we act to help to bring them about or their realisation influences us personally.

This understanding of respect for autonomy is consistent with the standard view that treating decisionally incapacitated patients in the ways they want to be treated is a matter of respecting their autonomy. Even though the treatment decisions are not made by the patient themselves, the treatments are given to them and can be more or less consistent with the life narrative they wanted for themselves. Inserting a breathing tube, for example, influences what happens to patients in ways they may or may not have endorsed. Notice, in contrast, that the *reasons* why the surrogate chose this course of treatment for the patient do not themselves become a part of the patient's life narrative. They are not things that the patient does or decides, nor are they things that happen to the patient. The former point seems clear, but the latter may be less so. The impact on the patient's life—what happens to them—is determined by the nature of the treatment, not by the reasons it was chosen. Hence, the reasons why the surrogate chooses a particular course of treatment are not relevant to respecting the patient's autonomy, even in this broader sense.

To further illustrate this point, consider the following example. As the result of a serious car accident, Tilquin is rendered unconscious. Her surrogate Suarez must decide whether Tilquin will receive aggressive treatment in the intensive care unit or supportive care until death. It is unclear which option makes sense medically and Suarez is unsure which option Tilquin would want to receive. He also knows that Tilquin was not proud of her place of birth and did not want it to be considered when making decisions on her behalf. The technician informs Suarez that if they take into account where Tilquin was born, there is a 75% chance the treatment choice the PPP outputs will match the treatment Tilquin would want to receive. If they don't input this information, there is a 55% chance the prediction will match the treatment Tilquin would want to receive.

Telling the technician Tilquin's place of birth significantly increases the chances that her life continues in the way she prefers and she receives the treatment she wants to receive. Thus, if the sole goal of surrogate decision-making is to promote patient autonomy, Suarez should tell the technician. At the same time, this involves a failure to respect Tilquin's wish that her place of birth not be taken into account when making decisions for her. Suarez thus faces a dilemma. Granting that, it is important to note that this dilemma does not involve how best to respect Tilquin's autonomy. Instead, it involves weighing the importance of respecting Tilquin's autonomy against the importance of respecting her wishes regarding how her surrogate makes treatment decisions for her.

In most cases, respecting patient autonomy seems more important than respecting their preferences regarding the decision-making process. But there will be exceptions, especially when the patient has strong preferences for how treatment decisions are made for them in the event of decisional incapacity. For example, some patients prioritise their family being involved in the decision-making process, even when their involvement decreases the chances they receive the treatments they want, and avoid the treatments they don't want. Analogously, if the patient did not want the PPP to be used in the process of making treatment decisions for them, that would provide an important reason against using it. Not using the PPP in these cases would show respect for the patient's wishes. This raises an empirical

question: what do we know about patient's preferences in this regard?

Existing data suggest most patients prioritise receiving the treatments they prefer and avoiding the treatments they don't want, not the process by which decisions are made for them.³² These patients' preferences regarding the process of making treatment decisions for them are informed by promoting this goal. This is why a strong majority of patients endorse using the PPP if its use increases the chances they receive the treatments they want and avoid the treatments they don't want. It is difficult to see how the use of the PPP would fail to respect the autonomy of these patients in any sense.

To see this, imagine you and your spouse become contestants on a new game show. Your spouse is presented with a series of challenging medical scenarios (eg, moderate Alzheimer disease and pneumonia requiring temporary intubation) and asked to guess your treatment preference in each scenario. If she gets them all right, you win a million dollars. If she gets any wrong, you receive nothing. How would you want your spouse to proceed? Would you want them to try to reason as you would reason, or would you rather they predict as accurately as possible, even if that involves appealing to factors you yourself do not regard as reasons for making medical decisions? We assume the vast majority of people would simply want the spouse to use whatever process is most likely to achieve the result that the contestant wants, in this case, winning the million dollars. In particular, many people are likely to think that getting all the questions correct is what matters and there is no independent value in how the decisions are made. Similarly, many patients prioritise receiving treatments they want and avoiding treatments they do not want. In most cases, patients do not have independent preferences for how their surrogates make treatment decisions for them in the event of decisional incapacity. In contrast, the fact that a patient opposed the use of the PPP would provide at least some reason not to use it, although not an autonomy based reasonⁱⁱⁱ.

CONCLUSION

We have reviewed six prominent criticisms of the PPP. We have found that none of these criticisms succeeds. To appreciate why autonomy-based criticisms of the PPP do not succeed, it is important to distinguish between predicting and predetermining, and between accurate predictions and appropriate reasons. The PPP does not predetermine that people are bound to the average preferences of their demographic group; it simply uses demographic information as correlative evidence for or against a choice. Moreover, the PPP relies on predictors, not reasons, and ultimately predictors need not be endorsed. While relying on non-endorsed predictors may fail to respect a patient's wishes regarding the decision-making process, they will better respect the patient's autonomy, by increasing the chances that the course of their life accords with their own preferences and values. Furthermore, use of the PPP differs from reliance on naked statistical evidence in important ways. The incorporation

ⁱⁱⁱAs previously noted, patients with strong preferences regarding the use of certain predictors could leave instructions regarding the use of the PPP. As the predictive technology becomes more sophisticated, individual patients may even be able to train and fine tune their own personalised sets of predictors. While appealing in principle, this approach would encounter the same challenges as current efforts to encourage patients to complete advance directives: only a small proportion of patients engage in advance care planning.

of a large number of predictors limits concerns about any single predictor. In addition, the PPP is used for the benefit of the patient and offers the most benefit in cases where other sources of evidence are unavailable. This analysis, though focused on the PPP, would also apply to artificial intelligence-based preference prediction. It supports proposals to develop a full-scale PPP and similar approaches, and to evaluate in practice whether they increase the chances that patients are treated consistent with their preferences and values and decrease the emotional burden on their surrogates.

Contributors EJJ wrote the first draft and subsequent drafts of the manuscript; DWen provided substantial comments and revisions on subsequent drafts; DWas provided comments and revisions on multiple drafts; all authors participated in the concept and arguments herein, approved the final manuscript as submitted, and agree to be accountable for all aspects of the work.

Funding This work was funded by: U.S. Department of Health and Human Services, National Institutes of Health, NIH Clinical Centre, Not applicable (Intramural Program/Programme).

Disclaimer The views expressed are the authors' own. They do not represent the position or policy of the NIH or DHHS.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement There are no data in this work.

ORCID iD

EJ Jardas <http://orcid.org/0000-0002-6168-6398>

REFERENCES

- 1 Sorger BM, Rosenfeld B, Pessin H, *et al*. Decision-making capacity in elderly, terminally ill patients with cancer. *Behav Sci Law* 2007;25(3):393–404.
- 2 Lepping P, Stanly T, Turner J. Systematic review on the prevalence of lack of capacity in medical and psychiatric settings. *Clin Med* 2015;15(4):337–43.
- 3 John S, Rowley J, Bartlett K. Assessing patients decision-making capacity in the hospital setting: a literature review. *Aust J Rural Health* 2020;28(2):141–8.
- 4 Silveira MJ, Kim SYH, Langa KM. Advance directives and outcomes of surrogate decision making before death. *N Engl J Med* 2010;362(13):1211–8.
- 5 Shalowitz DL, Garrett-Mayer E, Wendler D. The accuracy of surrogate decision makers: a systematic review. *Arch Intern Med* 2006;166(5):493–7.
- 6 Wendler D, Rid A. Systematic review: the effect on surrogates of making treatment decisions for others. *Ann Intern Med* 2011;154(5):336–46.
- 7 Su Y, Yuki M, Hirayama K. The experiences and perspectives of family surrogate decision-makers: a systematic review of qualitative studies. *Patient Educ Couns* 2020;103(6):1070–81.
- 8 Biller-Andorno N, Ferrario A, Joebges S, *et al*. AI support for ethical decision-making around resuscitation: proceed with care. *J Med Ethics* 2022;48(3):175–83.
- 9 Lamanna C, Byrne L. Should artificial intelligence augment medical decision making? The case for an autonomy algorithm. *AMA J Ethics* 2018;20(9):E902–10.
- 10 Rid A, Wendler D. Use of a patient preference predictor to help make medical decisions for incapacitated patients. *J Med Philos* 2014;39(2):104–29.
- 11 Rid A, Wendler D. Treatment decision making for incapacitated patients: is development and use of a patient preference predictor feasible? *J Med Philos* 2014;39(2):130–52.
- 12 Wendler D, Wesley B, Pavlick M, *et al*. A new method for making treatment decisions for incapacitated patients: what do patients think about the use of a patient preference predictor? *J Med Ethics* 2016;42(4):235–41.
- 13 Brock DW. What is the moral authority of family members to act as surrogates for incompetent patients? *Milbank Q* 1996;74(4):599.
- 14 Igel LH, Lerner BH. Moving past individual and 'pure' autonomy: the rise of family-centered patient care. *AMA J Ethics* 2016;18(1):56–62.
- 15 Berger JT, DeRenzo EG, Schwartz J. Surrogate decision making: reconciling ethical theory and clinical practice. *Ann Intern Med* 2008;149(1):48.
- 16 Mappes TA, Zembaty JS. Patient choices, family interests, and physician obligations. *Kennedy Inst Ethics J* 1994;4(1):27–46.
- 17 Wendler D, Wesley R, Pavlick M, *et al*. Do patients want their families or their doctors to make treatment decisions in the event of incapacity, and why? *AJOB Empir Bioeth* 2016;7(4):251–9.
- 18 Daveson BA, Bausewein C, Murtagh FEM, *et al*. To be involved or not to be involved: a survey of public preferences for self-involvement in decision-making involving mental capacity (competency) within Europe. *Palliat Med* 2013;27(5):418–27.

- 19 Jardas EJ, Wesley R, Pavlick M, *et al.* Patients' priorities for surrogate decision-making: possible influence of misinformed beliefs. *AJOB Empirical Bioethics* 2021;18(3):1–15.
- 20 SYH K. Improving medical decisions for incapacitated persons: Does focusing on 'accurate predictions' lead to an inaccurate picture? *J Med Philos* 2014;39:187–95.
- 21 Sharadin NP. Patient preference predictors and the problem of naked statistical evidence. *J Med Ethics* 2018;44(12):857–62.
- 22 Gardiner G. Legal burdens of proof and statistical evidence. In: *Routledge Handbook of applied Epistemology*, 2018: 179–95.
- 23 Pope TM, Bennett J, Carson SS. Making medical treatment decisions for unrepresented patients in the ICU. *Am J Respir Crit Care Med* 2020;201:1182–92.
- 24 White DB, Curtis JR, Lo B, *et al.* Decisions to limit life-sustaining treatment for critically ill patients who lack both decision-making capacity and surrogate decision-makers. *Crit Care Med* 2006;34(8):2053–9.
- 25 Varma S, Wendler D. Medical decision making for patients without surrogates. *Arch Intern Med* 2007;167(16):1711–5.
- 26 John S. Patient preference predictors, apt categorization, and respect for autonomy. *J Med Philos* 2014;39(2):169–77.
- 27 John SD. Messy autonomy: commentary on patient preference predictors and the problem of naked statistical evidence. *J Med Ethics* 2018;44(12):864.
- 28 Brudney D, Lantos J. Agency and authenticity: which value grounds patient choice? *Theor Med Bioeth* 2011;32(4):217–27.
- 29 Enoch D. Hypothetical consent and the value (s) of autonomy. *Ethics* 2017;128(1):6–36.
- 30 Parfit D. *Reasons and persons*. Oxfordshire, UK: Clarendon Press, 1987.
- 31 Phillips J, Wendler D. Clarifying and defending the endorsed life approach to surrogate decision-making. *J Med Ethics* 2015;41(9):736–8.
- 32 Rid A, Wesley R, Pavlick M, *et al.* Patients' priorities for treatment decision making during periods of incapacity: quantitative survey. *Palliat Support Care* 2015;13(5):1165–83.