**Patient values and inductive risk in disorders of consciousness**

Mona-Marie Wandrey

mmw54@cam.ac.uk

Forthcoming in the *European Journal for Philosophy of Science.*

Please cite the published version when available.

**Abstract**

Diagnosing patients with disorders of consciousness involves inductive risk: the risk of false negative and false positive results when gathering and interpreting evidence of consciousness. A recent proposal suggests mitigating that risk by incorporating patient values into methodological choices at the level of individual diagnostic techniques: when using machine-learning algorithms to detect neural evidence of responsiveness to commands, clinicians should consider the patient’s own preferences about whether avoiding false positives or false negatives takes priority (Birch, 2023). In this paper, I argue that this proposal raises concerns about how to ensure that inevitable non-epistemic value judgments do not outweigh epistemic considerations. Additionally, it comes with challenges related to the predictive accuracy of surrogate decision-makers and the decisional burden imposed on them. Hence, I argue that patient values should not be incorporated at the level of gathering evidence of consciousness, but that they should play the leading role when considering how to respond to that evidence.

**Keywords:** Inductive Risk, Values in Science, Disorders of Consciousness, Decision-Making under Uncertainty, Surrogate Decision-Making, Patient Autonomy

1. **Introduction**

Gathering and interpreting scientific evidence inevitably faces the problem of inductive risk: the risk of accepting a hypothesis that is, in fact, false (false positive), or rejecting a hypothesis that is, in fact, true (false negative). There is an overarching consensus among philosophers of science that scientists necessarily need to make non-epistemic value judgments when dealing with inductive risk: they need to consider the consequences of making a certain type of error, for example based on ethical, societal or economic considerations (Douglas, 2000; Rudner, 1953). However, it remains contentious at what point in the research process the influence of non-epistemic values is acceptable and what specific values researchers should employ (Holman & Wilholt, 2022; Schroeder, 2024).

 This paper focuses on inductive risk in the context of diagnosing patients with disorders of consciousness (DoC), such as coma, unresponsive wakefulness syndrome and minimally conscious state. Arguably, the problem of inductive risk is particularly pronounced in this context, as the evidence base for determining whether patients retain consciousness or not is fraught with uncertainty, and the ethical consequences of making false negative and false positive errors (i.e. diagnosing patients as unconscious, who are, in fact, conscious, and vice versa) are severe (Peterson et al., 2015). The introduction of novel neuroimaging and electrophysiological methods into the diagnostic process does not alleviate the problem of inductive risk. On the contrary, when determining, for example, whether a certain pattern of neural activity in an electroencephalogram (EEG) indicates responsiveness to commands, clinicians need to choose statistical methods that differentiate actual responsiveness from mere artefacts. Different statistical procedures can significantly influence the result of task-based EEG paradigms, with more conservative methods increasing the risk of false negative and more permissive methods increasing the risk of false positive results (Birch, 2023; Peterson et al., 2015).

The current European guideline on diagnosing DoC patients already recommends the supplementary use of functional neuroimaging and electrophysiological techniques whenever feasible, and these techniques are likely to play an even more important role in the future, as the use of machine learning algorithms for analyzing neural activity patterns becomes increasingly widespread (Kondziella et al., 2020). This emphasizes the need for developing strategies that address inductive risk in the DoC context. How should clinicians and decision-makers deal with the risk of false positive and false negative results, and at what stage of the diagnostic process do non-epistemic values play a legitimate role?

Philosophers of science broadly distinguish four stages of the scientific process where non-epistemic values might come into play: (1) steering science, (2) managing science, (3) doing science, and (4) using science (Elliott, 2022). Non-epistemic values clearly affect the stages of steering science (e.g. when selecting research questions or allocating funding), managing science (e.g. when developing ethical guidelines for handling data), and using science (e.g. when developing public policy based on scientific evidence). However, proponents of the value-free ideal in science argue that the process of conducting scientific research itself, i.e. gathering evidence and drawing conclusions from it (sometimes referred to as the “core part” of science), should be kept free from the influence of non-epistemic values (Elliott, 2022; Schroeder, 2024).

Yet, most philosophers of science acknowledge that the value-free ideal is not tenable, as the problem of inductive risk entails that even during the core part of scientific research non-epistemic value judgments are inevitable (Douglas, 2000). Nonetheless, there is a general consensus that non-epistemic values should play a secondary role to epistemic considerations during the process of gathering and interpreting evidence, and that scientists should make efforts to delineate legitimate from illegitimate value influences (Holman & Wilholt, 2022).

In this paper, I analyze how non-epistemic values affect the process of diagnosing and treating DoC patients, focusing on the stages of (1) gathering evidence of consciousness with individual diagnostic methods, such as behavioral examinations and neural techniques, (2) interpreting the evidence from these methods to determine an overall diagnosis, and (3) making further clinical decisions based on that diagnosis.

 So far, discussions of inductive risk in the DoC context have assumed a relatively traditional distinction between the stages of gathering and interpreting evidence (where epistemic considerations play the leading role, and non-epistemic value judgments, if made explicit at all, are secondary and reflect researchers’ own values) and the stage of making decisions based on that evidence (where non-epistemic values, particularly the patient’s own values, play the leading role) (Johnson, 2021; Peterson et al., 2015).

Recently, however, Jonathan Birch has advanced a proposal for addressing inductive risk in the DoC context that fundamentally diverges from this picture (2023). Birch proposes that, as value judgments are inevitable even at the stage of gathering and interpreting evidence, the patient’s own values should explicitly guide methodological choices at this stage. When assessing neural evidence of responsiveness in task-based EEG-paradigms with machine-learning algorithms, he argues for the use of Bayesian methods, proposing that Bayesian priors should be informed by the patient’s own attitudes towards the risks of false positive/negative results. While priors should be set using epistemic considerations as a starting point, with only minor adjustments to reflect patient values, this proposal still implies a fundamental departure from the view that the process of gathering and interpreting evidence should be kept as free from non-epistemic values as possible.

The remainder of this paper examines the advantages and disadvantages of this proposal, guided by the question which role patient values should play in the process of diagnosing and treating DoC patients. I will argue that implementing patient values at the stage of gathering evidence carries risks of its own, which have not previously been acknowledged. First, it raises concerns about how to delineate legitimate from illegitimate non-epistemic value influences in methodological decisions. Second, it complicates issues associated with surrogate decision-making. Surrogates might not be accurate predictors of patients’ values, potentially leading to methodological choices that are misaligned with the patient’s preferences. Additionally, involving surrogates into methodological choices risks increasing their decisional burden, as their assessment of the patient’s values would have an impact on the likelihood of detecting evidence of responsiveness in that patient.

Therefore, I argue that patient values should not be directly incorporated into methodological choices at the stage of gathering evidence. However, patient values play an important role when making decisions under uncertainty. Hence, I sketch a counterproposal for managing inductive risk in the DoC context that is based on assessing and communicating diagnostic uncertainty in a graded way and taking the patient’s own precautionary attitudes into account when considering how to respond to uncertain evidence.

The paper proceeds as follows: in section 2, I give a brief overview on the sources of inductive risk in DoC diagnosis. In section 3, I outline Birch’s proposal for implementing patient values into the process of gathering evidence and show why I regard some aspects of it as problematic. In section 4, I present my counterproposal. Section 5 concludes.

1. **Sources of inductive risk in disorders of consciousness**

As mentioned, diagnosing DoC patients is particularly affected by the problem of inductive risk. In what follows, I briefly outline how DoC patients are diagnosed and highlight sources of uncertainty in the diagnostic process.

DoC are conditions that are behaviorally characterized by loss or impairment of consciousness. Consciousness is clinically defined as consisting of two components, arousal (i.e. wakefulness) and awareness (i.e. subjective experience of self and environment) (Posner, 2019). While arousal can be detected by the presence of eye-opening and sleep-wake cycles, determining whether a patient is aware is more difficult. Awareness is fundamentally a subjective phenomenon, characterized by the fact that there is something it is like to be an experiencing subject from the first-person point of view (Nagel, 1974).[[1]](#footnote-1) As such, it cannot be measured directly, but has to be inferred by the presence of indirect markers of consciousness, such as introspective report and intentional agency. In the DoC context, patients cannot report their own conscious states, so intentional agency (i.e. the ability to produce purposeful, goal-directed behaviors as opposed to mere stimulus-driven, reflexive actions) is usually regarded as the most specific available marker of consciousness (Shea & Bayne, 2010).

 The standard approach for diagnosing DoC patients relies on a detailed neurobehavioral examination in which the clinician looks for signs of awareness in the form of purposeful behavioral responses to sensory stimuli (Giacino et al., 2014). Current diagnostic guidelines recommend the use of standardized behavioral rating scales, such as the Coma Recovery Scale-Revised (Giacino et al., 2018; Kondziella et al., 2020). Based on the results of these bedside examinations, DoC are distinguished into coma, unresponsive wakefulness syndrome (UWS)[[2]](#footnote-2) and the minimally conscious state (MCS). In coma, there are no signs of arousal or awareness. In UWS, patients show eye-opening and sleep-wake cycles, but no behavioral signs of awareness. In MCS, on the other hand, there are inconsistent, but reproducible signs of awareness in the form of purposeful responses to sensory stimuli or command-following (Harrison & Connolly, 2013).

 However, the rate of behavioral misdiagnoses in DoC is extremely high. Patients might be unable to behaviorally indicate their awareness to outside observers for a variety of reasons, for example due to sensory or motor impairments, aphasia, pain, fluctuating arousal or fatigue (Harrison & Connolly, 2013). Comparing the results of standardized behavioral rating scales to an unsystematic neurological examination alone reveals that up to 40% of patients are misdiagnosed as being in the UWS when they actually retain subtle behavioral signs of awareness (Schnakers et al., 2009).

 The high rate of behavioral misdiagnoses has sparked research into neural markers of consciousness, i.e. functional neuroimaging and electrophysiological methods that aim at analyzing whether the patient’s brain activity indicates awareness. These techniques can be distinguished into active, passive and resting state paradigms. Active paradigms analyze the patient’s brain activity in response to commands to determine whether the elicited brain activity matches that of healthy conscious controls. In passive paradigms, neural activity is analyzed in response to sensory stimuli, examining whether processing of these stimuli is comparable to conscious controls. In resting state paradigms, the patient’s brain activity is analyzed at rest to investigate functional connectivity or metabolic activity in certain brain networks that are correlated with awareness in healthy controls (Harrison & Connolly, 2013).

The implementation of these techniques into the diagnostic process has the potential to alleviate the high rate of behavioral misdiagnoses. For example, the results of active paradigms reveal that up to 25% of patients behaviorally diagnosed as being in the UWS are able to modulate their brain activity in response to commands, indicating a state of “covert consciousness” or cognitive motor dissociation (CMD) (Bodien et al., 2024). The detection of CMD in the intensive care unit has been associated with a higher likelihood of functional recovery one year after brain injury, which is why it is regarded as an important prognostic factor (Claassen et al., 2019; Edlow et al., 2021).

 However, the employment of neural techniques does not necessarily reduce the inductive risk associated with diagnosing DoC. For example, as mentioned, the likelihood of detecting evidence of responsiveness to commands in active EEG-paradigms depends a lot on which statistical method is chosen to differentiate real responsiveness from artefacts (Birch, 2023; Peterson et al., 2015). Hence, at the level of gathering evidence, there is a risk that an individual diagnostic test does not accurately detect the phenomenon it is intended to measure. Additionally, there is considerable debate about the question to what extent the results of individual tests support the hypothesis that a patient is conscious. For instance, if the task-based EEG detects responsiveness, can we infer that the patient is conscious? And if the EEG does not detect responsiveness, can we infer that the patient is unconscious?

Usually, active paradigms are regarded as relatively specific markers of consciousness: detecting a neural activation pattern in response to a command seems to require that the patient is consciously aware of the command and chooses to participate (Owen et al., 2007). However, active paradigms are not regarded as sensitive: patients might fail to activate brain areas in response to commands for similar reasons they might be unable to react to commands at the bedside, which is why negative results in active paradigms do not rule out awareness (Harrison & Connolly, 2013).

Similar considerations apply to passive and resting state paradigms. On the one hand, there is uncertainty about whether a method accurately reflects neural activity. For example, functional magnetic resonance imaging (fMRI) relies on the fact that oxygenated and deoxygenated blood have different magnetic properties. Increased neural activity results in increased blood flow and oxygenation levels in certain brain areas, but this effect is most pronounced in cortical areas, which is why fMRI might not detect neural activity changes in areas where blood flow does not change as much (Hillman, 2014; Watson et al., 2010).

Additionally, there is deep uncertainty about which neural networks are associated with conscious experience: neural theories of consciousness range from regarding the prefrontal cortex as necessary for consciousness (Dehaene & Changeux, 2011) to considering midbrain mechanisms present in all vertebrates as sufficient (Merker, 2007). Thus, there might be conflicting views on whether to interpret the results of passive and resting state paradigms as evidence of conscious awareness or not.

 This brief overview about the diagnostic process emphasizes that the epistemic uncertainty in the DoC context is extraordinarily high, both at the level of gathering evidence of consciousness with individual diagnostic techniques and at the level of interpreting to what extent the evidence supports an overall diagnosis. This leads to high levels of inductive risk.[[3]](#footnote-3)

 The ethical consequences of making false negative and false positive errors in the process of diagnosing DoC patients are severe (Peterson et al., 2015). False negative errors (missing consciousness) might cause significant harm to the patient. Families and healthcare workers might emotionally and financially withdraw, leaving the patient aware but unable to communicate their awareness through any forms of behavior. Additionally, life-sustaining therapies might be withdrawn prematurely (Johnson, 2021). Presumably, most people would regard the magnitude of that risk to outweigh the risk associated with false positive errors (misattributing consciousness). Yet, false positive errors might lead surrogate decision makers to continue life-sustaining therapies and cost-intensive rehabilitation measures in the false hope that the patient will recover. This might cause financial and emotional distress for the patients’ families, a factor that plays a relevant role for many patients who engage in advance care planning (Peterson et al., 2015; Rid et al., 2015).

 It is important to stress that the values of patients and their families in these situations are highly individual. For some, consciousness might not be the most relevant factor that determines what kind of care they would like to receive in such a scenario, as they might hold various ethical or religious beliefs independent of consciousness (Edlow & Fins, 2018). Additionally, information about the likelihood of functional recovery might be more relevant than diagnostic evidence to some decision-makers in these scenarios. Yet, the patient’s diagnosis influences prognostic considerations and treatment decisions to a significant degree, which is why the risk of false negative and false positive errors needs to be managed appropriately.

 As mentioned, previous strategies for managing inductive risk in the DoC context assume a relatively traditional distinction between epistemic and non-epistemic values when managing that risk. For example, Andrew Peterson has suggested that clinicians should consider a variety of sources of evidence to ensure that the diagnostic process is as reliable as possible, thus focusing on epistemic strategies to reduce inductive risk (2016). Other authors have focused on ethical strategies to mitigate inductive risk. For example, Syd Johnson has argued that acknowledging and communicating epistemic uncertainty should play a key role in the DoC context: rather than focusing on unreliable diagnostic distinctions between conscious and unconscious states, clinicians and caregivers should acknowledge that all DoC patients are potentially conscious persons whose welfare and autonomy need to be respected (2021).

Without analyzing these proposals in detail, I want to highlight that they are broadly based on the assumption that patient values play a central role at the decision-making stage of treating DoC patients. When the clinician presents the diagnosis to surrogate decision-makers (acknowledging uncertainty), the patient’s own values should be the guiding factor for further decision-making. However, at the initial stages of gathering and interpreting evidence, the patient’s own values are not explicitly factored in. For example, although the choice of statistical procedure when analyzing EEG-data entails an implicit value judgment regarding whether missing or misattributing responsiveness is the more severe error to make, this judgment will usually reflect the clinicians’ or researchers’ own non-epistemic values (Peterson et al., 2015). That said, the leading consideration for which statistical procedure to choose is guided by epistemic values, for example considering which procedure yields results most closely aligned with conscious controls (Cruse et al., 2013).

Birch, on the other hand, argues that patient values should be reflected already at the stage of gathering evidence to mitigate inductive risk in the DoC context (2023). This proposal would have significant consequences if implemented into medical practice. To my knowledge, these implications have not been discussed previously. Hence, in the following section, I discuss the advantages and limitations of incorporating patient values into the diagnostic process. While I take Birch’s proposal as a starting point, my analysis does not hinge on the specific details of the diagnostic technique Birch uses as an example, but engages with the question which role patient values should play when managing inductive risk more broadly.

1. **Incorporating patient values at the stage of gathering evidence**

The previous section has shown that making errors seems unavoidable when diagnosing DoC patients. Birch argues that it should partly be the patient’s own values that guide methodological choices at the level of individual diagnostic techniques. Using a recent task-based EEG study that employs machine learning to analyze evidence of responsiveness to commands as an example (Claassen et al., 2019), Birch highlights that the statistical methods that were chosen to differentiate responsiveness from artefacts reflected researchers’ implicit value judgments about the risk of false positive/negative results. He argues that, while these value judgments are unavoidable, they can lead to problems “if the values in question are misaligned with the patient’s own values” (Birch, 2023, p. 2).

Birch therefore advances a proposal for mitigating the inductive risk involved in diagnostic techniques like these that shifts the focus to the patient’s own values when making methodological choices. His proposal is based on three key elements: first, analyzing evidence of responsiveness in task-based EEG paradigms should not be based on categorical yes/no verdicts, but instead shift towards assessing degrees of evidence, employing Bayesian methods. Second, these degrees of evidence should be converted to probabilistic assessments (i.e. the probability that a patient responds to a command, given the evidence from the EEG). This requires considering prior probabilities. Bayesian priors should be set in a *patient-centered* way: they should be anchored in the highest realistic baseline estimates of covert consciousness, with small deviations that reflect not only the clinician’s expert judgment about the patient’s overall diagnosis, but also the patient’s own values regarding the risk of missing or misattributing responsiveness. Third, these probabilistic assessments should be communicated to caregivers and decision-makers as clearly as possible, using standardized language (Birch, 2023).

 Birch points out that the risk of incorporating value judgments that are misaligned with the patient’s own values is not unique to the machine learning technique in the DoC context described above, but might also arise in many other clinical settings where AI algorithms are employed to analyze data and give categorical yes/no verdicts (Birch et al., 2022). Hence, it seems essential to ensure that these techniques are as sensitive to the patient’s own values and risk attitudes as possible.

 Birch’s proposal is a promising strategy for managing uncertainty in the DoC context. Especially the emphasis on probabilistic assessments instead of categorical verdicts, as well as the suggestion to communicate these probabilities in a standardized way, add important nuance to the question of how to address diagnostic uncertainty. However, while I agree that it is essential to consider patient values when managing inductive risk, I will argue that incorporating patient values at the level of gathering evidence (such as when setting Bayesian priors to assess the probability of responsiveness from EEG data) carries risks of its own. First, it raises concerns about how to ensure that epistemic considerations remain paramount at this stage of the diagnostic process. Second, it risks complicating issues associated with surrogate decision-making, as surrogates might not predict patients’ values accurately, and the burden of responsibility imposed on them might increase if their assessment of patients’ values would affect the probability of detecting evidence of responsiveness to a small degree. In what follows, I explore these risks in more detail.

Regarding the first point, recall that there is a general consensus in the values in science literature that, during the process of gathering and interpreting scientific evidence, non-epistemic values should play a secondary role to epistemic considerations (Schroeder, 2024). However, what exactly that means is controversial. How much influence can non-epistemic values have without outweighing epistemic considerations?

Bennett Holman and Torsten Wilholt refer to the question of how to delineate legitimate from illegitimate non-epistemic value influences in science as the “new demarcation problem” (2022). In their view, the demise of the value-free ideal in science does not entail that “anything goes”, as there seems to be an important line between the inevitable management of inductive risk and an illegitimate distortion of scientific inquiry (Holman & Wilholt, 2022, p. 211). There are a variety of strategies for how to draw that line. For example, Heather Douglas argues that non-epistemic values should only play an *indirect* role at the core part of science: they should not serve as direct reasons to accept or refute hypotheses, but can come into play when there is no clear way to appeal to epistemic values to resolve decisions (2009). While it would be inappropriate for researchers to endorse a hypothesis because they want it to be true or because it aligns most closely with their preferred policy outcome, it would be appropriate to appeal to non-epistemic values to determine how much evidence is needed to accept a hypothesis. For example, the decision whether to classify borderline cases in pathology specimens as cancerous or not cannot be resolved by appealing to epistemic values alone, which is why non-epistemic values play a legitimate role at this stage (Schroeder, 2024).

In what follows, I analyze how the proposal to integrate patient values into the diagnostic process can address the new demarcation problem: how much influence could patient values legitimately have when making methodological choices, while ensuring that epistemic considerations remain paramount? Staying with the example of using patient-centered priors to convert evidence of responsiveness into probabilistic assessments in EEG paradigms, I think it would be difficult to ensure that non-epistemic value judgments do not outweigh epistemic considerations in an illegitimate way. When setting Bayesian priors, how much deviation from the epistemic anchoring point (baseline rate of covert consciousness) would be appropriate to reflect the patient’s individual values? Birch himself highlights the necessity to “guard against the risk of dogmatic priors”, e.g. priors that are intentionally set so low that even the strongest evidence of responsiveness in an EEG would not yield a high posterior probability of the patient responding (2023, p. 6). One could also imagine the opposite scenario where priors are set so high that even weak evidence of responsiveness would yield a high probability that the patient is responding. These two scenarios might be regarded as examples of illegitimate value judgments shaping methodological choices.

However, there are many possible scenarios where judging whether priors are influenced by non-epistemic values in a legitimate or illegitimate way is less clear. For example, if the clinician assumes, based on conversations with the patient’s family, that a patient would have regarded a false negative diagnosis as much worse than a false positive diagnosis, but the patient is behaviorally unresponsive at the bedside and her brain injury is extensive, so that the clinician regards the prior probability of her responding to commands as low, how much deviation from the baseline estimate would be appropriate to reflect the patient’s own values? Birch leaves this question open on purpose, as his hope is to start a debate on this issue rather than presenting a definitive solution. He suggests that diagnostic guidelines should recommend a range for appropriate priors, using realistic baseline estimates of covert awareness as an anchoring point and precluding the use of dogmatic priors.

However, the widespread theoretical disagreement about the neural prerequisites for consciousness entails that reaching consensus about the appropriate range of priors might be difficult. Priors that some clinicians would regard as realistic might be regarded as dogmatic by others. This makes it particularly challenging to demarcate the boundary between legitimate and illegitimate non-epistemic value-influences. There doesn’t seem to be a clear way to distinguish between scenarios where patient-specific adjustments are unavoidable to manage inductive risk, and scenarios where they outweigh epistemic considerations in an illegitimate way.

In my view, as long as we don’t have strategies to address the new demarcation problem in this context, it seems preferable to set priors based on the highest realistic baseline estimates of covert awareness alone, without additional value-based adjustments. Of course, using the *highest* baseline estimates as an anchoring point does reflect an implicit non-epistemic value judgment (that missing responsiveness is worse than misattributing responsiveness). The suggestion to diverge from this anchoring point to reflect individual values is motivated by the worry that a “one-size-fits-all” approach might be misaligned with patients’ own values (Birch, 2023, p. 4).

Yet, a standardized approach for setting priors would ensure that the results of different studies are comparable and that the inevitable value judgments involved in setting priors can be scrutinized by other scientists based on common standards. If priors were adjusted based on patients’ individual values, the probability of detecting responsiveness would vary from patient to patient. This would make it more difficult to use baseline rates of covert awareness as an anchoring point for Bayesian priors in the future, as these estimates would not only be based on epistemic considerations, but also reflect individual non-epistemic value judgments.

The importance of methodological conventions and a shared set of values among scientists has been highlighted as necessary to facilitate epistemic trust in science. For example, Torsten Wilholt argues that scientists should make non-epistemic value judgments based on “shared ideas about the value of true results and the dangers inherent in errors” (2013, p. 248). While a standardized approach for setting priors would not leave room for adjustments based on a patient’s individual values, making methodological choices like these based on a shared set of values (such as an agreement that missing responsiveness would be worse than misattributing responsiveness) would facilitate scientific coordination, thus decreasing the risk of eroding trust in science. This proposal would even be compatible with proposals that entail that the relevant set of values should be aligned with democratic or stakeholder values (Schroeder, 2021).

Hence, even though the worry that a “one-size-fits-all” approach for making methodological choices might be misaligned with the patient’s own values should be taken seriously, there are advantages to standardized strategies as well, as they seem more suited to delineate legitimate from illegitimate value judgments and facilitate epistemic trust and coordination in science.

I will now turn to the issues related to surrogate decision-making in this context. So far, I have discussed the proposal to integrate patient values into the diagnostic process under the idealized assumption that clinicians would have accurate knowledge of patients’ values when setting priors. In practice, however, patient values would need to be communicated to clinicians via the patient’s surrogate decision-makers (often family members), as DoC patients cannot communicate their values themselves. This creates two potential challenges, the first practical, the second ethical. First, surrogate decision-makers might not know what the patient’s values and risk attitudes are, which means that the risk of potential value misalignment would not be resolved. Second, involving surrogates into methodological choices might increase their decisional burden, as their assessment of the patient’s values would impact the probability of detecting evidence of consciousness in that patient.

 I will address the practical issue first. Surrogate decision-making is usually conceptualized as following a hierarchy of standards: the patient’s own wishes, substituted judgments and the patient’s best interests (Berger et al., 2008). Often, the patient’s own wishes are not explicitly known, as the majority of patients do not complete advance directives, and even if they do, it is often ambiguous how to interpret these in specific situations, especially in the context of end-of-life decisions (Wendler, 2017). Hence, decision-making often relies on the substituted judgment standard, where surrogates need to deliberate how the patient would have decided in that situation, based on knowing the patient’ general values and preferences, as well as receiving relevant medical information and guidance from the clinician.

However, surrogates are often unable to predict what decisions patients would have made (Berger et al., 2008; Wendler, 2017). A systematic review has shown that surrogates predict patients’ treatment preferences wrong in about a third of cases, which is not improved by prior discussion of the patient’s preferences and values (Shalowitz et al., 2006). This has a variety of reasons. Especially in end-of-life decisions, surrogates are not merely “channels” for applying the patient’s values to clinical dilemmas, but might be scared and confused by the possibility of losing a loved one and reluctant to take on responsibility for these decisions (Wendler, 2017, p. 30). Additionally, surrogates might find it difficult to differentiate their own values and wishes from the ones of the patient (Berger et al., 2008).

Predicting patient values in the DoC context seems particularly challenging as the diagnostic uncertainty is so high. Surrogates would not only need to assess patients’ treatment preferences in general, but also their attitudes towards the risk of false positive/negative diagnoses. While it is plausible that surrogates might have intuitions about patient values in this context, there seems to be a relevant risk of adjusting Bayesian priors in a way that does not accurately reflect the patient’s preferences. Hence, it seems questionable whether relying on surrogate decision-makers would alleviate the problem that methodological choices might reflect value judgments that are misaligned with the patient’s own preferences.

My second worry about incorporating patient values into methodological choices concerns the burden of responsibility imposed on surrogate decision-makers. As mentioned, surrogates can be profoundly impacted by having to make decisions on behalf of another person. This decisional burden can lead to negative psychological consequences, such as depression, anxiety and post-traumatic stress disorder (Dionne-Odom & White, 2021; Hickman & Pinto, 2014; Wendler, 2017). Surrogates frequently report feeling guilt and regret after making decisions, irrespective of the outcome, because of feeling uncertain whether they have made the right decision (Su et al., 2020).

Given the diagnostic uncertainty in the DoC context, the burden associated with surrogate decision-making seems particularly high. For example, determining whether continuing or withdrawing life-sustaining treatment would be most closely aligned with the patient’s preferences seems more difficult when it is uncertain to what extent the patient is capable of conscious experience. In my view, the burden of that decision should be shared between clinicians and surrogate decision-makers. Shared-decision making is often conceptualized as a process in which clinicians and patients (or their surrogates) arrive at decisions together, providing a middle way between paternalistic and autonomy-based approaches (Childress & Childress, 2020; Kon, 2010). After discussing the relevant medical evidence, surrogates can deliberate (together with clinicians) what decision would be most closely aligned with the patient’s preferences. Hence, surrogates need to rely on healthcare professionals to present them with the information they need to make decisions in line with the patient’s values.

There is a worry that integrating patient values into the diagnostic process might increase the burden on surrogate decision-makers, as the distinction between gathering evidence and making decisions based on that evidence becomes blurred. Recall that, when involving surrogate decision-makers into methodological choices by letting them convey information about the patient’s values and risk attitudes, the probability of detecting evidence of consciousness in that patient would shift by a small degree depending on how Bayesian priors are set. For example, if a surrogate assumes that a patient would have been averse to a false-positive result, this would entail that Bayesian priors should be set in a more conservative way, thus decreasing the probability of detecting evidence of responsiveness slightly. This might lead to rating the probability of responsiveness in an EEG-based paradigm as unlikely (whereas with more permissive priors, the probability might have been rated as realistic). This result might lead clinicians to regard the overall evidence for the patient retaining consciousness as weak. If the patient’s surrogate assumes that the patient would not have wanted to receive life-sustaining treatment in a state of unconsciousness, the result from this task-based EEG paradigm could influence medical decision-making significantly, potentially shifting the decision towards discontinuing life-sustaining treatment.

In this scenario, there seems to be an even heavier burden on the surrogate than usual: not only do they need to make an extremely challenging decision under uncertain evidence, but it would be their own assessment of the patient’s values and risk attitudes that affects how the evidence of consciousness is interpreted. Hence, contrary to the usual scenario where surrogate decision-makers deliberate, based on evidence they receive from clinicians, which decision would be most closely aligned with the patient’s values, in this example, the surrogate-decision maker would be involved, to a small degree, in influencing the very evidence this decision relies on. It seems conceivable that this could complicate the emotional burden for surrogate decision-makers.

Of course, this is a highly simplified example. In reality, the decision-making process is a lot more complex than that, and does not only depend on the patient’s presumed state of consciousness, but also many other factors, such as etiology, prognosis, comorbidities, etc. Additionally, it is still the clinician’s responsibility to determine the overall diagnosis from the available evidence. But I do think that this hypothetical case shows that the burden on surrogate decision-makers is an important factor to consider when analyzing how realistic it would be to incorporate patient values into methodological decisions. Reducing the burden on relatives is among the most important goals of patients who engage in advance care planning, and some patients prefer that clinicians make these decisions for them (Daveson et al., 2013; Rid et al., 2015). Hence, not all patients would want priors to be adjusted based on their own values.

In my view, the potential additional burden on surrogate decision-makers, as well as the worry that the problem of value misalignment might not be alleviated due to low predictive accuracy, provide further reasons not to adjust priors based on patient values. Rather, I think clinicians should follow a standardized approach for setting priors based on cautious baseline estimates of covert awareness and other clinical information specific to the patient. While this entails that methodological choices at the level of gathering evidence involve non-epistemic value judgments that might be misaligned with the patient’s own values, it offers clearer ways to distinguish between the process of gathering evidence (which is the clinician’s responsibility) and the process of making decisions based on that evidence (which is the shared responsibility of the clinician and the patient’s surrogate decision-maker).

To summarize, I have argued that patient values should not be directly incorporated into methodological choices at the level of individual diagnostic techniques, as there are no clear ways to delineate legitimate from illegitimate non-epistemic value influences and it complicates issues associated with surrogate decision-making. However, this does not entail that patient values should play no role when managing inductive risk. In what follows, I sketch a framework for making decisions under uncertainty that takes patient values into account as much as possible, shifting the focus from individual diagnostic techniques to the assessment and communication of diagnostic uncertainty overall.

1. **Incorporating patient values when making decisions under uncertainty**

Birch’s proposal focuses on managing inductive risk at the level of individual diagnostic techniques, specifically in the context of task-based EEG paradigms that have high false discovery rates. He correctly highlights that patients and their families are the ones who are most deeply affected by the consequences of false positive and false negative results, which is one of his leading motivations to argue that inductive risk management should incorporate the patient’s own values already at this level.

In my view, the risk that patients are harmed by false positive/negative results is particularly high if one makes direct inferences from categorical assessments of consciousness in individual diagnostic techniques (e.g. yes/no verdicts about responsiveness in EEG data) to categorical verdicts about consciousness in general. One of the main strengths of Birch’s proposal is the suggestion to move towards assessing and communicating degrees of evidence instead of providing categorical verdicts. I think that this could be a useful strategy at the level of the patient’s overall diagnosis as well, reducing the risk of translating false results at the level of individual techniques directly into false positive/negative diagnoses.

Given the deep uncertainty about the necessary and sufficient conditions for consciousness, the combined evidence from individual behavioral and neural tests will not support simple yes/no verdicts regarding a patient’s diagnosis. Rather, just as with individual techniques, the overall constellation of evidence will support the hypothesis that a patient is conscious to a varying degree. Hence, instead of making categorical distinctions between conscious and unconscious states (which would lead to harmful consequences in the case of false positive/negative results), clinicians could assess the overall strength of the evidence: how many diagnostic tests indicate evidence of consciousness, and how sensitive and specific are these tests? This assessment could be translated into degrees of diagnostic (un-)certainty: The stronger the evidence of consciousness, the higher the certainty about the patient’s diagnosis. For example, if all neural techniques (active, passive and resting state paradigms) show activity patterns indicative of consciousness, the evidence for consciousness might be regarded as stronger than if only techniques with high false discovery rates, or no neural techniques at all, show positive results. This way, a false positive result in an EEG-based paradigm would not directly lead to a false positive diagnosis of consciousness. Rather, it would lead to assessing the overall evidence of consciousness as stronger than might have otherwise been the case, but the deep uncertainty surrounding the overall diagnosis would be acknowledged and communicated either way.

Quantifying the strength of evidence and degree of diagnostic uncertainty in this way would give clinicians and surrogates the opportunity to consider the patient’s own attitudes towards risk and uncertainty when making clinical decisions. Instead of already incorporating patient values at the stage of *gathering* evidence, the focus would be on deciding how to *respond* to that evidence, making room for individual adjustments based on the patient’s own values.

When deliberating how to make decisions based on uncertain evidence, the precautionary principle can provide helpful guidance. Stephen John has suggested that the precautionary principle, an often-invoked norm in the context of making policy decisions under uncertainty, can be most adequately understood as an epistemic acceptance rule for policy-makers (2019). The bar of evidence that policy-makers should require to accept a hypothesis as a premise in practical reasoning should vary with the expected costs of error: if the priority is to avoid false positives (i.e. accepting claims that are false), more evidence should be required to accept a hypothesis than if the priority is to avoid false negatives (i.e. failing to accept claims that are true).

Decision-making in the DoC context could be interpreted in a similar way: the amount of evidence one regards as necessary to make particular clinical decisions will depend on one’s precautionary attitudes. In my view, clinicians and surrogates should consider the patient’s own attitudes as far as possible in this context. This way, potential value misalignments at the stage of gathering evidence could be mitigated when deciding how to respond to that evidence.

It is important to highlight that patient attitudes about the risk of false positive and negative errors will depend on the specific decision-making context. For example, in the context of pain management, the risk of false negative errors seems to outweigh the risk of false positive errors significantly. This is why several clinical guidelines highlight the need for precautionary pain treatment and symptom control in *all* patients with disorders of consciousness (including patients in which there are no signs of covert consciousness) (Fins & Bernat, 2018). No patient should be regarded as definitively incapable of experience in this context. Yet, in other contexts, such as decisions about life-sustaining treatment, patient attitudes about the risks associated with false positive/negative errors might be more variable. Hence, the amount of evidence one regards as necessary to make these decisions might vary. For example, if conversations with the patient’s family reveal that avoiding false negatives was a high priority for the patient, weaker evidence of consciousness would be required to sustain life-sustaining treatment than if the priority was to avoid false positives.

While the practical problem that surrogates might not predict patient values accurately remains, the risk of error would be comparable to the risk of making decisions that are misaligned with the patient’s preferences in general. However, surrogates’ assessment of patient values would not affect the probability of detecting evidence of consciousness directly. In my view, this modification would respect patient values in a way that avoids increasing the burden on surrogate decision-makers.

One might object that it is implausible to worry that incorporating surrogates’ assessment of patient values into the diagnostic process increases their decisional burden. Having to make these decisions is burdensome either way, so is it not conceivable that giving surrogates more autonomy alleviates their burden?[[4]](#footnote-4) Presumably, this depends on how transparent clinicians are about the extent to which they rely on surrogates’ assessments in the diagnostic process. If clinicians set priors based on general conversations with surrogates about the patient’s values, without informing surrogates about how their assessment affects the posterior probability of detecting responsiveness, this would not affect the burden imposed on surrogates. Rather, the decision-making process would still rely on the clinician presenting diagnostic information to the surrogate and then considering together how to proceed in line with the patient’s values.

Yet, this lack of transparency seems problematic. In most medical settings, clinicians have the sole responsibility for gathering diagnostic information and do not consult patients on methodological choices in the diagnostic process, whether directly or indirectly. In my view, a departure from this standard, even if justified, entails the necessity to transparently communicate to surrogates how their assessment of patient values affects diagnostic protocols. Yet, in this case, surrogates would be aware that the diagnostic process is influenced, to a small degree, by their assessment.

Of course, we should not underestimate people’s ability to process complex information. Given that non-epistemic value judgments are inevitable and priors need to be set in some way, some surrogates might prefer having more control over these decisions. However, it is equally conceivable that departing from the scenario where surrogates rely on the clinician to present them with the evidence they need makes an already difficult decision more burdensome, as surrogates might feel like the responsibility for the outcome of the diagnostic process is partially transferred to them.

Not integrating the values of individual patients into the diagnostic process does not entail an endorsement of paternalistic approaches where clinicians make all decisions for the patient. Highlighting the importance of patient values when considering how to respond to uncertain evidence is precisely motivated by the need to respect patient autonomy. Yet, relying on clinicians to bear the responsibility for the diagnostic process and make non-epistemic value judgments in a standardized and transparent way does not need to endanger patient autonomy. As mentioned, the non-epistemic value judgments affecting these choices can be scrutinized by the scientific community, and there are various proposals for how to ensure that the values in question are aligned with those of relevant stakeholders (Schroeder, 2021).

In cases where patients’ precautionary attitudes are not known, or in cases where surrogates want clinicians to make these decisions for them, clinicians could take the general recommendation of the European guideline for diagnosing DoC as a starting point, stating that patients should be diagnosed “with the highest level of consciousness” revealed by any of the available techniques (i.e. clinical examination, EEG and neuroimaging paradigms) (Kondziella et al., 2020, p. 751). This recommendation entails a precautionary attitude that prioritizes avoiding false negatives, as positive results from individual diagnostic techniques (including false positives) would be directly translated into the patient’s overall diagnosis. Hence, this proposal would need to be supplemented by considering the wider diagnostic picture. Additionally, the possibility of experience in patients with negative results in all neuroimaging paradigms would need to be incorporated as well, especially in the context of pain management.

To sum up, I have sketched a proposal for managing inductive risk in the DoC context that focuses on making context-specific decisions under uncertainty based on the patient’s own precautionary attitudes, while favoring a standardized approach for managing non-epistemic values at the level of individual diagnostic techniques. Of course, this proposal needs to be specified in more detail to analyze how exactly the overall strength of evidence and degree of diagnostic uncertainty should be assessed and communicated. Here, I just want to highlight that considering patient values when managing inductive risk seems possible without influencing the process of gathering evidence directly. At the stage of making decisions under uncertainty, patient values would not affect the probability of detecting evidence of consciousness, but determine how one responds to that evidence.

1. **Conclusion**

Diagnosing DoC patients involves inductive risk, both at the level of individual diagnostic techniques and the level of the overall diagnosis. When managing that risk, it is important to ensure that the patient’s own values are taken into account. However, the proposal to incorporate patient values into methodological choices at the stage of gathering evidence comes with its own challenges. First, it raises concerns about how to delineate legitimate from illegitimate value influences. Second, it faces difficulties associated with surrogate decision-making. For these reasons, I argued that patient values should not be directly incorporated into methodological choices when gathering evidence of consciousness.

However, patient values play an important role when making decisions under uncertainty. Assessing and communicating the strength of the available evidence and degree of diagnostic uncertainty would enable clinicians and surrogates to respond to uncertain evidence in a way that takes the patient’s precautionary attitudes into account. This strategy could mitigate inductive risk in the DoC context in a way that keeps the advantages of Birch’s proposal while avoiding some of its limitations.

**Acknowledgments**

I would like to thank Marta Halina, Jonathan Birch and two anonymous reviewers for helpful comments on this article. I am also grateful to the Values in Science Group Cambridge, in particular Stephen John, Ahmad Elabbar, Henrik Røed Sherling, Benjamin Chin-Yee, Alessandra Basso, Phillip Hintikka Kieval and Cristian Larroulet Philippi, for valuable feedback and discussions that helped me shape these ideas.

**References**

Berger, J. T., Derenzo, E. G., & Schwartz, J. (2008). Surrogate Decision Making: Reconciling Ethical Theory and Clinical Practice. *Annals of Internal Medicine*, *149*(1), 48–53. https://doi.org/10.7326/0003-4819-149-1-200807010-00010

Biddle, J. B., & Kukla, R. (2017). The geography of epistemic risk. In K. C. Elliott & T. Richards (Eds.), *Exploring Inductive Risk: Case Studies of Values in Science* (pp. 215–238). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190467715.001.0001

Birch, J. (2023). Medical AI, inductive risk and the communication of uncertainty: the case of disorders of consciousness. *Journal of Medical Ethics*, jme-2023-109424. https://doi.org/10.1136/jme-2023-109424

Birch, J., Creel, K. A., Jha, A. K., & Plutynski, A. (2022). Clinical decisions using AI must consider patient values. *Nature Medicine*, *28*(2), 229–232. https://doi.org/10.1038/s41591-021-01624-y

Bodien, Y. G., Allanson, J., Cardone, P., Bonhomme, A., Carmona, J., Chatelle, C., Chennu, S., Conte, M., Dehaene, S., Finoia, P., Heinonen, G., Hersh, J. E., Kamau, E., Lawrence, P. K., Lupson, V. C., Meydan, A., Rohaut, B., Sanders, W. R., Sitt, J. D., … Schiff, N. D. (2024). Cognitive Motor Dissociation in Disorders of Consciousness. *New England Journal of Medicine*, *391*(7), 598–608. https://doi.org/10.1056/nejmoa2400645

Childress, J. F., & Childress, M. D. (2020). What Does the Evolution From Informed Consent to Shared Decision Making Teach Us About Authority in Health Care? *AMA Journal of Ethics*, *22*(5), E423-429. https://doi.org/10.1001/amajethics.2020.423

Claassen, J., Doyle, K., Matory, A., Caroline, C., M, B. K., Angela, V., U, O. J., Jean-Rémi, K., Soojin, P., Sachin, A., David, R., Murad, M., Andrey, E., Sander, C., & Rohaut, B. (2019). Detection of Brain Activation in Unresponsive Patients with Acute Brain Injury. *New England Journal of Medicine*, *380*(26), 2497–2505. https://doi.org/10.1056/NEJMoa1812757

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., Laureys, S., & Owen, A. M. (2013). Reanalysis of “Bedside detection of awareness in the vegetative state: a cohort study” - Authors’ reply. *The Lancet*, *381*(9863), 291–292. https://doi.org/10.1016/S0140-6736(13)60126-9

Daveson, B. A., Bausewein, C., Murtagh, F. E. M., Calanzani, N., Higginson, I. J., Harding, R., Cohen, J., Simon, S. T., Deliens, L., Bechinger-English, D., Hall, S., Koffman, J., Ferreira, P. L., Toscani, F., Gysels, M., Ceulemans, L., Haugen, D. F., & Gomes, B. (2013). 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. *Palliative Medicine*, *27*(5), 418–427. https://doi.org/10.1177/0269216312471883

Dehaene, S., & Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Dionne-Odom, J. N., & White, D. B. (2021). Reconceptualizing how to support surrogates making medical decisions for critically ill patients. *Journal of the American Medical Association*, *325*(21), 2147–2148. https://doi.org/10.1001/jama.2021.6445

Douglas, H. (2000). Inductive Risk and Values in Science. *Philosophy of Science*, *67*(4), 559–579. https://doi.org/10.1086/392855

Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press. https://doi.org/10.2307/j.ctt6wrc78

Edlow, B. L., Claassen, J., Schiff, N. D., & Greer, D. M. (2021). Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies. *Nature Reviews. Neurology*, *17*(3), 135–156. https://doi.org/10.1038/S41582-020-00428-X

Edlow, B. L., & Fins, J. J. (2018). Assessment of covert consciousness in the intensive care unit: Clinical and ethical considerations. *Journal of Head Trauma Rehabilitation*, *33*(6), 424–434. https://doi.org/10.1097/HTR.0000000000000448

Elliott, K. C. (2022). *Values in Science*. Cambridge University Press. https://doi.org/10.1017/9781009052597

Fins, J. J., & Bernat, J. L. (2018). Ethical, palliative, and policy considerations in disorders of consciousness. *Neurology*, *91*(10), 471. https://doi.org/10.1212/WNL.0000000000005927

Giacino, J. T., Fins, J. J., Laureys, S., & Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: The state of the science. *Nature Reviews Neurology*, *10*(2), 99–114. https://doi.org/10.1038/nrneurol.2013.279

Giacino, J. T., Katz, D. I., Schiff, N. D., Whyte, J., Ashman, E. J., Ashwal, S., Barbano, R., Hammond, F. M., Laureys, S., Ling, G. S. F., Nakase-Richardson, R., Seel, R. T., Yablon, S., Getchius, T. S. D., Gronseth, G. S., & Armstrong, M. J. (2018). Practice guideline update recommendations summary: Disorders of consciousness: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and . *Neurology*, *91*(10), 450–460. https://doi.org/10.1212/WNL.0000000000005926

Harrison, & Connolly, J. F. (2013). Finding a way in: A review and practical evaluation of fMRI and EEG for detection and assessment in disorders of consciousness. *Neuroscience and Biobehavioral Reviews*, *37*(8), 1403–1419. https://doi.org/10.1016/j.neubiorev.2013.05.004

Hickman, R. L., & Pinto, M. D. (2014). Advance directives lessen the decisional burden of surrogate decision-making for the chronically critically ill. *Journal of Clinical Nursing*, *23*(5–6), 756–765. https://doi.org/10.1111/jocn.12427

Hillman, E. M. C. (2014). Coupling Mechanism and Significance of the BOLD Signal: A Status Report. *Annual Review of Neuroscience*, *37*, 161. https://doi.org/10.1146/ANNUREV-NEURO-071013-014111

Holman, B., & Wilholt, T. (2022). The new demarcation problem. *Studies in History and Philosophy of Science*, *91*, 211–220. https://doi.org/https://doi.org/10.1016/j.shpsa.2021.11.011

John, S. (2019). The Politics of Certainty: The Precautionary Principle, Inductive Risk and Procedural Fairness. *Ethics, Policy & Environment*, *22*(1), 21–33. https://doi.org/10.1080/21550085.2019.1581418

Johnson, L. S. M. (2021). *The Ethics of Uncertainty: Entangled Ethical and Epistemic Risks in Disorders of Consciousness*. Oxford University Press.

Kon, A. A. (2010). The Shared Decision-Making Continuum. *Journal of the American Medical Association*, *304*(8), 903–904. https://doi.org/10.1001/jama.2010.1208

Kondziella, D., Bender, A., Diserens, K., van Erp, W., Estraneo, A., Formisano, R., Laureys, S., Naccache, L., Ozturk, S., Rohaut, B., Sitt, J. D., Stender, J., Tiainen, M., Rossetti, A. O., Gosseries, O., & Chatelle, C. (2020). European Academy of Neurology guideline on the diagnosis of coma and other disorders of consciousness. *European Journal of Neurology*, *27*(5), 741–756. https://doi.org/10.1111/ene.14151

Laureys, S., Celesia, G. G., Cohadon, F., Lavrijsen, J., León-Carrión, J., Sannita, W. G., Sazbon, L., Schmutzhard, E., von Wild, K. R., Zeman, A., & Dolce, G. (2010). Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Medicine*, *8*, 68. https://doi.org/10.1186/1741-7015-8-68

Merker, B. (2007). Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *The Behavioral and Brain Sciences*, *30*(1), 63–134. https://doi.org/10.1017/S0140525X07000891

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, *83*(October), 435–450.

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., Jolles, D., & Pickard, J. D. (2007). Response to Comments on “Detecting Awareness in the Vegetative State.” *Science*, *315*(5816), 1221 LP – 1221. https://doi.org/10.1126/science.1135583

Peterson, A. (2016). Consilience, clinical validation, and global disorders of consciousness. *Neuroscience of Consciousness*, *2016*(1), niw011–niw011. https://doi.org/10.1093/nc/niw011

Peterson, A., Cruse, D., Naci, L., Weijer, C., & Owen, A. M. (2015). Risk, diagnostic error, and the clinical science of consciousness. *NeuroImage. Clinical*, *7*, 588–597. https://doi.org/10.1016/j.nicl.2015.02.008

Posner, J. B. (2019). *Plum and Posner’s diagnosis and treatment of stupor and coma* (C. B. Saper, N. D. Schiff, J. Claassen, & F. Plum, Eds.; Fifth edit). Oxford University Press.

Rid, A., Wesley, R., Pavlick, M., Maynard, S., Roth, K., & Wendler, D. (2015). Patients’ priorities for treatment decision making during periods of incapacity: Quantitative survey. *Palliative and Supportive Care*, *13*(5), 1165–1183. https://doi.org/10.1017/S1478951514001096

Rudner, R. (1953). The Scientist Qua Scientist Makes Value Judgments. *Philosophy of Science*, *20*(1), 1–6. http://www.jstor.org/stable/185617

Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., Moonen, G., & Laureys, S. (2009). Diagnostic accuracy of the vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology*, *9*(1), 35. https://doi.org/10.1186/1471-2377-9-35

Schroeder, S. A. (2021). Democratic Values: A Better Foundation for Public Trust in Science. *British Journal for the Philosophy of Science*, *72*(2), 545–562. https://doi.org/10.1093/bjps/axz023

Schroeder, S. A. (2024). The Influence of Values on Medical Research. In A. Broadbent (Ed.), *The Oxford Handbook of Philosophy of Medicine (forthcoming)*. https://philpapers.org/archive/SCHTIO-145.pdf

Shalowitz, D. I., Garrett-Mayer, E., & Wendler, D. (2006). The Accuracy of Surrogate Decision Makers: A Systematic Review. *Arch Intern Med*, *166*, 493–497.

Shea, N., & Bayne, T. (2010). The Vegetative State and the Science of Consciousness. *The British Journal for the Philosophy of Science*, *61*(3), 459–484. https://doi.org/10.1093/bjps/axp046

Su, Y., Yuki, M., & Hirayama, K. (2020). The experiences and perspectives of family surrogate decision-makers: A systematic review of qualitative studies. *Patient Education and Counseling*, *103*(6), 1070–1081. https://doi.org/10.1016/j.pec.2019.12.011

Watson, C., Kirkcaldie, M., & Paxinos, G. (2010). *The Brain: An Introduction to Functional Neuroanatomy*. Elsevier Science & Technology.

Wendler, D. (2017). The Theory and Practice of Surrogate Decision-Making. *Hastings Center Report*, *47*(1), 29–31. https://doi.org/10.1002/hast.671

Wilholt, T. (2013). Epistemic Trust in Science. *The British Journal for the Philosophy of Science*, *64*(2), 233–253. https://doi.org/10.1093/bjps/axs007

1. In what follows, I use the terms consciousness and awareness interchangeably to refer to subjective experience, for consistency with the philosophical literature. I use the term arousal to refer to wakefulness. [↑](#footnote-ref-1)
2. Unresponsive wakefulness syndrome (UWS) is also referred to as the vegetative state. As this term has pejorative connotations and the question whether patients in this state are actually unaware is contested, the merely descriptive term UWS is preferable (Laureys et al., 2010). [↑](#footnote-ref-2)
3. Some of the risks involved in misdiagnosing and treating DoC patients might better be characterized as epistemic risks. This broader category refers to the general risk of mistakes and false conclusions during the process of generating empirical knowledge under uncertainty (Biddle & Kukla, 2017). For example, adopting false background assumptions about which neural networks are associated with conscious experience might be most appropriately classified as an epistemic risk, not as inductive risk in the narrower sense. While my discussion of managing uncertainty in the DoC context is connected to these broader risks, the main focus is on how to manage the inductive risk of false positive and negative inferences, given the evidence from diagnostic tests. Thanks to an anonymous reviewer for encouraging me to clarify that point. [↑](#footnote-ref-3)
4. I am grateful to an anonymous reviewer for raising that objection. [↑](#footnote-ref-4)