

Reevaluating the Dead Donor Rule

MIKE COLLINS*

Mount Sinai School of Medicine, New York, NY, USA

*Address correspondence to: Mike Collins, PhD, The Bioethics Program,
Mount Sinai School of Medicine, One Gustave Levy Place, Box 1108, New York,
NY 10029, USA. E-mail:mike9943@gmail.com

The dead donor rule justifies current practice in organ procurement for transplantation and states that organ donors must be dead prior to donation. The majority of organ donors are diagnosed as having suffered brain death and hence are declared dead by neurological criteria. However, a significant amount of unrest in both the philosophical and the medical literature has surfaced since this practice began forty years ago. I argue that, first, declaring death by neurological criteria is both unreliable and unjustified but further, the ethical principles which themselves justify the dead donor rule are better served by abandoning that rule and instead allowing individuals who have suffered severe and irreversible brain damage to become organ donors, even though they are not yet dead and even though the removal of their organs would be the proximal cause of death.

Keywords: *brain death, dead donor rule, death, organ donation, organ transplantation*

I. ORGAN TRANSPLANTATION AND BRAIN DEATH

Human organs are sensitive to damage from ischemia. Because of this, the majority of solid organs are removed from donors who have been declared dead by neurological criteria yet are maintained on somatic life support equipment and maintain a spontaneous heartbeat right up until the moment of removal. In this way, the organs to be donated are perfused with oxygenated blood for as long as possible.

No one questions the societal benefit that has been brought about by the advent of organ transplantation technology. What has been questioned for many years is the coherence and acceptability of determining death by neurological criteria. Although compelling arguments have been made both philosophically

and empirically regarding the conceptual confusion of equating brain death with death and the unacceptable level of uncertainty involved in making the diagnosis of brain death, both the clinical practice and the unrest over the legitimacy of this practice remain. The debate at this point needs to be refocused.

There are different aspects to the concept of death. In particular, there is a biological notion of death, which seemingly contrasts with a socio-legal notion, in which what is at issue is not a strict metaphysical/biological question of what death is or when it has occurred but is a question involving the moral and legal standing of individuals. What is particularly relevant to the present issue is the moral difference between permissible actions that may be taken upon a dead human body versus those permitted on a living human body. The dead donor rule is related to this socio-legal notion of death: it states that it is permissible to remove vital organs only from a dead body but never from a living body.

In this paper, I argue the following. First, brain death is not death. This is not a new claim, however, the force of the arguments in its favor, and more importantly, the confusion in the arguments that oppose it are not sufficiently appreciated. For this reason, I summarize and review the arguments on both sides of this debate. Once we have clarified the dialectic, we see that the arguments that identify brain death with death are mostly non sequiturs. Second, I argue that there are fundamental ethical principles that justify our current guidelines for organ donation (which include the dead donor rule). By making those principles explicit, we can see that there is a better way to implement them into practical guidelines. The new proposed guidelines take into account the moral and socio-legal aspects of the concept of death by noting when it is morally permissible to remove vital organs yet avoid the confusion and inconsistencies inherent in identifying death with brain death. I conclude the paper by suggesting that the true nature of the brain death debate is also underappreciated. This is not a metaphysical or scientific debate about the nature of death. It is a normative debate about when, if ever, it is morally acceptable to remove vital organs from a terminally ill yet living human body. The brain death debate only serves to obfuscate that.

II. WHY DECLARING DEATH BY NEUROLOGICAL CRITERIA IS BOTH UNRELIABLE AND UNJUSTIFIED¹

Several arguments have been advanced over the years against using neurological criteria to determine when an individual is dead. The most trenchant and pressing of these arguments can be organized around two principles: the unreliability of the diagnosis of brain death and the lack of coherence of the physiological criterion for brain death with the concept of the biological death of an organism.

The debate typically centers on acknowledging the distinction between the conceptual definition of death, the physiological criteria that must be

satisfied in order for death to have occurred, and the diagnostic tests involved in determining whether or not those criteria have been satisfied.² Conceptually, death is the cessation of the integrative unity of the various metabolic processes involved in maintaining homeostasis and resisting entropy. The physiological criteria for that to have occurred is the irreversible cessation of all functions of the entire brain (including the brain stem). The diagnostic tests involved in discovering whether or not irreversible cessation of all brain function has occurred involve three essential elements: unresponsiveness, absence of cranial nerve function as determined by clinical bedside tests, and apnea as determined by elevated levels of arterial carbon dioxide pressure in the absence of spontaneous respiratory effort. In addition, the cause of coma must be established (this satisfies the “irreversibility” component of the definition), and confounding factors such as hypothermia, metabolic disturbances, neuromuscular blocking agents, and shock must be ruled out. Confirmatory tests such as electroencephalograms (EEGs) or angiograms are not necessary to make the diagnosis except under special circumstances (Beecher, 1968; Bernat, Culver, and Gert, 1981; President’s Commission, 1981; American Academy of Neurology, Quality Standards Committee, 1995; Wijdicks, 1995).

The first major problem with the standard justification for current practice is that the diagnosis of brain death is unreliable: there are significantly many cases of false positives. The evidence for this can be conveniently placed into two categories, the first of which involves noninvasive electrophysiological testing such as EEG and evoked potential studies, whereas the second is clinically apparent evidence of neurological functioning in the so-called “brain dead” patients.

Early in the brain death literature, it was discovered that some individuals who meet the diagnostic requirements for brain death retain both cortical and subcortical organized electrical activity. For example, one study has shown that 11 out of a pool of 56 consecutive patients diagnosed as brain dead maintained spontaneous cortical electrical activity, showing low-voltage theta or beta activity (Grigg et al., 1987), and another study described a woman meeting the diagnostic criteria for brain death yet maintaining spontaneous electrical activity resembling sleep spindles (Rodin et al., 1985; see also Deliyannakis, Ioannou, and Davaroukas, 1975). Further evidence was brought to light regarding the maintenance of sensory processing pathways through the use of evoked potential studies. These studies have demonstrated the functional integrity of afferent pathways in the form of normal (or near-normal) auditory (Barelli et al., 1990), visual (Ferber et al., 1986), and somatosensory (Anziska and Cracco, 1980) evoked potentials.

Defenders of the status quo reply that the EEG studies do not show evidence of function but merely activity (Bernat, 1998, 2006). However, first, the brain is far too complex of an organ for us to claim to know which activity should count as “functional” and which should not. By way of contrast,

consider the pulseless electrical activity that is sometimes seen with cardiac arrest. There is electrical activity, but that activity does not cause the myocardium to contract and thereby circulate blood. In that case, we can be justified in claiming that there is electrical activity that does not serve its function because, clearly, the function of the heart is to pump blood. The brain, by contrast, is far more complex and does not admit of such transparent attributions of function to any particular pattern of electrical activity. Second, the spatial resolution of the EEG is quite poor, thus rendering attributions of function even more difficult. Third (and more importantly), the evoked potential studies do indeed demonstrate the preservation of function.

Maintaining receptivity to the environment in such a way as to transduce ambient energy into electrochemical signals is certainly one of the many functions of the central and peripheral nervous systems. The demonstration of normal or close-to-normal evoked potentials involving transduction of various forms of ambient energy into organized waves of depolarization and the release of neurotransmitter necessary to propagate those signals clearly demonstrates a connection to and a form of (electrical but not behavioral) responsiveness to the environment. Finally, the debate over whether these studies show evidence of neural function (as opposed to activity) is obviated by the second category of evidence demonstrating the unreliability of the diagnosis of brain death.

Diabetes insipidus is clinically apparent through polyuria, and its absence is conspicuous by the lack of polyuria. Further differentiation between central and nephrogenic diabetes insipidus is possible through serum markers of the antidiuretic hormone arginine vasopressin, which is secreted by the pituitary under the control of the hypothalamus. Surely the regulation of vital physiologic parameters such as the maintenance of free water homeostasis is a function of the brain. As it turns out, many patients who have been diagnosed as brain dead either do not develop diabetes insipidus or develop nephrogenic and not central diabetes insipidus (Grenvik, 1978; Schrader et al., 1980; Outwater and Rockoff, 1984; Fiser et al., 1987; Fackler, Troncoso, and Gioia, 1988; Hohnegger et al., 1990). This provides essentially incontrovertible evidence that individuals might meet the diagnostic requirements for brain death in terms of unresponsiveness, apnea, and loss of cranial nerve reflexes, yet those patients are clearly not brain dead: It is not the case that all functions of the brain have ceased irreversibly since the hypothalamus still maintains vital physiologic parameters through the regulation of neuroendocrine secretion.

Further evidence involves a cardiovascular response to tissue damage that occurs as a result of organ retrieval. Organ donors have been observed to develop tachycardia and hypertension in response to organ removal (Wetzel et al., 1985; Pennefather, Dark, and Bullock, 1993; Hill, Munglani, and Sapsford, 1994), suggesting to some the maintenance of neurological function at a supraspinal level (Truog, 1997), although others have suggested that this

might be a purely spinal cord–mediated response (Shewmon, 2001). What has been clearly established is that individuals can satisfy the currently acceptable diagnostic standards for brain death yet still maintain spontaneous organized cortical electrical activity, transduce physical stimuli from the environment into electrochemical neurological signals and pass them along the usual pathways suggesting electrical connectivity to the environment, as well as maintain internal homeostasis of vital physiological parameters including free water balance, and (perhaps) further attempt to maintain internal homeostasis through hemodynamic responses to tissue damage. It thus follows that when a physician uses the accepted standards to diagnose a patient as brain dead, this is by no means a reliable indication that *all* functions of that individual's brain have irreversibly ceased.

The next step for the defender of the brain-oriented concept of death is to demarcate the significant functions from the insignificant ones, so that it is acceptable to diagnose an individual as brain dead by simply ignoring certain residual functions (Bernat, 1998, 2006). This step in the dialectic is both ad hoc and a non sequitur. It is ad hoc because there is no reason, prior to this debate, for anyone to claim that sensory transduction and information encoding, the maintenance of free water homeostasis, and the regulation of hemodynamic responses to tissue damage (if in fact this is mediated supraspinally in the cases under consideration) are insignificant functions, whereas pupillary and corneal reflexes (which are tested in order to assess cranial nerve function) are significant (this point was also made in Truog and Robinson, 2003). It is a non sequitur because which functions are significant and which are not is irrelevant to this debate: according to the standard justification for current practice and the Uniform Determination of Death Act, an individual is dead if *all* functions of the brain have ceased, not just the ones that are deemed “significant.” Demarcating the significant from the insignificant functions of the brain leads only to the essentialism of the higher brain concept of death, which is not at issue here.

James Bernat is one of the principal architects of the brain death doctrine and commits the above fallacy. In their seminal paper, Bernat, Culver, and Gert (1981) defined death as “the permanent cessation of functioning of the organism as a whole” (390) and identified “permanent loss of functioning of the entire brain [as] perfectly correlated with the permanent cessation of functioning of the organism as a whole” (391), and hence, as the ideal physiological criterion to satisfy the definition of death. Due to the sorts of considerations mentioned above (especially those involving neurohormonal functioning), Bernat (1998, 17) later modified the definition of death to “the permanent cessation of the *critical* functions of the organism as a whole” (my emphasis). He also modified the criterion for death, from permanent loss of functioning of the entire brain, to “the irreversible cessation of all *clinical* functions of the entire brain” (Bernat, 1998, 18; my emphasis).

In so doing, Bernat has subtly but importantly changed the dialectic. The question for Bernat is no longer whether brain death is death; he has already conceded that it is not. Brain death, recall, is the cessation of all functions of the brain. If some neurological functions may remain (i.e., those that are not “critical”) yet the organism is nonetheless dead, then brain death is sufficient but not necessary for death. For Bernat, something weaker is now required: the cessation of the “critical” functions, whose criterion is the cessation of “clinical” neurological functions, is necessary for death. It is important to recognize that the dialectic has now shifted from brain death to partial brain dysfunction. Because of this, the move from the declaration of neurohormonal functions as noncritical to the claim that brain death is death is fallacious. However, let us briefly engage this new dialectic: If the argument works, then there is no reason we cannot just redefine “brain death” as the permanent cessation of the *clinical* functions of the brain rather than *all* functions of the brain.

How then shall we understand “critical function”? As follows: “I use ‘criticality’ to refer to the extent to which a given function of the organism as a whole is necessary for the maintenance of life, health, and unity of the organism” (Bernat, 1998, 17). This, however, is circular. We cannot define “death” as the loss of critical functions of the organism and then define “critical functions” as the functions necessary for life, i.e., the functions without which an organism is dead). These definitions may both be true, but they are also tautologous and hence trivial. Additionally, Bernat’s choice of clinical functions, which are “important functions of the organism that are readily observable or measurable on bedside neurological examination” (Bernat, 1998, 18), as the criteria for critical functions, is ad hoc. It serves only to rule out those neurological functions that are not observable at the bedside, such as neurohormonal functions (which may require blood tests to differentiate central from nephrogenic diabetes insipidus) or the sensory transduction of ambient energy (which require an evoked potential study). But those are the very functions that have been used to challenge the brain death doctrine. Even assuming that the notion of a critical function could be made noncircular, the epistemic access of neurologists is not a relevant consideration when determining whether a set of functions constitutes physiological criteria for critical functions. Clinical observability is irrelevant. So we should not change the dialectic in the way that Bernat would like to, because the notion of a critical function is essentially undefined, and the use of clinical functions as physiological criteria for the undefined critical functions is arbitrary and ad hoc.

The tests that are standardly used to diagnose brain death, thus, are not reliable. But this does not show that brain death is not death, it only shows that our tests for brain death are not acceptable. I have provided a summary of the findings demonstrating the unreliability of the standard tests, first, in order to argue for the unreliability of those tests but more importantly in

order to clarify the dialectic. Digressions into discussions of function versus activity are misguided (because ad hoc, and there exists obvious preservation of function anyway). Discussions of significant versus insignificant (or critical versus noncritical) functions are irrelevant to this particular debate: If we would like to argue that individuals who have lost most (but not all) neurological functions are in fact dead, that is fine, but that constitutes *a change of subject*, to some kind of a partial-brain or higher brain conception of death. What is at issue is whether, as stated in the Uniform Determination of Death Act, an individual who has lost *all* functions of the brain is dead.³

Even if we had absolutely perfect tests for determining the cessation of all brain function, however, it would not matter. An individual with a completely nonfunctioning brain who is maintained on a ventilator (and has a spontaneous perfusing cardiac rhythm) is no more dead than is an individual with a completely nonfunctioning renal system who is maintained on a dialysis machine.

The justification for choosing the irreversible cessation of all neurological functions as the criterion for death is as follows. It is the loss of integrative unity, or the cessation of the functioning of the organism as a whole, that constitutes the death of an organism. Without the brain serving to integrate and unify the dynamic and metabolic processes of the various organ systems, the body is no longer a unified whole that acts together to maintain homeostasis and resist entropy and disintegration but is a mere collection of parts. Thus, without the brain, integrative unity is lost, and the organism as a whole no longer exists; hence, it is dead (Beecher, 1968; Bernat, Culver, and Gert, 1981; Bernat, 1998).

This line of reasoning has been challenged on both philosophical and empirical grounds. Perhaps most telling is a simple counterexample (or rather, several). Shewmon (2001) has provided a list of numerous functions that can be carried out in the absence of any organizing or modulating influence by the brain. (This list arose from several documented patients that were brain dead for a significant period of time yet continued to carry out the following functions.) These include elimination, detoxification, and recycling of wastes, wound healing, nutrition, cellular respiration, fighting of infection and development of the febrile response, and even dramatic examples such as the gestation of a fetus by a brain dead woman and the sexual maturation and proportional growth of a brain dead child. These are each unifying functions that serve to maintain homeostasis and resist entropy yet can be carried out in the absence of neurological modulation. Thus, pace Bernat, Culver, and Gert (1981), complete loss of functioning of the entire brain does not perfectly correlate with permanent cessation of the functioning of the organism as a whole.

Conceptually, this notion is challenged first by the claim that integration does not imply the need for an integrator or a central organizing unit. This is made obvious by complex systems such as plants and embryos, which

operate as integrated units made of many parts that jointly work together to resist entropy and disintegration but require no central integrator without which unity is lost (this point was made in [Shewmon, 2001](#), 473). Second, a very close analogy, in both theory and practical guidelines for care, exists between individuals who suffer a global brain infarction down to the foramen magnum and those who suffer a transient spinal shock at the cervico-medullary junction ([Shewmon, 1999](#)). In both cases, the body is completely divorced from communication with the brain,⁴ yet in the spinal shock case, we do not hesitate to say that both the person and the body are alive. By parity of reasoning then, it could not be the lack of influence by the brain qua organizing and integrating unifier that is the difference between a living body and a dead body. But this is exactly the justification presented for considering an individual who has lost all brain functions as a dead body. Thus, the integrating unifier rationalization for brain death as the physiological criterion of death is unjustified.

It is important to note that consciousness is a nonfactor in this particular argument. Patients who suffer from spinal shock do in fact maintain consciousness, whereas brain death patients do not. However, the justification for choosing the physiological criterion of total lack of brain function is the cessation of integrative unity not the lack of consciousness. An individual in a deep coma or persistent vegetative state is not dead, even though that person is also not conscious. So the fact that the spinal shock patient is conscious is irrelevant to the point that Shewmon makes with respect to the standard rationale for equating brain death with death; consciousness is not the dividing line between life and death.

In his 1998 article, Bernat makes this very mistake. He argues, first, that there is a distinction between a function and the mechanism that subserves this function (citing [Tomlinson, 1984](#), on this point) and that what matters for being a criterion for death is the function not the mechanism. Then he argues that there is a qualitative difference between the functions of the brain and the functions of other organs; the functions of the brain cannot be mechanically implemented in the way that ventilation and waste secretion can:

The final proof of the validity of the whole-brain criterion centers on the criticality and irreplaceability of the brain. . . Brain function is unlike circulatory, respiratory, or renal function in a qualitative way. Circulation, ventilation, and excretion can be replaced by machines that reproduce the functions of the heart, lung, and kidneys quite well. By contrast, although some of the brain's regulatory functions may be replaced mechanically, the brain's functions of awareness, sentience, and sapience, and its capacities to experience and communicate cannot be reproduced or simulated by any machine. ([Bernat, 1998](#), 19)

If we assume that the functions and not the mechanism that implements those functions are what matter, then it is not the *death* of the brain that matters, it is the loss of its functions. Further, Bernat apparently holds (at least

in the context of the above argument) that regulatory functions of the brain are not critical functions since they can be mechanically implemented. Consciousness and cognition, by contrast, we cannot as of yet implement with machines. Thus, it is not the death of the brain that matters, nor is it the loss of its “noncritical” functions. Rather, it is the loss of the qualitatively different functions of consciousness and cognition that divides life from death.

Bernat’s argument is as follows: The brain serves some functions that cannot be mechanically implemented and some that can. But it is the *irreplaceability* of the brain that justifies the choice of brain death as the criterion for death. Since it is the irreplaceability of those functions that matter, then it is not the regulatory functions, which are replaceable, that justify the brain death doctrine. Rather, the irreplaceable functions of consciousness and cognition are what matter. Thus, it is not qua integrating unifier, it is qua consciousness-producer, that the brain is the critical organ without which the organism is dead.

However, since it is consciousness and cognition *themselves* that matter (as opposed to the neural mechanisms that implement them), then the integrating unity argument is irrelevant, and all that matters are consciousness and cognition. But in that case, any individual who suffered the permanent cessation of consciousness and cognition would lack those very functions and thus would be dead. Patients in persistent vegetative states lack these functions (as far as we know), but they are not dead.

It is important to recognize that I have not argued, as a general interpretation, that Bernat argues that consciousness is the difference between life and death. He does not, and that would be a misinterpretation of his work. However, what we must do, and what Bernat has not done in the argument above, is remain clear on the difference between the following claims: (i) The death of the brain is a criterion for the death of the organism because the brain, *qua integrating unifier*, is the organ without which the organism is dead, and (ii) The death of the brain is a criterion for the death of the organism because the brain, *qua consciousness-producer*, is the organ without which the organism is dead. Bernat (1998, 17) has argued that consciousness is one of the “critical” functions (recall that “criticality” is essentially undefined in this context). Even granting him that point,⁵ still, the category of integrating unifier subsumes the category of consciousness. It is not qua conscious that consciousness is a critical function (he argues), rather, it is qua integrating unifier that consciousness is a critical function. Because of this, consciousness is a red herring here, and the fact that we do not yet have machines that can simulate neural activity to a degree sufficient to produce conscious intelligence is irrelevant to the question of whether brain death is death. Shewmon’s counterexamples and arguments have shown that (i) above is false. That individuals (such as those in a persistent vegetative state) can be both permanently unconscious and alive shows that (ii) is false. We must guard against conflating the two, as this leads to yet another non sequitur.

In his 2006 article, Bernat argues that although the brain death doctrine may be imperfect, we must see it within the field of public policy and be willing to make some compromises in order to achieve good outcomes or good social policy. But this is just another way to change the subject. “Is it good policy?” is different from “is it true?”. Social policy must be built around reality. We cannot shape the world to fit our policies but must shape our policies to fit the world. Once we have an adequate grasp of just how imperfect the brain death doctrine really is, it becomes a separate question to ask if we should nonetheless maintain the brain death fiction as social policy. This is a separate question, to which I will return in Section V.

Hence, the empirical hypothesis that the brain is the body’s integrator, without which the body would cease to function as a unified whole, is false. Second, the revelation that biologically living organisms exist that do not have a central integrator shows that the empirical hypothesis, qua proposed justification for equating death with brain death, was a confused one to begin with. An organism with a completely nonfunctioning brain but maintained on a ventilator is no more dead than an individual with nonfunctioning kidneys that is maintained on a dialysis machine. Finally, this allows us to see that it does not matter if Bernat or someone else can give us a noncircular definition of “critical functions,” as well as provide a non-ad hoc justification for using clinically observable functions as their criteria. The retreat to critical and clinical functions is an attempt to allow for those cases where the loss of neurological function is incomplete (such as when neurohormonal function remains), as cases of death. Since an individual with a *completely* nonfunctioning brain is not necessarily dead, then an individual with only a partially nonfunctioning brain is not necessarily dead either.

As we can see from above, there is a great deal of sliding back and forth from one question to another in the standard arguments on behalf of the brain death doctrine. The positive arguments in favor of the brain death doctrine, and the replies to the numerous objecting arguments, are almost all non sequiturs. This results in a great deal of confusion in the literature. Another distinct issue that also looms large (and is a further source of confusion and dialectic shifting) is the relation of organ donation and the dead donor rule to brain death. Let us now turn to that, but as our discussion continues, keep in mind that whether brain death is death is *independent* of whether we ought to abandon the dead donor rule. Although they are related, they are separate questions.

III. REEVALUATING THE DEAD DONOR RULE

An individual who has suffered massive brain damage is no longer able to breathe on her own, is not conscious nor will ever regain consciousness, and has a terribly poor prognosis: She will likely suffer cardiac arrest within a few

days even with intensive support. However, as demonstrated by the arguments discussed above, while being close to death this individual is not yet dead. There seems to be tension between the concept of biological death that is implicit in the arguments above and a socio-legal or moral notion of death, which would imply that it is morally permissible to remove this individual's vital organs and hence, coupled with the dead donor rule, implies that she is in fact dead. By taking careful notice of the moral principles that underlie the dead donor rule, I suggest that we will find a resolution of this tension.

We begin by examining the dead donor rule, which states that an individual must be dead before organs can be procured and further that such procurement cannot be the proximal cause of death. There are two fundamental ethical principles that justify the dead donor rule: the principle of non-maleficence and the principle of respect for autonomy.

Non-maleficence is the duty to do no harm, and the dead donor rule instantiates this principle by ensuring that vital organs are not removed from living individuals, which at least *prima facie* constitutes a grave harm. Respect for autonomy in the health care setting is realized by (among other things) allowing patients to make their own decisions regarding their course of treatment, including the refusal of treatment even for potentially life-threatening conditions. It also encompasses the duty to obtain informed consent. Although the dead donor rule does not directly instantiate this ethical principle, current organ procurement practice does, by requiring informed consent for organ retrieval either through an advance directive or by a proxy decision maker. Our ethical standards for organ donation, which include the dead donor rule, are justified by these more fundamental principles. Irrespective of the problems with declaring death by neurological criteria, it is important to investigate if these more basic principles are fully realized by our current guidelines or if there is a better way to implement them.

I suggest the following. Individuals that have met certain requirements, some of which might be meeting the current diagnostic standards for what is now known as brain death, ought to be allowed to donate their organs even though they are in fact not yet dead. Other possible requirements, for example, would be to include those individuals who do not meet the diagnostic criteria for brain death yet are terminally ill and maintained on life support and who have chosen the removal of life support. These individuals are currently considered candidates for donation after cardiac death. Rather than attempting to orchestrate their death in the operating room, waiting for cardiac death as their vital organs are starved of oxygen (as is current practice), my proposal allows for this class of individuals' organs to be removed while allowing the organs to be perfused with oxygenated blood right up until the moment of retrieval, thus avoiding the problem of warm ischemia time and perhaps increasing the pool of viable organs.

At first glance my suggestion and current practice might seem radically opposed since I suggest that it is morally permissible, in some highly circumscribed

cases, to remove vital organs from an individual who is not yet dead, whereas, on the other hand, the dead donor rule is supposed to protect organ donors from this very practice. And the two sets of guidelines are very different, but with deep and important similarities as well. Both seek to follow the principle of non-maleficence and protect organ donors from harm. The difference is that on my proposal, in specific cases, individuals would be allowed to donate organs, even though they are not yet dead. Potential donors are protected from harm equally well in both cases since (given the requirements on being a donor) the harm of death is both unavoidable and minimal: Either the donor must meet the diagnostic standards for what is currently known as brain death (and hence is ventilator dependent, has suffered massive brain injury, and is permanently unconscious) or, as in the second case, the person is terminally ill and on life support equipment and has already chosen the removal of such equipment. Further, and this is a crucial point, there would be no difference at all with respect to the allowable candidates for donation.⁶ The same individuals that are candidates for brain death donation now would be candidates under my proposal (the only difference is that we recognize, under my proposal, that they are not yet dead). The same individuals that are candidates for donation after cardiac death would also be candidates under my proposal, and we may set the requirements in such a way that no one else would be a candidate for donation. Since the very same individuals would be allowable candidates under both proposals, it follows that the same protections from harm are in place in both instances.

There would be one slight difference in candidates for donation that we cannot avoid. Sometimes a candidate for donation after cardiac death chooses to donate, is in the operating room, the ventilator is disconnected, and then the patient begins to respire spontaneously and fails to undergo cardiac arrest despite all medical predictions to the contrary. In that case, under current standards the patient would be taken back to the ICU and would not become a donor, whereas under my proposal the patient's organs would have been removed while she was still on the ventilator and would have died in the operating room as an organ donor. It seems that the same protections from harm are not in place under both proposals since this patient would have been protected from death on current standards but not on my proposal. As I have said above, in these sorts of cases, death is a minimal and unavoidable harm. In the rare case under consideration here, although the patient does not undergo immediate cardiac arrest, she is still at the end stage of a terminal illness. So long as the patient (or her proxy decision maker) has made an informed, autonomous decision to donate her organs under the present circumstances, the harm of dying in the operating room today in order to donate her life-saving organs is no different than the harm of her dying tonight or tomorrow in the ICU. Although the harm of death is the same in both circumstances, the benefit of organ donation is present under my proposal, but not under current standards.⁷

Additionally, it is worth mentioning that although I recommend abandoning the dead donor rule, we do not currently follow the dead donor rule anyway. As demonstrated by the arguments above, brain death patients maintained on a ventilator are still alive at the time of organ removal, and the process of removal is the proximate cause of their death. So in a certain sense, I am not proposing any changes at all to current organ *removal* practices. What I am proposing is a change to the process of obtaining consent.

Both sets of guidelines are similar in that they seek to implement the principle of respect for autonomy by mandating informed consent and allowing for refusal to donate. Current practice, however, is woefully inadequate in this respect. Although consent is obtained, it is not by any means an informed consent because the physician informs the relatives of a patient that their loved one is already dead, when this is not in fact the case. This constitutes a failure to respect autonomy since the decision makers are not provided the benefit of reasonably complete and accurate information about a particularly important element of the decision-making process. The proposal on offer, by contrast, allows for greater respect for autonomous decision making by mandating that decision makers (whether in the form of surrogates or through an advance directive) understand the circumstances surrounding organ donation before they consent, thus allowing for an informed consent.

A crucial source of confusion lurks here. As briefly mentioned in the introductory section, there seems to be more than one concept of death, or more than one meaning of the word “death,” that is at issue in these debates. On the one hand, there is the biological notion of death, which is not species specific and applies to all (or at least most, sufficiently similar) organisms. But, on the other hand, there is a socio-legal or moral concept (which is specific to our species) in which what is at issue is the moral standing of individuals as members or nonmembers in a moral community. For example, and most relevantly for this discussion, when is it permissible to remove vital organs? The concept of brain death and the practice of declaring death by neurological criteria could be thought of as an outgrowth of this latter notion of death, coupled with the well-intentioned desire to free up ICU beds, relieve families of undue suffering and confusion, and make organs available for donation. Thus, when a physician declares an individual brain dead and thus, dead, and further informs the family of this diagnosis, she is really using the word “dead” in an extended or metaphorical sense. This is acceptable, so the argument might go, since many words have evolved to take on various meanings, and using “death” in this way is actually a reflection of a social policy decision and a moral decision that it is permissible to treat this particular body as if it had lost all biological functioning, even though it has really only lost most, but presumably the most crucial, biological functions.

This response, in addition to engendering confusion about death and brain death, constitutes a failure to respect autonomy. When a physician declares a patient dead by cardiorespiratory criteria and offers condolences to the

family, the word “dead” has one meaning (for clarity, let us call this “dead-1”), but when a physician declares a patient dead by neurological criteria and offers the option of organ donation, the word “dead” (let us call this “dead-2”) has quite a different meaning. Instead of signaling the irreversible loss of biological functioning, in this second context, “dead-2” means that physicians, ethicists, and policy makers have decided that it is morally permissible to remove this individual’s vital organs. But this is not what the physician presumably tells the family; instead, the physician informs the family that, in fact, their loved one is *dead*. However, no one outside of the medical or bioethics community is privy to this new, technical usage of the word. Even if it is the case that “death” has taken on this new meaning, the use of it as such constitutes technical jargon. The use of jargon in a conversation for informed consent (for any procedure) is inappropriate, misleading, and can even be disingenuous, whether intentional or not. In the case under discussion, it is particularly unacceptable because “dead-2”, used as technical jargon, is homonymous with the more colloquial word “dead,”⁸ and thus can only create confusion. Rather than having an emotionally charged word take on extended meanings, it is better to make our policy decisions explicit. Thus, whether “dead” has taken on a new meaning or if “dead” has the same meaning regardless of how death was declared, current ethical standards fail to respect patient autonomy by providing false or misleading information to those involved in making the decision about whether to donate organs. The proposal on offer does not suffer this defect (see Collins, 2009, for a more detailed defense of the claim that consents obtained for organ removal are not informed).

A similar proposal has been made by Truog and Robinson (2003), although with the following important difference. These authors suggest making a distinction between the legal interpretation of the cause of the patient’s death and the actual proximal cause of the patient’s death, in such a way that the procurement of organs is not deemed to be the legal cause of death but rather that the underlying disease or injury is. The authors argue that their approach has the advantage of focusing on the more salient ethical issues of protection from harm and respect for persons (which I have been treating as respect for autonomy, but it comes down to the same thing in this context), while avoiding the inconsistencies and conceptual confusion inherent in current practice. I support their proposal, with the exception that introducing an artificial distinction between the legal cause of death and the actual cause of death only trades one conceptual confusion for another.

Removing vital organs from a living body is the proximal cause of death, although in the circumscribed cases, I as well as Truog and Robinson argue that to do so is not to commit a morally prohibited but a morally permitted act. As such, it should not be considered a crime. As an analogous situation, consider that to insert an endotracheal tube with the patient’s permission (or presumed permission) is not a crime, but with the explicit denial of permission to do so is to commit assault. Either way, the proximal *cause* of the

tube's insertion in the patient's trachea is the physician or paramedic. In the one instance, it is morally permissible (or perhaps required) and not a crime, but in the other, it is morally prohibited and a crime.

Similarly, to remove someone's vital organs is to be the proximal cause of death. In an instance in which either the person did not meet the proposed requirements or in the absence of consent, to do so is to wrongfully cause death, is morally prohibited, and should be considered a crime. But when a patient meets those diagnostic standards and consents to the procedure either through an advance directive or proxy decision maker, to remove her organs should not be a crime any more than inserting an endotracheal tube under the appropriate conditions is assault.

There is a normative aspect that is essential to the concept of murder. To murder is to cause wrongful death. Causing death, yet not *wrongfully* causing death, is not murder, although it is homicide. Most normative theories, as well as all legal codes, make a distinction between causing death (homicide) and wrongfully causing death (murder). What I am arguing is that under the appropriate conditions, to cause death by the removal of organs is morally permissible, and thus does not constitute the wrongful causing of death, and hence should not be criminal. Although there is no need to introduce artificial distinctions between the legal and actual cause of death, there is a need to clarify homicide laws. Some homicides are legally justifiable, such as the use of lethal force under the appropriate circumstances when acting in the capacity of a law enforcement officer. Similarly, laws will need to be revised to reflect the fact that causing death under the appropriate circumstances when acting in the capacity of a transplant team is both morally and (should be made) legally justifiable.

In review: The standard diagnostic tests for determining brain death are unreliable since a patient can meet the requirements of those tests yet maintain certain neurological functions. Further, the loss of all neurological function does not equal the death of the organism any more than the loss of all kidney function does. Current guidelines that include the dead donor rule are morally justified, but they are justified by more fundamental ethical principles—the principles of non-maleficence and respect for autonomy. However, our current guidelines are not the best *implementation* of those principles. My proposal avoids the uncertainties and inconsistencies of declaring death by neurological criteria and further constitutes a better implementation of those principles that justify our current guidelines. It protects the very same individuals from harm and does a far better job at respecting autonomy. It does, however, face formidable obstacles. In the following section, I address some of these challenges.

IV. CHALLENGES

First, there are legal and political obstacles. Whether or not my arguments with respect to the dead donor rule and organ donation are accepted, the

Uniform Determination of Death Act has gotten it wrong. Brain death is not death, and so an individual who has suffered complete loss of neurological function is not necessarily dead. State legislatures need to revise the legal definition of death.⁹ This is independent of whether it is deemed acceptable for physicians to be the proximal cause of a patient's death by removing their organs under the appropriate circumstances. I also propose further revisions to the legal code, where causing death by organ removal under the appropriate circumstances would be legally justified. I have argued that it is morally justified; the legal justification should follow from that.

There is an important political obstacle that deserves mention and that comes from the transplantation community itself. Many people have built careers around their role in a transplant team. Hospitals and medical centers gain prestige, and hence business, by being large and successful transplant centers. Physicians gain credentials, stature, and money by being proficient transplant surgeons (see also Rhodes, 2002, 349, for a similar point). This creates a financial motivation for maintaining the status quo. Although I have argued that organ *removal* practices should remain the same, we must be realistic about what I am proposing as many people may find it radical or eccentric.

That the intentional killing of an innocent person is wrong, regardless of what good consequences may ensue, and regardless of whether that person competently requests it, is one of the most firmly entrenched principles in western society. Once careful consideration is given to the various fallacious arguments in defense of the brain death doctrine, the public in general will begin to realize that brain death is not death and that organ removal kills organ donors. It is altogether likely that at that point, Truog's, Robinson's, and my suggestions regarding the abandonment of the dead donor rule will not be adopted, and instead, the organ transplantation enterprise will be severely restricted to things like kidney or partial liver transplantation, which can be removed without killing the donor. Thus, publicly recognizing the deep flaws in the brain death doctrine may very well result in the near collapse of the transplantation enterprise. This is why there is a financial motivation for maintaining the status quo. (Whether a utilitarian argument for preserving the status quo may be made on this basis is something that I will address in the following section.)

In addition, consider the following moral objections. First and foremost, many people hold the injunction to refrain from killing an innocent person, regardless of consequences or its competent request, as inviolable. Second, I have claimed that under some circumstances, the harm of death is both unavoidable and minimal and is outweighed by the benefit of allowing the autonomous decision to make an anatomical gift by choosing the manner of one's death, as well as the benefit to the organ recipient. However, some may object that the harm of death is never minimal and cannot be outweighed. Or at least, the sense in which that harm is minimal needs clarification.

Most ethical systems make the distinction between causing death and wrongfully causing death. Killing as a last resort in self-defense, or defense of an innocent victim, for example, is permitted or justifiable under almost all moral codes. Because of this, it follows that killing is not *always* wrong, so the dialectic must then move to cases: under which circumstances would killing be acceptable?

The dispute here is not whether some cases of killing are morally permitted, as most everyone distinguishes killing from wrongfully killing. The dispute is whether directly and intentionally killing an innocent person (i.e., one who is not physically assaulting another in a life-threatening way) is always wrong. In order to get at that question, let us first examine whether the harm of death, in the cases under consideration, can be considered “minimal” as I have claimed above.

Recall that with the criteria for acceptable donors that I have suggested, the patient would either have satisfied the diagnostic tests for brain death or would have satisfied the current criteria for donation after cardiac death. Arguments against my claim that the harm of death is minimal that are based on considerations of *quality* of life would not work since their quality of life is by hypothesis nonexistent (they are permanently unconscious) or unbearable (because they have already independently opted for the removal of life support, knowing that death will occur). But if it is not quality of life that matters in the objection under consideration, then it must be life itself. This argument must depend on one of three claims. First, there is the claim that all killing is intrinsically wrong, but as we have noted above, this position should not be accepted. Second, there is the claim that death itself is intrinsically bad and is a harm to be avoided. I do not accept this claim, but I also do not know how to argue directly against it, for the very reason of its fundamental nature. However, in general, we as a society do not accept this claim. If we did, then we would find it morally repugnant to allow a person to discontinue or refuse life support or to choose to die at home without excessive medical intervention. Attempting cardiopulmonary resuscitation would be a mandate, not an option, and the very idea of a do not resuscitate order would not be contemplated. So although I do not have an argument against this second option, no one (or almost no one, I suppose) accepts it anyway. Finally, there is the claim, not that all killing is inherently wrong, but rather that the intentional killing of an innocent human is inherently wrong and must be avoided even if that individual desires or asks for death and has made that request with full decisional capacity.

The heart of the disagreement lies here. Note that if the intentional killing of an innocent human being cannot be justified, then our current practice of removing organs from brain dead but living donors cannot be justified either. However, there do seem to be scenarios where compassion or mercy seems to override the normal prohibition against intentionally killing. Imagine, for example, a person trapped in a fire, with no means of being

removed from the fire or the fire being extinguished. Imagine further that there is an onlooker with a loaded gun who is a skilled marksman. Out of a compassionate desire to relieve this person of her suffering, the marksman shoots the trapped victim several times in the head and chest, ensuring a rapid death. I contend that in this case, the actions of the marksman, although difficult and perhaps even distasteful, are morally permitted. In this particular case, it is morally acceptable to intentionally cause the death of another (innocent) human being. If this is so, then this case constitutes a counterexample to the general, universally quantified principle that says that intentionally killing an innocent person is *always* wrong. Without that general principle however, the dialectic must then move to cases.¹⁰

Thus, it is not that killing is always wrong or even that the intentional killing of an innocent person is always wrong because plausible counterexamples exist for both claims. It is not that death itself is always a harm. Nor is it the loss of life's quality or the quality of experience that would make death harmful in the cases under consideration. If not any of these, then what?¹¹ Without a doubt, the intentional killing of anyone is *almost* always wrong, but it is not always or intrinsically wrong, so the dialectic must move to cases. As far as I can see about the particular cases under consideration, it is morally permitted because patients are protected from harm by requiring that they meet current standards for donation such as brain death or ventilator-dependent terminal illness, and their autonomy is respected by providing them or their surrogates with adequate information to make an informed, autonomous decision.

At the core of my position is the claim that we must respect patient autonomy and protect individuals from harm. Why then do I place the requirements for allowable candidates for donation where I do? It is consistent with my position that those requirements are much different from where they are now. What about, for example, patients in a persistent vegetative state or even those who simply choose rational suicide? Should they be allowed to become donors?

Prima facie, there is an important difference between a person who has suffered a herniated brainstem and subsequent brain death from a teenager with a badly broken leg and in severe pain, who, in the throes of great pain, asks to become an organ donor. In the latter sort of case, a paternalistic refusal to allow the donation is justified. By contrast, consider a person who is far more gravely ill and competently requests to be a donor yet does not satisfy the requirements for acceptable candidates suggested above. In this case, it is not so clear that refusing the request would be justified. There is a line here, but it is blurry and difficult to discern.

I have chosen the requirements that I have because they are relatively conservative, and precisely because they coincide with the class of individuals that are currently considered to be candidates for donation. Cognizant of the fact that just as the practice of declaring death by neurological criteria did

not develop overnight, neither will its abandonment, I have chosen conservative requirements for the allowable candidates in the hope that it may smooth the process. It is possible to implement the changes that I, Truog, and Robinson recommend, without coming to a consensus on where the acceptable dividing line is between candidates and noncandidates for donation. Although I do think that placing that dividing line in a somewhat less conservative position is morally defensible, that is a separate debate that can be deferred for now. It is more important to begin to implement some change than to attempt to change too much all at once.

V. IMPLICATIONS AND PUBLIC POLICY

Since brain death is not death, surgeons are already killing their patients by removing organs from brain dead but living patients. Further, they are doing so under the false assumption that the patient has already died. This is disingenuous, and it constitutes a failure to respect the autonomy of those who seek to be (or not be) organ donors, by withholding a crucial piece of information from the decision-making process and precluding the possibility of an informed consent. I do not claim that it is intentionally disingenuous or that the killing that occurs is morally wrong. In fact, I have argued that the killing would be morally acceptable, were adequately informed consent obtained. However, current organ procurement practice does suffer from a moral failing because transplant teams fail to respect their patients' right to make autonomous decisions.¹²

I do not dismiss the very strong pull of a blanket prohibition against killing, even when it apparently causes no harm or serves a greater benefit, and is competently consented to or requested. There is something distasteful about assisted suicide, euthanasia, etc., and what I propose here falls into the same class. That distaste perhaps arises from some other principle that I have failed to recognize, such as respect for the sacredness of life. However, it is not clear that respecting the sanctity of life requires continuing it at all costs, and if it does not, then one must show how there is any *moral* difference between allowing inevitable death to occur and causing inevitable death to occur for the sake of furthering other persons' lives and well-being. And that argument cannot be based on the inherent wrongness of killing itself, because as I have argued, although killing is wrong, it is not inherently or always wrong, and so there are cases where it is permitted. The difficult questions involve teasing the very few permitted cases apart from the rest and on what grounds. Perhaps the distaste arises from the difficulty of that task, and perhaps the distaste is a purely emotional reaction, empty of moral justification, and in need of cognitive revision in light of considered, rational moral deliberation. Either way, if that distaste cannot be assuaged in the public to any significant extent, then the only other acceptable option is to discontinue retrieving vital organs from brain dead yet living donors.

It may seem that I have missed an obvious point, which is that there is another option, and that other option is to clarify a definition of death to be used for social and legal purposes, which respects the intuition that it is permissible to remove organs from brain dead individuals, while at the same time respecting the intuition that it is not permissible to kill. In addition, that this definition and practice have become widely accepted shows that we have indeed hit upon an acceptable solution. But we have not hit upon a solution; we have created a bigger moral problem. We are doing what so many people find morally repugnant anyway by killing patients but further, are dissembling about it, not only preventing the possibility of an informed consent for the reasons discussed above but also preventing forthright public debate about when it is acceptable to remove vital organs by framing the debate as a technical medical/scientific debate about whether brain death is death.

Death, like life, is a biological phenomenon. Certainly our reactions to death and ways of coping with its occurrence are partially culturally constructed. But just as we cannot legislate the cancer patient's cancer out of her, so also we cannot simply *legislate* whether someone is dead or not. What we can as a group decide upon is when it is morally permissible to remove organs.

Essentially, we have already made that decision. Where we have gone wrong is in hiding that decision behind new meanings of old words. By framing the debate in terms of whether or not brain death is death, the debate gets framed as a technical medical debate rather than what it is: It is a moral discussion about when it is acceptable to remove vital organs, or it is a normative debate about when an individual loses her standing in the moral community. But rather than so framing it, when we claim that brain death is identical to death, we have already cut off that normative debate. Everyone agrees that dead bodies are not subject to the same protections as living bodies. Autopsies and cremation are acceptable, so why wouldn't organ removal also be? Then, when the technical claim that brain death is death gets challenged, as it has been challenged for many years now, one response is that considering brain death to be death is a clarification of the definition of "death" that is to be used for socio-legal purposes and is not the same as the biological conception of death. But this only shows that the debate never was a technical medical/scientific debate about the nature of death in the first place. If organ donation had never been a possibility, the debate over whether brain death is death would never have occurred. This discussion is and always has been a moral debate about when it is acceptable to remove organs.

I agree with the proponents of the brain death doctrine that it is morally acceptable to remove organs from brain dead donors. But I also advocate that the *reasons why* must be made explicit, open to debate and considered opposition, and not hidden behind misleading legislation or technical

jargon. Organ removal from brain dead individuals is justified because the harm of death is minimal or nonexistent, and, assuming the donor or her surrogate has given an informed consent, her autonomy is respected. It is *not* justified on the grounds that the donor is already dead, because she is not yet dead, and passing legislation that allows us to *say* that she is dead does not make that statement true.

Let us finally address the public policy issue. In his 2006 article, while acknowledging that his defense of the brain death doctrine does not address all valid criticisms, Bernat argues that “[i]n the real world of public policy on biological issues, we must frequently make compromises or approximations to achieve acceptable practices and laws”, and further, “[t]hose scholars attacking the established whole-brain death formulation have a duty to show that their proposed alternative formulations not only more accurately represent biological reality, but also can be translated into successful public policy” (Bernat, 2006, 41). Bernat’s objection that those who oppose the established wisdom ought to be able to translate their opposition into successful public policy is legitimate but very easily can lead us into yet another fallacy. There is a difference between (i) we ought to be able to translate our opposition into public policy suggestions and (ii) it would be a good thing if p were true, or, it would make for good public policy if p were true, therefore, p is true. We may call (ii) the *reverse naturalistic fallacy*. Just as “is” does not imply “ought”, neither does “it would be good policy were p to be true” imply that p is true. Hence, although I do agree with Bernat that it would be good for public policy were the brain death doctrine true, unfortunately it is not true, and its goodness as public policy is irrelevant. Statement (i), on the other hand, does have something to be said for it.

What are the criteria for a policy’s success? Bernat (2006) writes that a policy is successful if it “maintains public confidence in physicians’ accuracy in death determination and in the integrity of the organ procurement enterprise” (41), but this begs the question. Whether physicians are indeed accurate in their death determination is exactly what is at issue. When physicians use neurological criteria to determine death, they are systematically *inaccurate* in their determination of death, and so the public’s confidence is ill-founded. Certainly, public trust in the medical field is an extremely important desideratum, and is in fact necessary for the very practice of medicine. But the desire to maintain public confidence does not constitute an argument that brain death is death. It may constitute a utilitarian argument that we ought to maintain the brain death fiction despite the pervasive and insurmountable flaws in the brain death doctrine, but this is a different issue. Besides, maintaining the error so that the public does not lose confidence sounds like a fantastic way to erode public confidence.

Bernat has also said that policy is successful if it maintains public confidence in the integrity of the organ transplantation enterprise. But the very criticism on offer here is that we do *not* have a successful public policy. Our

current policy fails to respect personal autonomy. Those who wish to be organ donors have a right to understand that although they would be very close to death and permanently unconscious, they would not yet be dead. Many (I suppose) would choose to donate nonetheless. However, as I have admitted above, my moral position on killing is not mainstream, so perhaps many would choose not to donate. In the real world of politics and public debate, once it becomes public knowledge that the brain death doctrine is false, as I have mentioned above, it is certainly possible that the transplantation enterprise may be severely restricted.

Yet another separate issue is whether a utilitarian argument can be made to the effect that the consequences of significantly curtailing the transplantation enterprise are so great, and are so dire, that utility trumps respect for autonomy. On this argument, it is better to continue to save lives with transplants, and not erode public trust, by maintaining the brain death fiction *even though* the brain death doctrine is deceptive. We must note that currently we maintain the brain death fiction because of an honest mistake, not intentional public deception. The utilitarian argument on offer here proposes the intentional public deception by the medical community, since one could not propose such an argument without simultaneously acknowledging the error of the brain death doctrine. This argument may be the strongest (and only) reason to maintain current public policy, and I readily admit that it is initially quite appealing. Is autonomy so valuable that it is worth even more than human life? Are lives worth less than some academic nicety? Brain dead individuals are “as good as dead”, so what is the harm? These are reasonable questions. However, since trust is essential to and a prerequisite of the practice of medicine, widespread and intentional public deception on the part of the medical community is antithetical to the practice and institution of medicine. For that reason, we cannot accept this utilitarian argument. Further, it is not a mere academic pedantry that is at issue here.

Although I certainly concede the point that wrapped up in the brain death doctrine is the need for practical public policy, what has not been appreciated thus far is that the objections that I have discussed above are not mere scholarly discontent. This is not solely an academic integrity problem; it is a *moral* problem. Persons who wish to be or not be organ donors do not make an informed decision.

As we have seen above, the brain death doctrine has *no* good arguments in its favor and many good arguments in opposition. Insisting on the brain death fiction in spite of this, however, is for bioethicists and physicians to take the decision on whether the living donation of vital organs should be permitted in our society, out of the hands of society at large. Thus, the objections to the brain death doctrine are not made solely on scholarly grounds. Our current public policy on brain death is morally flawed.

I conclude this essay by noting that although I have been very critical of the proponents of the brain death doctrine, I must make clear that

I understand, and agree with, what I take to be their underlying motivation. Ultimately, the desire to both respect the dead donor rule and make life-saving organs available for transplantation is motivated by a respect for the sanctity and dignity of human life. That respect constitutes an admirable motivation for social policy construction, and it is something that we can all agree on. However, our social policy must be shaped first and foremost by reality, and the reality is that brain death is not death, and our defining it as such does not make it so. Respecting the dead donor rule and removing vital organs from brain dead patients are *incompatible*: one or the other has to go. In order to move forward with this issue, we must reframe the terms of this debate. Respecting the sanctity of life does not imply that we must continue it all costs, nor that we may not, in some cases, intentionally end it. This debate is, and always has been, about whether there are any circumstances under which removing vital organs from a terminally ill yet living human being is morally permitted. I argue that, sometimes, there are.

NOTES

1. A summarized version of this section appears in my paper “Consent for Organ Retrieval Cannot be Presumed” (Collins, 2009) where I argue that the concept of presumed consent does not justify removal of organs in the absence of explicit consent. A review of the brain death literature and my justification for the claim that brain death is not death are necessary to both papers, which is why I repeat them.

2. These distinctions are widely accepted in the debate over brain death but not universally. Chiong (2005) argues that the entire debate is philosophically misguided and naïve precisely because the participants accept these distinctions as foundational. He correctly notes a contemporary trend in the philosophy of language and metaphysics that rejects the possibility of “pure conceptual analysis” and with it a “naïve realism” that seeks necessary and sufficient conditions, the satisfaction of which infallibly signal the instantiation of the property in question or membership in the category at issue. In its place he suggests an approach to brain death as being a somewhat arbitrary yet reasonable and acceptable dividing line between life and death, the separation of which is inherently indeterminate. Chiong is correct to note these subtle and fundamental metaphysical issues but misses an essential point. On a deeply theoretical level that deals with the nature of meaning, identity, category membership, and the relation between thought and reality, Chiong’s point is well taken. But on a slightly less theoretical and more applied level, the question at issue *just is* whether brain death is a reasonable or acceptable dividing line between life and death. The language of definitions, criteria, and tests provides a structured framework in which to articulate and defend slightly different views that exist along a continuum of reasonableness, but the adoption of which provide striking differences both for ethical theory and for medical practice. It is therefore justified and not misguided to continue to construe the debate in these terms.

3. Many of the empirical articles cited above are dated by medical standards, but they are still relevant. First, the American Academy of Neurology has not yet changed its 1995 recommendations on the diagnosis of brain death, and the above-cited studies are relevant to that set of recommendations. Second, although the once ancillary EEG and angiogram are now becoming more commonly used in the diagnosis of brain death (see Greer et al., 2008), the confusions about significant/insignificant functions (or critical/noncritical functions) and function versus activity are still prevalent, and those confusions arose as a result of these original empirical studies. Third, the debate those studies engendered persists to this day. Finally, as I state in the text, it would not matter even if we had absolutely reliable diagnostic measures for brain death since the reasons for equating a dead brain with a dead human remains confused and unjustified.

4. This has qualifications. Blood still circulates carrying hormones with it, cerebrospinal fluid may circulate as well, and the 12th cranial nerve exits the skull at a different point than the rest so that its

subcranial functions (movement of the shoulder and neck) are sometimes preserved. But these qualifications do not obviate the point made above.

5. We need not grant this point, however. Bernat (1998) claims that “consciousness . . . is required for the organism to respond to requirements for hydration, nutrition, and protection, among other needs” (19), but this is not the case. Responsiveness to the environment (both internal and external), or wakefulness, is what is required to serve these functions. This has been termed “creature consciousness,” and it is distinct from “transitive consciousness,” or consciousness *of* something. This is in turn usually thought of as distinct from “state consciousness,” which is a property of mental states not organisms (Rosenthal, 1997, introduced these terms). Finally, “phenomenal consciousness” is usually associated with the subjective, qualitative aspects of what it is like to be a conscious organism (Block, 1995, introduced this term). I have no wish to enter into a debate about the nature of consciousness here, but it suffices to note that it is not so simple a phenomenon as Bernat wishes to treat it. Because of this, we need not accept his argument that consciousness is a critical function (assuming, of course, that we can get a noncircular definition of “critical function”).

6. Or at least, it is consistent with my proposal to set the standards for allowable organ donors in such a way that there is no difference between those currently considered candidates and those considered candidates under my proposal. I have more to say on who should be allowed to become donors in the text.

7. There is a further ancillary point that deserves mention, at least in an endnote. The nature of consciousness is one of the greatest theoretical quandaries of our time. Progress has been made by dissociating wakefulness and alertness from the subjective feel of phenomenal consciousness or qualia, and it is legitimate and justified to claim, for example, that if the ascending reticular activating systems are permanently damaged, then that person will not wake up again. However, as far as some residual level of subjective phenomenal consciousness or the ability to feel pain is concerned, it is highly speculative on both sides to make claims about what the person can or cannot “feel” or “be aware of” partly since these terms themselves are so theoretically laden and encompass many different uses. Especially when we consider the fact that tachycardic and hypertensive responses to the midline incision for organ retrieval have been observed, it seems prudent to afford every individual the protection of general anesthesia, just in case. Again, this is an ancillary point, and none of my arguments rest on an understanding of consciousness or any claims to the effect that brain dead individuals have some residual awareness of pain during organ retrieval. I mention it in this endnote, however, because it would be a tragedy to allow even the possibility that some individual might suffer pain during organ donation. My proposal allows for anesthesia and thus does a better job of protecting from harm. It would be inconsistent under current guidelines to anesthetize organ donors since the very justification for the removal of their organs is that they are already dead. We do not anesthetize corpses before autopsy, so why would we anesthetize “corpses” before organ retrieval? This is not such a bizarre proposal as it might perhaps sound: There is a debate among anesthesiologists about whether organ donors should in fact receive anesthesia for organ procurement (Dalglish, 2000; Keep, 2000; Poulton and Garfield, 2000; Young and Matta, 2000).

8. Note that the meaning of the colloquial word “dead” is far more similar to “dead-1” than to “dead-2.” Both the family dog and the family matriarch can be dead (colloquial use here), and they can both be dead-1, but only a human can lose her standing in a (human) moral community due to massive and irreversible brain injury. I use the parenthetical “human” because all sentient beings should have moral consideration and thus included in the moral community. But in this context, we are interested in a subset of the full community that includes only humans.

9. I advocate revising that definition to a circulatory-respiratory one. Arguing for that is outside the scope of this paper. However, none of my arguments in this paper depend in any crucial way on that further thesis.

10. There are of course other scenarios that seem to make the same point, such as Philippa Foot’s famous Trolley Problem. If it is morally acceptable to avert the trolley car, then it is morally acceptable, in at least one case, to intentionally cause an innocent person to die. If so, this also constitutes a counterexample to the universally quantified statement that killing an innocent person is *always* wrong. As I say in the text, without that general principle, the dialectic must then move to cases.

11. The friends and family members of the deceased of course suffer harm from their loss. However, (i) respecting autonomy does not dictate that if the patient makes an informed decision with full capacity to, say, discontinue life support, even without family agreement or acceptance, then the family’s wishes override that of the patient. A similar principle would apply here. Further, (ii) in the cases under

consideration the harm to the family is unavoidable since the patient's death is impending anyway, so there is no difference in the calculus of harm done to the family in that respect. Finally, (iii) it is at least plausible that some small comfort for the family can be taken in knowing that their loved one had chosen to help someone else upon their death.

12. Note that respecting the decision of a proxy is, in essence, respecting the autonomy of the patient. When the patient lacks capacity and has no written or otherwise known wishes regarding the matter, the proxy serves, literally, as a stand-in for the patient, and the proxy's decisions are to be respected as if they were those of the patient. There are of course special issues that arise in evaluating surrogate decision making, but these are irrelevant to the points made here.

ACKNOWLEDGEMENTS

I wish to thank my anonymous reviewers at the *Journal of Medicine and Philosophy*. This essay has been substantially revised, and improved, as a result of their comments and suggestions. Additionally, I presented an early draft of this paper at the 2007 meeting of the American Society for Bioethics and Humanities, and I thank the audience and panel members for lively and thoughtful discussion.

REFERENCES

- American Academy of Neurology, Quality Standards Subcommittee. 1995. Practice parameters for determining brain death in adults (summary statement). *Neurology* 45:1012–4.
- Anziska, B. J., and R. Q. Cracco. 1980. Short latency somatosensory evoked potentials in brain dead patients. *Archives of Neurology* 37:222.
- Barelli, A., F. Della Corte, R. Calimici, C. Sandroni, R. Proietti, and S. I. Magelli. 1990. Do brainstem auditory evoked potentials detect the actual cessation of cerebral functions in brain dead patients? *Critical Care Medicine* 18:322–3.
- Beecher, H. K. 1968. A definition of irreversible coma. Report of the Ad Hoc Committee of the Harvard Medical School to examine the definition of brain death. *Journal of the American Medical Association* 205:337–40.
- Bernat, J. 1998. A defense of the whole-brain concept of death. *Hastings Center Report* 28:14–23.
- . 2006. The whole-brain concept of death remains optimum public policy. *Journal of Law, Medicine, and Ethics* 34:35–43.
- Bernat, J., C. Culver, and B. Gert. 1981. On the definition and criterion of death. *Annals of Internal Medicine* 94:389–94.
- Block, N. 1995. On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18:227–47.
- Chiong, W. 2005. Brain death without definitions. *Hastings Center Report* 35:20–30.
- Collins, M. 2009. Consent for organ retrieval cannot be presumed. *HEC Forum* 21:71–106.
- Dalglish, D. 2000. Brain stem death: Healthcare workers have difficulty accepting current management. Letter. *British Medical Journal* 321:635.
- Deliyannakis, I., F. Ioannou, and A. Davaroukas. 1975. Brainstem death with persistence of bioelectric activity of the cerebral hemispheres. *Clinical Electroencephalography* 6:75–9.
- Fackler, J. C., J. C. Troncoso, and F. R. Gioia. 1988. Age-specific characteristics of brain death in children. *American Journal of Diseases in Childhood* 142:999–1003.
- Ferbert, A., H. Buchner, E. B. Ringelstein, and W. Hacke. 1986. Isolated brainstem death. *Electroencephalographic Clinical Neurophysiology* 65:157–60.
- Fiser, D. H., J. F. Jimenez, V. Wrape, and R. Woody. 1987. Diabetes insipidus in children with brain death. *Critical Care Medicine* 15:551–3.

- Greer, D. M., P. N. Varelas, S. Haque, and E. F. M. Wijdicks. 2008. Variability in brain death determination guidelines in leading US neurologic institutions. *Neurology* 70:284–9.
- Grenvik, A. 1978. Cessation of therapy in terminal illness and brain death. *Critical Care Medicine* 6:284–9.
- Grigg, M. M., M. A. Kelly, G. G. Celesia, M. W. Ghobrial, and E. R. Ross. 1987. Electroencephalographic activity after brain death. *Archives of Neurology* 44:948–54.
- Hill, D. J., R. Munglani, and D. Sapsford. 1994. Haemodynamic responses to surgery in brain-dead organ donors. *Anaesthesia* 49:835–6.
- Hohnegger, M., M. Vermes, W. Mauritz, G. Redl, P. Sporn, and P. Eiselsberg. 1990. Serum vasopressin levels in polyuric brain-dead organ donors. *European Archives of Psychiatry and Neurologic Science* 239:267–9.
- Keep, P. J. 2000. Anaesthesia for organ donation in the brainstem dead. Letter, comment. *Anaesthesia* 55:590.
- Outwater, K. M., and M. A. Rockoff. 1984. Diabetes insipidus accompanying brain death in children. *Neurology* 34:1243–6.
- Pennefather, S. H., J. H. Dark, and R. E. Bullock. 1993. Haemodynamic responses to surgery in brain-dead organ donors. *Anaesthesia* 48:1034–8.
- Poulton, B., and M. Garfield. 2000. The implications of anaesthetizing the brainstem dead: 1. *Anaesthesia* 55:695.
- President's Commission, for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. 1981. *Defining death: Medical, legal, and ethical issues in the determination of death*. Washington: US Government Printing Office.
- Rhodes, R. 2002. Justice in transplant organ allocation. In *Medicine and social justice* (pp. 345–61), eds R. Rhodes, M. P. Battin, and A. Silvers. New York, NY: Oxford University Press.
- Rodin, E., S. Tahir, D. Austin, and L. Andaya. 1985. Brainstem death. *Clinical Electroencephalography* 16:63–71.
- Rosenthal, D. M. 1997. A theory of consciousness. In *The nature of nonconsciousness: Philosophical debates* (729–53), eds N. Block, O. Flanagan, and G. Güzeldere. Cambridge, MA: The MIT Press. A Bradford Book.
- Schrader, H., K. Krogness, A. Aakvaag, O. Sortland, and K. Purvis. 1980. Changes of pituitary hormones in brain death. *Acta Neurochirurgica* 52:239–48.
- Shewmon, D. A. 1999. Spinal shock and “brain death”: Somatic pathophysiological equivalence and implications for the integrative-unity rationale. *Spinal Cord* 37:313–24.
- . 2001. The brain and somatic integration: Insights into the standard biological rationale for equating “brain death” with death. *Journal of Medicine and Philosophy* 26:457–78.
- Tomlinson, T. 1984. The conservative use of the brain-death criterion—A critique. *Journal of Medicine and Philosophy* 9:377–93.
- Truog, R. D. 1997. Is it time to abandon brain death? *Hastings Center Report* 27:29–37.
- Truog, R. D., and W. M. Robinson. 2003. Role of brain death and the dead-donor rule in the ethics of organ transplantation. *Critical Care Medicine* 31:2391–6.
- Wetzel, R. C., N. Setzer, J. L. Stiff, and M. C. Rogers. 1985. Hemodynamic responses in brain dead organ donor patients. *Anesthesia and Analgesia* 64:125–8.
- Wijdicks, E. F. M. 1995. Determining brain death in adults. *Neurology* 45:1003–11.
- Young, P. J., and B. F. Matta. 2000. Anaesthesia for organ donation in the brainstem dead: Why bother? *Anaesthesia* 55:105–6.