

HOW MUCH OF THE BRAIN MUST BE DEAD?

Baruch A. Brody

The proponents of the standard criterion of brain death (death occurs at that point in time when there is an irreversible cessation of functioning of the entire brain) encounter difficulties in reconciling it with the definition (irreversible loss of integrative functioning) and the clinical tests (no stem reflexes, no respiratory efforts, no responsiveness) normally associated with that criterion. The solution to this problem is neither to defend the standard criterion by modifying the tests or the definition nor to look for another criterion based on another definition and to employ other tests that confirm the satisfaction of that criterion. Rather, one should recognize that (1) criteria of death postulate a particular point as an answer to a series of questions, (2) death is a process rather than an event that occurs at a particular point in time, and (3) the answers to these different questions are to be found at different points in the process, so that no one point can be picked as the moment of death. Rather than seeking a point in the process to serve as the criterion of death and as an answer to these questions, one should choose different points in the process as appropriate answers to the different questions.

Halevy and I made these arguments in a 1993 paper.¹ This chapter expands on that position. In the first section, "Restating the Problem," I briefly review our evidence of the difficulties faced in reconciling the tests, the criterion, and the underlying definition. Under "Possible Responses," I amplify our criticisms of several suggestions that have been made in response to these difficulties, including the suggestion that advances in neurology might provide better tests that would resolve the difficulties.

Editor's note: Many footnotes have been deleted. Students who want to follow up on sources should consult the original article.

Reprinted with permission of the author and publisher from Stuart J. Youngner, Robert M. Arnold, and Rene Schapiro, eds., *The Definition of Death: Contemporary Controversies*. Baltimore: The Johns Hopkins University Press, 1999.

In the final section, "The Halevy-Brody Response," I explain our theory and use it to evaluate the current debate about procuring organs from anencephalics.

First, however, I want to make it clear that part of the intellectual background to our 1993 paper is the acceptance of the fundamental insight of fuzzy logic, namely, that the world does not easily divide itself into sets and their complements. Death and its complementary property life determine mutually exclusive but not jointly exhaustive sets. Although no organism can fully belong to both sets, organisms can be in many conditions (the very conditions that have created the debates about death) during which they do not fully belong to either. That is why you cannot find the answers to the questions by finding the right moment in the process to serve as the moment for belonging to the set of the dead. Death is a fuzzy set.

RESTATING THE PROBLEM

To understand the difficulties Halevy and I raised, one needs to remember that the whole-brain criterion of death, with its associated clinical tests, is put forward on the basis of a definition that provides its rationale. According to the definition, the organism is alive only when its functioning is integrated. Given that both the cortex and the stem play central roles in the integration of the functioning of the organism, the organism dies only when all of these integrative functions of all of the parts of the brain irreversibly cease. This is the criterion of death. The clinical tests (such as those for responsiveness/voluntary movements and apnea) test for the presence of these integrative functions.

Both the cortical criterion of death and the cardiorespiratory criterion of death are also based upon definitions that provide their rationales. According to the cortical definition, life requires the functioning of a person. Given that the cortex is the physiological location for functions (such as consciousness, thought, and feeling) that are essential

for the existence of a person, death occurs when the cortex irreversibly loses the capacity for those functions. According to the cardiorespiratory definition, the organism is alive only when the vital "bodily fluids"—air and blood—continue to flow through the organism. Given that this flow requires respiration and circulation, the organism dies when those two functions cease.

For each of these definitions, there are, of course, problems either with the definition or with the relation between the definition and its associated criterion. Parts of the body other than the brain help integrate the organism's functioning, so why does the first definition lead to the criterion of the cessation of the integrative functions of only the brain? Is it sufficient for death, as the second definition maintains, that the person has stopped functioning, or must other functions also cease before death has occurred? If the flow of the "vital fluids" is maintained artificially, is the organism still alive according to the third definition, especially if the organism is conscious and capable of responding and moving spontaneously? I shall return to aspects of these problems below. What I want to note for now is that adherents of these three competing criteria have recognized the importance of there being justifying definitions for the criteria; without such a definition, all that you have is an arbitrarily chosen criterion. This point is central for understanding the difficulties raised in our paper.

Halevy and I called attention to the fact that there are organisms who satisfy all of the standard clinical tests for whole-brain death but who have not lost all of the integrative functions of the brain. The most important example is neurohormonal regulation. The presence of this residual neurohormonal regulation in a significant percentage of organisms who satisfy the usual tests for brain death and whose respiration and circulation are being maintained artificially (usually to allow for the possibility of organ donation) is well documented. Most crucially, this regulation is just as much and as important an example of the integrative functioning of the brain as is the brain's control of respiration or of responsive movements. Given the definition behind the whole-brain criterion, this functioning of the brain should have to cease before the criterion is really met. As the usual tests do not ensure this,

they are inadequate as tests for the satisfaction of the criterion *given the definition that supposedly justifies that criterion*.

In the article in the *Annals of Internal Medicine*, Halevy and I also call attention to two other functions of the brain that do not necessarily cease when the normal clinical tests are met: (1) continued functioning of the auditory pathways as evidenced by brainstem evoked potentials and (2) continued cortical functioning as evidenced by EEG readings. I would today put less emphasis on them. To begin with, the latter is present only in very special cases and the extent of the former is not clearly known. Second, and more important, while they constitute brain functions, it is not clear that either integrates the functioning of the entire organism. If not, then both should be irrelevant to the death of the organism. According to the definition that supposedly justifies the whole-brain criterion, only the integrative functioning of the brain must irreversibly cease before the organism dies, so these functions may not count.

Halevy and I emphasized these other two functions to illustrate another problem. Patients who meet the normal clinical tests for brain death may not satisfy the criterion used by the President's commission and embodied in the law, the "irreversible cessation of all functions of the brain." These examples are very relevant to illustrate the dissonance between the clinical tests and that legal criterion. The point I am making here is that they may not be relevant to illustrating the dissonance between the clinical tests and the whole-brain criterion justified by the associated definition, a criterion that refers only to the cessation of *integrative* functions.

In short, our difficulty may be stated as follows: (1) the true whole-brain criterion of death is that the organism dies when all of the integrative functioning of the entire brain ceases; (2) when the normal clinical tests are met, at least one form of integrative functioning of the brain, neurohormonal regulation, has often not ceased, and there may be other forms of integrative functioning that have not ceased; (3) there is, therefore, an incongruity between the normal clinical tests and the whole-brain criterion as understood in light of the definition that justifies it.

POSSIBLE RESPONSES

Given the above argument, one can identify a series of possible responses: (1) The incongruity between tests and criterion exists, but we should not worry about it because it occurs in only a few cases or because it makes no difference, as current practice works well. (2) These residual functions are either not integrative functioning or not integrative functioning of the right type, so there is really no incongruity between tests and criterion. (3) There is an incongruity between tests and criterion which cannot be ignored, and we should resolve it by improving the tests employed through advances in neurology. (4) There is an incongruity between the clinical tests and the criterion which cannot be ignored, and we should resolve it by adopting some other criterion based on some other justifying definition. (5) The incongruity is indicative of a fundamental problem that is best resolved by giving up on the search for a single criterion of death that answers all of the questions that a criterion of death is traditionally understood as answering. This last response is the conclusion Halevy and I adopted.

Should we worry about the incongruity, since it occurs in only a few cases? The claim that it occurs in only a few cases is mistaken. If, of course, organisms are declared dead on the basis of the usual tests and removed from life support, neurohormonal regulation (and many other functions of the organism) will soon cease. But that may show only that taking dependent living organisms off life support soon produces death. The crucial question about the extent of the incongruity is how often these functions are still occurring when the usual tests are first met and before the organism is taken off life support. The data cited in our 1993 paper show that this occurs in a significant percentage of cases.

Should we worry about the incongruity even if it occurs in many cases, since the use of the current tests works well? It all depends upon what you mean by "the tests work well." They certainly have enabled the organ procurement programs to harvest a significant number of additional organs by declaring death on the basis of the tests without waiting for neurohormonal regulation to cease. They also have enabled physicians to discontinue

life support in many cases where the families insisted on doing everything so long as death had not occurred. The need to respect that family preference stops when death is pronounced on the basis of the current tests, even when neurohormonal regulation has not ceased. But does that mean that the tests have worked well? Not if the organisms in question still are alive, as they are according to the current criterion when regulation has not yet ceased. If they are still alive, the use of the current tests has in many cases resulted in killings to harvest organs and in discontinuing life support by misleading families about when death has occurred. It is hard to understand why that should be described as working well. I suggest below that the claim that the current tests work well can be reconstructed to make some sense once one drops, as Halevy and I have advocated, the search for a criterion of death. But until one does so, the incongruity means that the current tests are not working well.

Are the residual functions integrative functions or integrative functions of the right type? Bernat has suggested that they are not:

However, in some cases, a critical number of neurons have been destroyed but a few continue to function in isolation. For example, some unequivocally whole-brain-dead patients continue to manifest rudimentary but recordable electroencephalographic activity or hypothalamic neuroendocrine activity sufficient to prevent diabetes insipidus. Because these isolated nests of independently operating neurons no longer contribute critically to the functions of the organism as a whole, their continued activity remains consistent with the whole-brain criterion of death.² (569)

I am not convinced by this objection. Although it is true that the residual cortical activity is separated from the functioning of the organism as a whole and is in that sense an isolated nest of operating neurons, this is just not true of the neurohormonal regulation that, by definition, is integrated with the functioning of the rest of the organism. Another residual functioning that Bernat does not mention, intact auditory pathways, is harder to classify, although it is certainly not a clear-cut example of purely isolated nests of operating neurons. That is why I said above that, for the purposes of this chapter, the residual neurohormonal functioning deserves the most attention. It is without any doubt

residual integrative functioning of the very sort that is supposed to mean that the patient is still alive, according to the justifying definition that lies behind the whole-brain criterion of death. This last point also serves as a criticism of Pallis's recent response to our argument:

What is the *philosophical* significance, for instance, of a given TSH level detected a specified number of hours after a clinical diagnosis of brain death? . . . A "concept" which "dares not speak its name" in fact often lurks behind most such "challenges." It is that death on neurological grounds should mean but one thing: the irreversible loss of function of the totality of the intracranial contents . . . This is advanced without specification of the functions, the loss of which would demarcate the living from the dead. This approach hardly warrants being described as a "philosophical" concept of death.³ (21)

The relevant functions are, of course, the brain's integrative functions, and the level of thyroid-stimulating hormone (TSH) is evidence that some of them are still intact. None of this requires, of course, that all of the intracranial contents should have stopped functioning.

I confess that I am surprised by Dr. Bernat's response. After all, his 1981 paper with Culver and Gert is a landmark paper precisely because it clarified for the first time the justifying definition of continued integrative functioning ("functioning of the organism as a whole") as lying behind the whole-brain criterion.⁴ In that paper, the first example of the functioning of the organism as a whole is neuroendocrine control (390). Why, then, does Bernat now describe it as a purely isolated nest of operating neurons?

Isn't the obvious response to modify the clinical tests so that the criterion of the irreversible cessation of the integrative functioning of the entire brain is satisfied when the new set of tests is met? Couldn't we test for neurohormonal regulation and (if we were concerned about it as integrative functioning) for intact auditory and visual pathways? In fact, the use of such tests has been suggested by various authors. There is, however, a major problem, which Halevy and I noted, with this suggestion. Data from the transplant community which has studied this question to determine whether hormone replacement should be part of the management of potential donors, suggest that hormonal

levels due to residual neurohormonal functioning may remain intact for more than 72 hours. These are some of the best data available. They include patients where angiography indicated a complete cessation of intracranial blood flow, indicating the presence of a dual blood supply.

Consequently, the adoption of these new tests would mean a serious challenge to the transplant community in maintaining both the viability of organs and the willingness of families to donate. There could be a significant loss of organs. Moreover, putting aside the transplantation setting, maintaining organisms on life support until the new tests are met, when the families insist that everything be done until death occurs, can be very expensive, so we should not rush to add additional tests.

There is a crucial difference between the criticism of this response and the criticism of the other responses. The problem with the first three responses is that they fail on intellectual grounds. The justification for the brain-death criterion means that the functions which remain and are the source of the dissonance cannot be disregarded if one wants to maintain intellectual honesty. The problem with the fourth response is that it fails on practical grounds. We want to be able to harvest organs and to disconnect life support unilaterally long before the suggested new tests are satisfied. Some might suggest that this is irrelevant. If we want theoretical soundness, we must pay the practical prices. One of the merits of the proposal Halevy and I put forward is that it offers the opportunity to be both theoretically sound and practical. I will return to this point below.

Perhaps the best response is to modify the criterion of death and the justifying definition. Three versions of this response are found in the literature: (1) Adopt as the criterion of death the permanent cessation of respiratory activity or the permanent cessation of cardiac as well as respiratory activity (the two options most advocated in the debate in the Orthodox Jewish community about brain death and organ transplantation). (2) Adopt as the criterion of death the permanent cessation of respiratory activity and of consciousness (Pallis's brainstem criterion).³ (3) Adopt as the criterion of death the permanent cessation of consciousness (the higher-brain criterion).

There is more to be said about each of these suggestions than is possible in this chapter, so I will confine myself to just a few observations. (1) The adoption of the view that death requires the irreversible cessation of both cardiac and respiratory functioning may mean a significant and expensive prolongation of the dying process as well as the end of organ transplantation as we know it. A strong futility policy might avoid the former, while a modification of the Pittsburgh protocol might preserve some transplantation. Unless we have powerful intellectual reasons for preserving that criterion, other than adherence to the traditional definition, it is a poor suggestion on practical grounds. (2) Neither a purely respiratory criterion nor a combined respiratory/consciousness criterion lends itself to a justifying definition. The former criterion involves only one of the traditional vital "bodily fluids," and it is hard to see why one is to be preferred to the other. The latter criterion comes from two very different definitions, and it is hard to see why the two criteria should be combined. Pallis points out quite correctly that they are "embedded in coherent historical and cultural matrices" (21). However, the fact that each is embedded in a coherent matrix does not ensure that their combination is embedded in a coherent matrix. (3) The suggestion that we adopt a higher-brain criterion for the death of the person, based upon the definition that the person dies when the cognitive and affective functioning required to be a person ceases, makes a lot of sense when discussing the death of the person. But are we only looking for an account of the death of the person? Perhaps we really want an account that encompasses the death of the full organism? We certainly seem to want that type of account before burial or cremation.

THE HALEVY-BRODY RESPONSE

Our response to the incongruity begins by recognizing that the death of the organism is a process rather than an event. Consider the organism that suffers damage to its brain so that it is no longer conscious and can no longer engage in responsive or voluntary movements. At some later stage, it loses the capacity to breathe on its own so that its respiration must be supported artificially. At a later

stage, its capacity to regulate hormonal levels stops. Somewhere during this time period, its auditory pathways stop functioning. Finally, its heart stops beating. Is it really meaningful to suppose that the organism died at some specific point in this process? Isn't it more reasonable to say that the search for a criterion of death (a specific moment) made sense when these points were always close in time to each other because medicine lacked the capacity to protect some of the functions when the others had stopped, but no longer makes sense today when medicine can, and sometimes has good reasons to, keep some of the functions going for longer periods? Isn't it more reasonable to say that the organism was fully alive before the chain of events began, is fully dead by the end of the chain of events, and is neither during the process. Fuzzy logic enables us to say that in a precise fashion.

But don't we have to identify a specific point of time at which the organism died? Aren't there important questions which need to be answered and can only be answered by identifying the precise point in the process at which the organism died? These questions include when life support can unilaterally (without patient or surrogate concurrence) be withdrawn, when organs can be harvested, and when the organism can be buried or cremated. Perhaps not. While traditionally it has been thought that the way to answer these questions is to find that precise moment of death, perhaps that is the mistake. Perhaps these questions need to be examined and answered each on its own, with the answer to one question (some point in the process) not necessarily being the answer to the other questions. That is the heart of Halevy's and my proposal.

In our paper, we suggested that life support could in these cases be unilaterally withdrawn when the organism no longer composes a person because the cortex no longer functions. We emphasized that allowing for this unilateral withdrawal would constitute an appropriate stewardship of social resources. Elsewhere, I have argued that even those moral and religious traditions that place great emphasis on the value of the life of human organisms can accept such appeals to stewardship. Notice, by the way, that such an argument would not apply to those *rare* cases where the resources of the patient or family paid for the full costs of the continued care.

In our paper, Halevy and I suggested that organs could be harvested at that stage in the process after the loss of cortical functioning when the organism can no longer breathe on its own. This, of course, corresponds to current practice. We defended it, however, not by adopting some criterion of death justified by some definition of death. Instead, we argued for it on the grounds that it preserves the proper balance between trying to maximize the supply of organs to save lives and trying to preserve public support for organ transplantation by not harvesting organs in cases that would be socially unacceptable.

This approach offers, I believe, a basis for evaluating AMA approval—later withdrawn—of harvesting organs from still-breathing anencephalics, allowing for a reasoned consideration of their proposal while rejecting their justification of it. The AMA continues to accept the current criterion of death, with its implication that such anencephalics are still alive. They also recognize that harvesting organs means, on their own assumptions, killing the anencephalic organism, although they avoid using that word, preferring to talk instead about “sacrificing” it. To justify their conclusion, they argue that anencephalics, who have never and who never will experience consciousness, can be killed because they have no interest in being alive and there are no compelling social interests in preserving their life. This argument succeeds only if one is willing to change deontological constraints (“thou shall not kill living human organisms”) into teleological rules (“killing human organisms is wrong when their interests or social interests are harmed”). The implications of this are very disturbing.

I respectfully suggest that the Council on Ethical and Judicial Affairs adopted this change without even arguing for it because the council, following much of the recent bioethical literature, does not understand deontological constraints. Things would be very different if they argued that anencephalics are in that class of in-between organisms that are neither fully alive nor fully dead. Then, they might argue that the deontological constraints do not apply to them and that we should settle the question by balancing the benefits of additional organs (needed, e.g., by other newborns with hypoplastic left hearts) against the risks to public acceptance of organ procurement if the public does

not see anencephalics as being in this in-between category. That, I submit, could be the basis for a reasoned discussion of the AMA proposal, one following the framework presented in our paper.

There is, however, one further complication that must, I now believe, be taken into account. In our article and in the analysis just presented, the assumption is that the deontological constraint against killing human organisms applies only to those who are fully alive; once the organism is in the in-between range, we need only consider the policy trade-off. But is that assumption necessarily true? What happens to deontological constraints in a world of fuzzy sets? This additional issue will require a reasoned discussion, although the contours of the discussion are at the moment quite unclear.

What about burying or cremating the organism? Here, we suggested, maximum leeway could be given to respecting family sentiments by waiting for asystole, which usually occurs soon after all support ends. We can adopt that approach, saying that, on the basis of the traditional definition, the organism is fully dead only at that point, because that does not require us to wait for asystole before withdrawing life support unilaterally or harvesting organs.

In conclusion, then, our response answers the three questions in ways that are both theoretically defensible and practically useful. It is able to do so only because it does not answer them by adopting a consonant definition, criterion, and test of death. The dissonance we identified makes that impossible in a world that also needs to harvest organs and control health care expenditures. It is able to do so, instead, because it recognizes the implications of the fact that death is a process in a world governed by fuzzy logic.

REFERENCES

1. Halevy A, Brody B. Brain death: Reconciling definitions, criteria, and tests. *Ann Intern Med* 1993; 119:519-25.
2. Bernat JL. Brain death occurs only with destruction of the cerebral hemispheres and the brain stem. *Arch Neurol* 1992; 49:569-70.
3. Pallis C. Further thoughts on brainstem death. *Anaesth Intensive Care* 1995; 23:20-23.
4. Bernat JL, Culver CM, Gert B. On the criterion and definition of death. *Ann Intern Med* 1981; 94:389-94.