To be or not to be: legal and ethical considerations in suicide prevention†

Jaydip Sarkar

SUMMARY
Making potentially critical clinical decisions in complex cases with the real risk of death by suicide is a most challenging job in psychiatry. Sadly, risk assessment and management of harm to self is a largely ignored area compared with risk of harm to others. The legal and ethical challenges are more nuanced, and contemporary training schemes and textbooks on psychiatry have not always done justice to this area, where front-line clinicians require probably most assistance. This article is an attempt to integrate the seemingly disparate threads from legal, ethical and clinical realms to assist decision-making, and it introduces a set of principles for managing these in clinical practice. It refers in particular to legislation in England and Wales, but the clinical and ethical issues discussed are universal.

DECLARATION OF INTEREST
None.

‘There is but one truly philosophical problem and that is suicide. Judging whether life is or is not worth living amounts to answering the fundamental question of philosophy.’

Albert Camus (The Myth of Sisyphus, 1942)

Psychiatrists in England have found themselves at the centre of government strategy following its proposed reduction of the suicide rate by one fifth by 2010 (Department of Health 1999). However, Camus’ words reflect the enormity of the challenge faced by doctors and nurses asked to assess a person who is thought to be suicidal. The act of suicide is no longer illegal in the UK (in light of the Suicide Act 1961), but aiding and abetting a suicide remains so. Although mental health legislation does not mandate clinicians to prevent a person putting themselves at risk, failure to prevent suicide in an identifiable victim may put hospital trusts and clinicians at risk of civil suits for negligence. Recent case law has found trusts liable for failure to prevent the suicide of patients known to pose a ‘real and immediate’ risk of taking their own lives (Savage v South Essex Partnership NHS Foundation Trust 2008) and has considered the deaths of both formal and informal patients to be the result of negligence (Rabone v Pennine Care NHS Foundation Trust 2012), ordering trusts to pay pecuniary damages to next of kin (Reynolds v United Kingdom 2012). Heading a five-member judicial bench, Lord Walker in the Rabone case stated: ‘if there was a real and immediate risk of suicide at [the material] time of which the trust was aware or ought to have been aware, then in my view the trust was under a duty to take reasonable steps to protect [the patient] from it’. He defined real risk as ‘a substantial or significant risk and not a remote or fanciful one’ and immediate risk as ‘present and continuing’ (Rabone v Pennine Care Foundation Trust 2012).

Notably, there is major criticism of the way that risk was quantitatively calculated by the psychiatric experts in the Rabone case. The defence expert cited a risk ‘of the order of 70%’, whereas the trust expert accepted a ‘50% risk’ (Rabone v Pennine Care Foundation Trust 2009). No rationale was provided as to how these risk probabilities were calculated, and Large et al (2012) suggest that they might be affected by hindsight bias.

It is notoriously difficult to gauge the risk of suicide in self-harming individuals or to decide what treatment should be provided, who should provide it and where. One of the best validated instruments, Beck’s Suicide Intent Scale, has a positive predictive value (PPV) of only 4%, i.e. it successfully predicts only 4 out of 100 self-harming patients who will later die by suicide (Harris 2005). Nevertheless, individual professionals are likely to be seen as negligent if a patient at high risk of suicide dies in their care. Perhaps a doctor does not detain the patient ‘at the right time’ or a nurse meant to be observing the patient leaves their post for a few minutes. They are now more than ever likely to be exposed to legal action in the wake of these case judgments. At the very least, they will face disciplinary and professional hearings by their respective regulatory bodies.

In this article, I discuss three aspects of clinical decision-making with regard to suicidal or self-harming patients. First, I set out the empirical challenges encountered by clinicians in assessing risk. Second, I explore the ethical questions that arise from cases involving suicide and self-harm: especially the balance between respect for patient...
autonomy and the exercise of healthy paternalism in a climate of risk-averse health policy. Finally, I offer a set of principles that busy clinicians can use to deal with common clinical, legal and ethical problems in practice. I do not claim that these are definitive or exhaustive: only that they are based on practical experience of managing cases of high-risk, repetitive self-harm by patients with extremely complex needs.

**Clinical conundrums**

Clinical conundrums come in the form of problems such as how to differentiate a patient who is suicidal from one who is self-harming but has no wish to die. As it is notoriously difficult to treat a behaviour, for that is what suicidal acts are, should one focus on clinical syndromes? But patients with clinical syndromes such as personality disorders are very difficult to manage as they alienate carers, and those with substance use disorders cannot be treated involuntarily under English mental health law. Should one instead focus on high-risk periods or methods of self-harm?

**Whom to treat?**

Although the risk factors for suicide are well established, predicting who will actually take their own life is not so easy. For example, patients with a history of self-harm are more than four times as likely to die by suicide than those without such a history. However, people who self-harm outnumber those who take their own life, i.e. although most people who die by suicide have self-harmed, only a subgroup of those who self-harm go on to kill themselves (Appleby 2006). As noted earlier, the use of psychometric risk measures adds little in these cases.

There is also a complication and expense in enforcing mental health legislation. Many who repeatedly self-harm are not seeking help or care; indeed, some actively refuse it or assault those who are trying to prevent them from self-harming (Sarkar 2011a). Reducing their risk of dying by suicide often requires detention under the Mental Health Act 1983, at a time when there are fewer beds for this purpose. Detention may also imply support under Section 117 of the Act, which includes assertive attempts at engaging patients (e.g. telephone calls, home visits or contact with the patient’s family) (National Confidential Inquiry into Suicide and Homicide by People with Mental Illness 2009). Following the Wooltorton case (Box 1), capacitous patients may refuse medical treatment after self-poisoning and may also refuse to engage with psychological therapies recommended by clinicians. It seems clinically futile to detain someone under the Mental Health Act to prevent them from harming themselves if they are simultaneously refusing to be helped. However, the legal position is that if there is ‘real and immediate risk to life’ about which the authorities know or ought to know, they have a duty to do ‘all that reasonably could [be] expected’; failing to do so will lead to (legal) consequences that can be devastating, not just for the patient but also for the staff and the service (Savage v South Essex Partnership NHS Foundation Trust 2008). There is an interesting overlap here in terms of the ethical duties of doctors. Treating every individual who self-harms as though they are at potential risk of suicide will entail a significant increase in clinical resources. It may also result in so many false positives that true positives may be missed. In terms of best use of resources, it might be better to focus on high-risk individuals, and not only on anyone who cuts themselves (Runeson 2010). However, such an approach means that

---

**BOX 1 Capacitous refusal of life-saving treatment for self-harm**

In 2007, Ms Kerrie Wooltorton was not prevented from dying by suicide by her local hospital. She was a 26-year-old woman with a history of depression and some evidence of disordered personality traits. In the previous 12 months, she had been admitted to mental health units on several occasions; and there were nine separate incidents where she had drunk industrial antifreeze. In September 2007, in accordance with the Mental Capacity Act 2005, Ms Wooltorton made an advance refusal of life-saving medical treatment in the event of her poisoning herself. She indicated that she was fully aware of the consequences of refusing such treatment. Three days after drafting the advance refusal, she made her final suicide act of drinking antifreeze. She called an ambulance, not from a desire to have life-saving treatment, but because she did not want to die in her flat alone.

At the subsequent Coroner’s hearing, there was considerable controversy. On the one hand, a young woman was deemed to have capacity to make a decision that her life was no longer worth living, despite having recognised mental disorders and having had previous psychiatric treatment. Her history of mental disorder and treatment were not considered adequate reasons preventing her from making a capacitous decision to end her life. On the other hand, it was possible to see the mental health services as having abandoned their duty of care to a young person in a highly distressed and disordered state of mind. Note that her history and presentation contained several high-risk factors associated with completed suicide, which often lead to detention for treatment:

- she had a known, treatable mental disorder
- in the previous 12 months she had made several attempts to take her life
- she had received previous psychiatric treatment
- she gave advance notice of her intention to die through a potentially fatal method of self-harm.

Ms Wooltorton’s parents threatened to sue for medical negligence, and the government of the day stated that ‘it was not the intention of the [mental capacity] law to give legal force to such suicidal advance decisions’ (Bingham 2009a). Contrary to the case of Savage (Savage v South Essex Partnership NHS Foundation Trust 2008) the Coroner’s Court did not criticise the clinicians, but acknowledged that doctors who were aware of her history had judged Ms Wooltorton to be capacitous at the time of drafting the advance refusal. The coroner concluded that it would have been ‘unlawful’ for doctors to intervene in the wishes of a person with ‘full knowledge’ of what she was doing at the time she drafted her advance directive (Bingham 2009b).
some individuals will die by suicide because no algorithm is 100% accurate, and cases deemed to be low-risk will turn out to be occult ‘high-risk’.

**What to treat?**

Might it make more sense clinically to focus on high-risk factors such as diagnosis, not behaviour? This is true to an extent. The two most common single diagnoses among those who die by suicide are affective disorders (46%) and schizophrenia (19%); next, a large proportion (29%) have diagnoses of personality disorder, drug and alcohol misuse, anxiety or adjustment disorders (Appleby 2006). So even if clinicians targeted the two categories of serious mental illnesses that can be easily treated with pharmacotherapy, there would still be a sizeable group who would need more complex care packages; and (as already mentioned) might refuse to engage. Similarly, the sheer numbers would be a challenge for current resources. For example, if every person with a drug or alcohol problem is assessed for suicide risk, and offered treatment, this would mean establishing which of the people with these risk factors are actually at high risk of completing a suicidal act. Further, it is in the nature of suicide (as it is in fatal violence to others) that some deaths occur impulsively, in the absence of any known risk factors, or unintentionally. Hindsight bias may make such cases appear ‘obvious’ later, but this is a cognitive illusion (Kahneman 2011).

**Ethical considerations**

Philosopher and ethicist Tom Beauchamp has argued that, in Western society, people are allowed to take risks with their lives if their decisions are made capacitiously, which requires that they have ‘the ability to grasp, appreciate the significance of, form relevant intentions and not be controlled by either internal or external forces that [they] cannot resist’ (Beauchamp 1993). This capacious autonomy over one’s life is protected if a decision emerges out of rational logic, even though it results in one’s death – the so-called ‘rational suicide’ (Brandt 1980). Most modern societies consider this to be morally permissible as long as the decision is understandable or the individual’s actions bring benefit to society (e.g. soldiers going to war, firefighters risking their lives or astronauts going into outer space). However, when a risky course of action by an individual is driven solely by personal wishes and desires that appear to lack logic or reason, or have little or no benefit to society, most modern societies do not unconditionally support the individual’s personal autonomy and self-determination. For example, people who engage in extreme sports may not be able to take out life insurance: even though it is not their intention to kill themselves, society speaks through tacit discouragement of such behaviours.

But what of individuals who are under the legal age of consent, or who have intellectual disabilities or mental disorder? Are they permitted a ‘rational suicide’? No. Our society considers that the decision by such a person to take their own life is ‘irrational’ (as judged by clinicians and approved by law) and therefore morally impermissible (Brandt 1980), for their lives are considered to be ‘worth living’ (Glover 1990).

The notion of sanctity of life is championed by theologians of all hues, and most religions assert that life is a gift from the creator and taking it shows disrespect towards God (Buel 1980). It is likely that someone who is depressed forgets all that they usually get out of life. This is probably true regardless of the cause of the depression: long-term severe depressive disorder, a transient episode associated with a personality disorder, or a challenging life event such as bereavement, injury or unemployment. On the other hand, someone who fears death, possibly because of the belief that...
they would go to hell, may wish they had never been born, but still not want to die.

Who decides?

Clinicians are sometimes required to make judgements on life and death issues on behalf of their patients. The risks of allowing someone other than the person whose life it is to determine which lives are ‘worth living’ and ‘worth saving’ is a highly controversial matter (Glover 1990). It may be difficult to separate one’s own views from those of society about what constitutes a life worth living or a life worth saving. This moral decision is what often determines to which individuals a society will ‘cause death to save other lives’ (Glover 1990), which individuals are ‘allowed’ to take their own lives, which ones are forced to endure lives of tremendous hardship and pain, and which individuals are taken care of by society irrespective of how intensely they wish to die. The eugenics programmes of several nations in the early 20th century and the excesses of Nazi Germany are not too distant reminders of the risks implicit in such a process.

Mental health law is a social response to the profound ethical dilemmas posed by adults who are a risk to themselves or others as a result of mental disorder or distress. The ethical tension is between respect for autonomy and liberty to be left alone and the public duty of society to protect those who are vulnerable. All mental health legislation is a trade-off between the principles of respect for autonomy and liberty on one side and the principle of respect for welfare on the other (Adshead 2005), and on occasion it leads to medical paternalism.

Medical paternalism

For a clinician to decide whether it is ‘appropriate’ to prevent someone from taking their own life is an example of medical paternalism. The Mental Health Act 1983, which governs England and Wales, allows for individuals with a recognised (in international classificatory systems) mental disorder to be detained for appropriate treatment if their illness is of a sufficient nature or degree to require hospital treatment. This ultimately boils down to an assessment as to whether the person, as a result of the mental disorder, poses a significant danger to themselves, their own health or to others. Amendments to the Act in 2007 require a demonstration by the detaining authority that ‘appropriate medical treatment’ exists to treat the condition that the potentially detained patient has. In terms of a suicidal patient, this requires the clinician to consider whether it is appropriate to detain the person simply to prevent suicide or whether appropriate treatment can ameliorate the condition to the extent that the person no longer remains at risk of suicide at the end of treatment. This determination allows that clinician to override the person’s wishes, even if they retain mental capacity to make a reasoned decision.

There has been particular concern about the mental legislation of England and Wales because it is strongly paternalistic in that it allows for the detention and involuntary treatment of capacitous, and therefore autonomous, individuals because clinicians consider it is ‘appropriate’ to do so. The consequence of this is a coercive anomaly in that, if one has full mental capacity, one can refuse any medical treatment; but it is not possible to refuse admission to a psychiatric hospital or even physical/pharmacological treatment. An attempt to introduce a capacity-based mental health act (whereby one would have to show that a person lacked capacity before involuntary detention and/or treatment) (Richardson 1999) failed under the Labour government of 1999 and the current coalition government has done nothing to change the status quo. This suggests a political will to ensure that those considered to be mentally disordered are treated differently from the rest of the society when it comes to autonomy; a sad commentary on Liberal Democrat principles.

Autonomy and best interests

In psychiatry, respect for autonomy may mean that a clinician will allow a patient to make decisions free of coercion, even if this may appear to go against the patient’s best interests – the basis of assessing capacity in statute law. Ethical principles for assessing capacity are now enshrined in law, in the Mental Capacity Act 2005 for England and Wales. They are as follows:

- people are assumed to have capacity to make their own decisions unless otherwise proven
- everything should be done to help a person arrive at a decision before declaring them incapable of making it
- whatever is done for a person lacking capacity should be done in their best interests
- the decision maker should choose the least restrictive intervention.

For instance, where someone is found in a state of semi-consciousness following a large overdose and their life is at stake, the Mental Capacity Act would require that immediate action be taken in the person’s best interests to save life. It would not be practical or reasonable to go on an exhaustive fact-finding mission before initiating treatment. However, beyond the emergency phase, while a person still lacks capacity, the Act requires a much
more concerted assessment of the person’s capacity, the views of their carers or appointed attorneys, formal assessment of their best interests and consideration of any advance decisions. Within this Act, an advance decision is specifically a decision to refuse treatment made in advance and in writing. If it was prepared at a time when the patient had capacity, an advance decision is legally binding unless there are legitimate questions as to its validity or applicability (see the Wooltorton case outlined in Box 1). Further guidance is contained in the Code of Practice accompanying the Mental Capacity Act (Department for Constitutional Affairs 2007).

Challenges in judging capacity

Doctors, by virtue of their professional training, have access to people at their most vulnerable; and duties that entail physical contact that may even cause harm or damage. For this reason, the law on consent is a dominant component of medical jurisprudence. No medical intervention can be imposed on an individual without consent; and to do so would amount to an assault. The legislation on consent reflects the respect of the law for individual autonomy and choice, as indicated in Collins v Wiklox [1984]: ‘The integrity of every person’s body, save by consent, is established in law’.

This principle was made explicit in the case of Re T (Adult: Refusal of Treatment) [1992], where it was specifically emphasised that, ‘prima facie, every adult has the right and capacity to decide whether or not he will accept medical treatment, even if a refusal may risk permanent injury to his health or even premature death’.

It is now established that competent adults can refuse life-saving treatment. This is consistent with the repeal of the Suicide Act 1961, and in light of other legal decisions about the capacity to make unwise decisions. Debate still continues with regard to situations in which competent adults want to end their lives, but need assistance to do so. This debate is typically couched in terms of voluntary euthanasia or the ‘right to die’. In some jurisdictions, such as Oregon, USA, and The Netherlands, it is legally possible for competent adults to be assisted to die by licensed physicians.

Such debates are problematic for psychiatric services, where the wish to die is usually taken to be prima facie evidence of a disordered mind. Indeed, in both the USA and The Netherlands those who wish to end their lives are required to undergo a psychiatric assessment that excludes the possibility of the decision being a ‘symptom’ of any of the diagnoses known to be associated with a risk of suicide. As described earlier, a person who attempts suicide can be detained under the Mental Health Act if they have a mental disorder and if it is appropriate to detain them. The purpose of the detention is to prevent them from killing themselves, even if they possess the capacity to make the decision to do so.

The law currently implies that it is clinically straightforward to distinguish between a wish to die that is the result of a competently made decision that must be respected in law, and a wish to die that is a symptom of a treatable mental illness. In reality, it is not as simple.

Fluctuating capacity and the tripartite model

What is a clinician to do when a patient’s capacity changes from day to day and week to week? Severe disturbances of affect and arousal regulation cause rapid oscillations in a person’s perceptions about themselves, their world and their anticipated future (Sarkar 2006); and equally oscillating decisions about accepting or refusing life-saving care and treatment (Sarkar 2008). These are states of mind in which there is fluctuating capacity, a phenomenon that has been largely ignored in capacity judgments to date (Sarkar 2011c). Despite a slew of legislation that can bewilder clinicians, fluctuating capacity raises the problem that theoretical principles do not necessarily address real, complex clinical issues (Mogg 2005). The bar in assessing capacity must be set high when clinicians make best interests decisions when someone’s life is at stake, as in the case of suicidal patients. This approach was recognised by the Court of Appeal in Re MB (An Adult: Medical Treatment) [1997] when it was stated that ‘the graver the consequences of the decision, the commensurately greater the level of competence required to make that decision’.

Case law on capacity suggests that temporary factors such as confusion, shock, fatigue, pain, drugs or panic can completely erode capacity (Royal College of Psychiatrists 2004). Clinicians will add to this list other factors, such as dissociation and problems of self-identity, rapidly fluctuating mood and behaviours, and concomitant risk of harm to others (Gallagher 2010; Sarkar 2011a). This highlights the real clinical challenge of working out a person’s real intent and choices and decisions when they are in great distress. I argue that it is imperative to consider a third element in capacity decisions (Sarkar 2008). This is the determination of how ‘stable’ or ‘settled’ a decision judged to be capacitous actually is. If it is affected by or a product of underlying mental disorder, it may change with the illness over time. Therefore, a written decision cannot simply be
I propose an ‘enhanced’ approach to the balance model of capacity judgements if there are relatively rapid changes in the person’s mental state, leading to rapid changes in capacity and risks. The ‘enhanced’ tripartite model requires three elements to be assessed: the rationale or logic behind the decision made; the conviction with which the decision is held and is aligned with the patient’s life narrative, beliefs and values; and the clarity with which this decision is communicated to others (Table 1). It is suggested that these three elements, when assessed together, are likely to address the confounding effect that fluctuating capacity has on issues of consent (Mogg 2005), which can often lead to reluctance of clinicians to treat in difficult cases (McCall Smith 2001). By trying to ensure that the decision is stable, enduring and consistent with the patient’s frameworks of value and meaning, this approach provides the clinician with a strategy to limit paternalism as much as possible. The clinician’s tasks therefore include:

- judging whether the decision (e.g. to self-harm, discharge from voluntary admission, go on leave from hospital) is informed and balanced
- judging whether it is communicated clearly and unambiguously
- judging the conviction with which it has been made by assessing its stability and endurance over time: the more important the decision and the more final its consequences, the longer the time frame should be over which its stability is assessed.

### Conclusions

Managing self-harm and risk of suicide are complex tasks. Although recent case law directs the clinician’s mind to assessing risk of suicide and acting according to it in a longitudinal way, current clinical practice often lags behind on account of categorisation of such risk. Risk of self-harm and suicide and, by extension, a patient’s capacity to make rational choices are most often categorised as ‘present’ or ‘absent’. Clinical decisions that flow from such cross-sectional categorical determinations of risk are likely to be truncated judgements on a matter that needs to be judged in a more dynamic and longitudinal manner. The proposed tripartite model can be used to eliminate difficulties associated with fluctuating capacity in patients with complex needs and risks. In practice, it has been found to add a certain degree of rigour to clinical decision-making. Time will tell whether such an approach is widely adopted by clinicians.
Legal and ethical considerations in suicide prevention

References


Bingham J (2009a) Living wills law could be ‘revisited’ after Kerrie Wooltorton suicide. The Telegraph 4 October.


Collins v Wilcox [1984] 1 WLR 1172.

Re T (Adult: Refusal of Treatment) [1992] 4 All ER 649.


Re MB (An Adult: Medical Treatment) [1997] 2 FCR 541.


Savage v South Essex Partnership NHS Foundation Trust [2008] UKHL 74.

MCQs

Select the single best option for each question stem.

1 The balance model of Eastman & Hope states that capacity:
   a is a binary concept
   b decisions are easy to make
   c is either present or absent
   d is graded into degrees
   e need not be always assessed.

2 Fluctuating capacity refers to:
   a capacity that is either present or absent
   b decisions that are rapidly changing
   c decisions made in an intoxicated state
   d the approach recommended in the Mental Capacity Act 2005
   e an issue routinely addressed by clinicians.

3 In the Rabone suicide case, the Court made a landmark judgment in terms of:
   a making hospitals not liable for irresponsible acts of their patients
   b making hospital authorities remove all ligature points in psychiatric wards
   c making consultants legally responsible for deaths of their patients
   d holding the carers of patients legally responsible for their deaths
   e making hospitals legally responsible for preventing suicide where a ‘real and immediate’ risk has been identified.

4 Under the mental health legislation of England and Wales, people with terminal physical illness but no mental illness:
   a are permitted to take their own life regardless of their mental capacity
   b can get help from their carers to assist them in taking their own life
   c can appeal to the European Court of Human Rights if they are prevented from taking their own life
   d cannot be prevented from taking their own life if they are demonstrably capacitous in the opinion of consultant psychiatrists
   e are not required to demonstrate capacity.

5 Suicidal individuals with mental illness who possess capacity:
   a pose significant challenges in terms of society’s duty to save their lives
   b can be prevented under the Mental Capacity Act from taking their own life
   c are always allowed to take their own lives, under mental health legislation of England and Wales
   d should not be detained under the Mental Health Act
   e are not encountered in practice.