

Predicting suicide is a difficult, complex challenge

By Joseph Franklin, *The Conversation*

Who is going to die by suicide? This terrible mystery of human behavior takes on particular poignance in the wake of suicides by high-profile and much-beloved celebrities Kate Spade and Anthony Bourdain. It is only natural that people want to know why such tragedies occur. Those closest to those who take their lives are often tormented, wondering if there is something they could have – or should have – known to prevent their loved one's suicide.

As a scientist who has focused on this question for the past decade, I should have a pretty good idea of who is and isn't going to die by suicide. But the sad truth is, I don't. The sadder truth is, neither do any other suicide experts, psychiatrists or physicians. The sum of the research on suicide shows that it does not matter how long we've known someone or how much we know about them. In my research, my colleagues and I have shown that we can only predict who is going to die by suicide slightly more accurately than random guessing.

The need for answers

The fact that suicide is so hard to predict unfortunately took about 50 years for most scientists to appreciate. About the same time that this recognition became widespread a few years ago, a new hope emerged: a form of artificial intelligence called machine learning. As several research groups have demonstrated in recent years, machine learning may be able to predict who is going to attempt or die by suicide with up to 90 percent accuracy.

To understand why this is, and why we humans won't ever be

able to accurately predict suicide on our own, one needs to take a step back and understand a little more about the nature of human cognition, suicide and machine learning.

As humans, we love explanations that have two qualities. First, explanations should be simple, meaning that they involve one or a small number of things. For example, depression is a simple explanation for suicide.

Second, explanations should be determinate, meaning that there is one set explanation that accounts for all or most of something. For example, the idea that depression causes most suicides is a determinate explanation. This simple and determinate explanatory style is highly intuitive and very efficient. It's great for helping us to survive, procreate, and get through our days.

But this style of thinking is terrible for helping us understand nature. This is because nature is not simple and determinate. In recent decades, scientists have come to recognize that nearly everything – from physics to biology to human behavior – is complex and indeterminate. In other words, a very large number of things combined in a complex way are needed to explain most things, and there's no set recipe for most physical, biological or behavioral phenomena.

I know that this latter idea of indeterminacy is especially counterintuitive, so let me provide a straightforward example of it. The math equation $X + Y = 1$ is indeterminate. As humans, we instinctively try to find one solution to this equation (e.g., $X = 1$, $Y = 0$). But there is no set recipe for solving this equation; there are nearly infinite solutions to this equation. Importantly, however, this does not mean that "anything goes." There are also near infinite values for X and Y that do not solve this equation. This indeterminate middle ground between "one solution" and "anything goes" is difficult for most humans to grasp, but it's how much of nature works.

The sum of our scientific evidence indicates that, just like most other things in nature, the causes and predictors of suicide are complex and indeterminate. Hundreds, and maybe thousands, of things are relevant to suicide, but nothing predicts suicide much more accurately than random guessing. For example, depression is often considered to be an extremely important predictor of suicide. But about 2 percent of severely depressed people eventually die by suicide, which is only slightly higher than the 1.6 percent of people from the general United States population who eventually die by suicide. Such a pattern is consistent with complexity because it suggests that we must put a lot of factors together to account for suicide.

Empathy will always matter

So how should we put all of these factors together? One intuitive solution is to add many of these factors together. But even when summing hundreds of factors, this doesn't work – prediction is still only slightly more accurate than random guessing.

A much better solution would be to somehow find an optimized combination of tens or even hundreds of factors. How can we do this? One promising answer is machine learning. In short, machine learning programs can process a large amount of data and learn an optimal combination of factors for a given task. For example, most existing machine learning studies have used data from electronic health records, spanning hundreds of factors related to mental health diagnoses, physical health problems, medications, demographics and hospital visit patterns. Results from several groups in recent years have shown that this approach can consistently predict future suicide attempts and death with 80-90 percent accuracy. Multiple groups are currently working on applying these algorithms to actual clinical practice.

One important thing to keep in mind is that there isn't, and

never will be, a single algorithm or recipe for suicide prediction. This is because suicide is indeterminate, much like the $X + Y = 1$ equation. There are likely near-infinite algorithms that could predict suicide with 80-90 percent accuracy, as a number of studies have shown. Research has already demonstrated that no particular factors are necessary for a good algorithm, and many different types of algorithms can produce accurate prediction. But again, this indeterminacy also means that there are near-infinite bad algorithms, too.

All of this research shows that suicide is unfortunately too complex and indeterminate for humans to predict. Neither I nor anyone else can accurately predict who is going to die by suicide or truly explain why a particular person died by suicide (this includes the suicide decedents themselves). Machine learning can do a much better job of approximating the complexity of suicide, but even it falls far short. Although it can accurately predict who will eventually die by suicide, it cannot yet tell us when someone will die by suicide. This "when" dimension of prediction is critical, and we are likely still many years away from accounting for it.

In the meantime, what can we humans do? While we don't have the ability to know whether someone is going to die by suicide or not, we do have the ability to be supportive and caring. If you believe that someone may be struggling, talk with them and let them know about resources such as the National Suicide Prevention Lifeline (800.273.8255).

Joseph Franklin is an assistant professor of psychology, Florida State University.