'You may be hacked' and other things doctors should tell you

November 3 2020, by Maximilian Kiener

Credit: Karolina Grabowska from Pexels

On September 9 2020, a woman died during a cyber-attack on a hospital in Düsseldorf, Germany. The woman was in a critical condition and about to be treated when hackers disabled the computer systems of the

hospital. Unable to avert the attack, medical staff had to transfer the woman to another hospital, but the help came too late and the woman died.

This incident was the first reported case of death after a cyber-attack and shows that such attacks are not just a threat to our data anymore, but also to our lives. In fact, the situation is urgent. We know that cyber-attacks on medical devices and hospital networks are a growing threat. During the current pandemic, some types of cyber-attacks have increased by 600%.

And it's not just old computer systems that are vulnerable. Even the very best artificial intelligence (AI) in medicine can be compromised. Academic research continually reveals new ways in which state-of-the-art AI can be attacked. Such attacks can block life-saving interventions, undermine diagnostic accuracy, administer lethal drug doses, or sabotage critical moves in an operation.

Inherent risks

Doctors need to do everything they can to keep patients safe, but as a matter of general medical disclosure, should they have to tell patients about the risk of a cyber-attack, at least when their healthcare critically relies on computers? After all, patients have to give their informed consent to medical procedures and doctors are required to warn patients about potentially harmful consequences.

In some US legal cases, judges have argued that doctors need to disclose a risk only if it is "inherent" in a medical procedure, that is, a risk that "exists in and is inseparable from the procedure itself". Relying on such a view, one may argue that the risk of cyber-attacks is not an "inherent" risk and so does not require disclosure. Many equate "inherent" risks with "medical" risks and thereby rule out the "criminal" risk of a cyber-

attack.

This view against disclosure raises an important point. There is indeed a connection between the requirement of disclosure and the expertise of a doctor as a medical professional. Doctors need to disclose inherent medical risks because they are, unlike laypeople, especially well placed to know about them. But doctors cannot be expected to predict whether certain people will target their patients through cyber-attacks. After all, doctors are not criminologists. So they are not really able, let alone obliged, to disclose those risks.

On the other hand, this view against disclosure underestimates several important aspects. To begin with, the growing digitalisation and use of computer systems in medicine will render the risk of cyber-attacks ubiquitous in healthcare. Even though it may not be an "inherent" risk, it will certainly be an inevitable part of future clinical reality, and if we want patients to make well-informed decisions, they should know about such a risk.

Also, even though doctors don't need to disclose general criminal risks, they are required to disclose the risks that their medical equipment poses to patients. After all, being subject to medical procedures leaves people vulnerable in important ways, and if certain computer-based procedures introduce new vulnerabilities, an informed patient will need to know about them.

Finally, unlike traditional cyber-attacks, the risk of some new cyber-attacks may become "inherent," as defined above. Consider the case of medical AI. In so-called "input attacks" on medical AI, an attacker can change the pixel value of an MRI scan so that the AI system will categorize tissue as falsely malignant with a confidence rate of over 99% when it would correctly categorize it as benign with the same confidence rate in absence of the attack. The human eye is unable to detect such

changes. The attacker would only have to scatter some well-placed digital dust over the image.

The only way to detect an attack is to detect the intrusion in another computer system where the medical images have been stored. But even here, we may not know whether, in addition to the intrusion into the database and the potential theft of medical data, attackers made any changes to medical images at all, what their motives might have been, and what consequences could await patients as a result.

So, unlike other cyber-attacks, input attacks no longer compromise their target system. The AI system itself, its algorithm, and how it works can be left completely untouched. In other words, the AI system would still work normally, not be affected by any bug or interference, and the doctor performing or supervising the procedure would act as professionally as possible.

Therefore, no such AI-based procedures can avoid the vulnerability to input attacks. But if this is so, then the risk of input attacks does become "inherent" to certain medical procedures, as defined earlier.

Be prepared

There are sound reasons to require the disclosure of cyber-risks to patients, at least in certain high-stakes medical procedures. However, cyber-risks are only one new type of risk that patients may face in the future. When algorithms play an increasingly large role, we also need to think about whether doctors should disclose the risk that these algorithms are systematically biased or the risk that, because of the opacity of certain AI systems, doctors may no longer be able to understand and double-check the AI's decisions.

In any case, the growing reliance on computer-systems and AI demands

that we think afresh about medical disclosure and which risks to disclose to patients. Otherwise, our clinical practice will be unprepared for the major transformations that await it.

Provided by The Conversation