The Hidden Harms of Crying Wolf

By Chuck Dinerstein — January 17, 2019

Courtesy of wikipedia [1]

Snap Judgment, a podcast that features stories that, literally, have a beat, reminded me that this week marks the anniversary of perhaps the greatest false-positive alarm our country may have ever experienced: the notification given to Hawaii of the immediate impact of a ballistic missile strike.

The podcast did a far better job than I'm currently doing here in conveying the devastating emotional effect on individuals, as they waited for the end of times. It got me to thinking about how I might have responded. I have had more experience with false positives than most Hawaiians since physicians frequently have to deal with falsely positive information.

The ascendency of algorithmic medicine, now cloaked in mysterious Deep Artificial Intelligence (AI), has not been allowed to act autonomously. It refers patients to a doctor who over-reads the report, who determines whether their judgment is the same as that of the computer acting upon the information. But these systems, like their human counterparts, can be wrong.

From the point of view of safety, it's better to be falsely positive and do more diagnostic testing or interventions than to be falsely negative and let a problem, like cancer, fester. In malpractice litigation involving false-negative mammography, a suspicious lesion is seen more than 25% of the time on the test before the "false" negative test is the subject of the litigation. So while we are fallible, having to look at a test result that looks fine to you but that the algorithm says is suspicious is uncharted territory. What to do? Use your judgment or that of a computer that sees some indiscernible "pattern?"

There is another far more frequent source of false positives that all clinicians encounter daily. They are those alarms that we have installed on our intravenous pumps, telemetry monitors and

electronic medical records. Anyone who has walked the hallways of a busy "med-surg" floor recognizes the various chirps and buzzes; it's a cacophony of confusion. Those so-called "cries of wolf" have become so bad that they've engendered a new term – alarm fatigue.

One study [2] found that each patient in an intensive care unit generated 197 alarms a day. As of 2016, the Joint Commission, the NGO accrediting hospitals, have made dealing with alarm fatigue one of its national patient safety goals. But from the sounds of it (pun intended), officials there have made little progress.

With the introduction of computerization of physician orders alarms were again a problem. While only a tiny percentage of orders might generate an alert, every patient requires lots and lots of orders, so alerts and notifications quickly amplified. In many instances we continue to respond to the same "concern," like allergies we answered two alerts previously.

The predictable results are that most of these alerts are over-ridden, and you could say that from the design point of view: no harm, no foul. But the silent harm is the way that these intrusions distract us. And distracted care, like distracted driving, is unsafe. We've become increasingly insensitive to them, and in the cognitive battle to separate the real from unreal we turn the whole system off. Therefore, paradoxically, safety alarms can make us less safe.

At its root are the good intentions of system designers, who "out of an abundance of caution" have made the alarms very sensitive. But at the same time, the system is not too specific, which then requires a human to come, look, and make a judgment – just as the designers of AI are asking us to do again.

Algorithms are behind the alarms of telemetry, devices and electronic health records. They have an established track record; providing some marginal improvement for patients, but taking a big toll on the clinicians required to provide oversight and act.

So perhaps, before this next algorithmic wave of AI diagnostics, the designers and engineers can turn their attention to addressing this "safety paradox." Before we are presented with more deep AI diagnostics, we need some deep AI to look at how to reduce the "cries of wolf." That's not as shiny and monetizable, but so much more necessary.

Take a few moments to listen to that *Snap Judgment* podcast, and consider how it plays out every day for patients and clinicians, and the toll it exacts on our emotional well-being.

###

Source: Snap Judgment [3], produced by WNYC.

---

[3] https://www.wnycstudios.org/story/snap-1001-not-drill