Is the 510(k) Pathway Robust Enough for the Era of AI?

Hugh Harvey, MBBS BSc (Hons), FRCR, MD (Res), FBIR

The US Food and Drug Administration's (FDA) 510(k) regulatory pathway allows medical device manufacturers to identify and leverage an already marketed device to apply for premarket clearance of their own substantially equivalent device. The FDA's logic is that two devices with the same intended use can be similar enough to either share the same technological characteristics or not raise any new safety or effectiveness concerns.[1]

An obvious problem of this pathway is grandfathering. Not only does grandfathering lead to a nearly endless line of devices all claiming to be identical (but "novel" to their targeted market), but it potentially creates a failure waterfall when a predicate device is subsequently recalled.

One would think that if a predicate is withdrawn from the market under FDA orders, all other substantially equivalent devices must also be withdrawn. Not so. The FDA treats every company's new 510(k) cleared device individually once it hits the market, even though it may have received marketing clearance under the very same parameters of the device's named predicates.

Affiliation: Dr Harvey is managing director of Hardian Health, based in Haywards Heath, England.
Disclosures: None.

Published literature on recalled pelvic mesh devices readily demonstrates issues created by the 510(k) pathway,[2-4] which speeds time-to-market while treating real patients as guinea pigs, rather than mandating rigorous premarketing clinical investigations.

When it comes to artificial intelligence (AI) imaging products, the FDA 510(k) process risks allowing technology on the market without full consideration of subtle differences between the products. AI relies on bespoke programming and "black box" machine learning to perform specific actions. In the case of medical imaging device software, these can range from triage to diagnosis. Computer programming comes in many languages and styles, with developers each having their own preferences and methodologies. No two developers will ever produce the same lines of code, let alone two identical AI models.

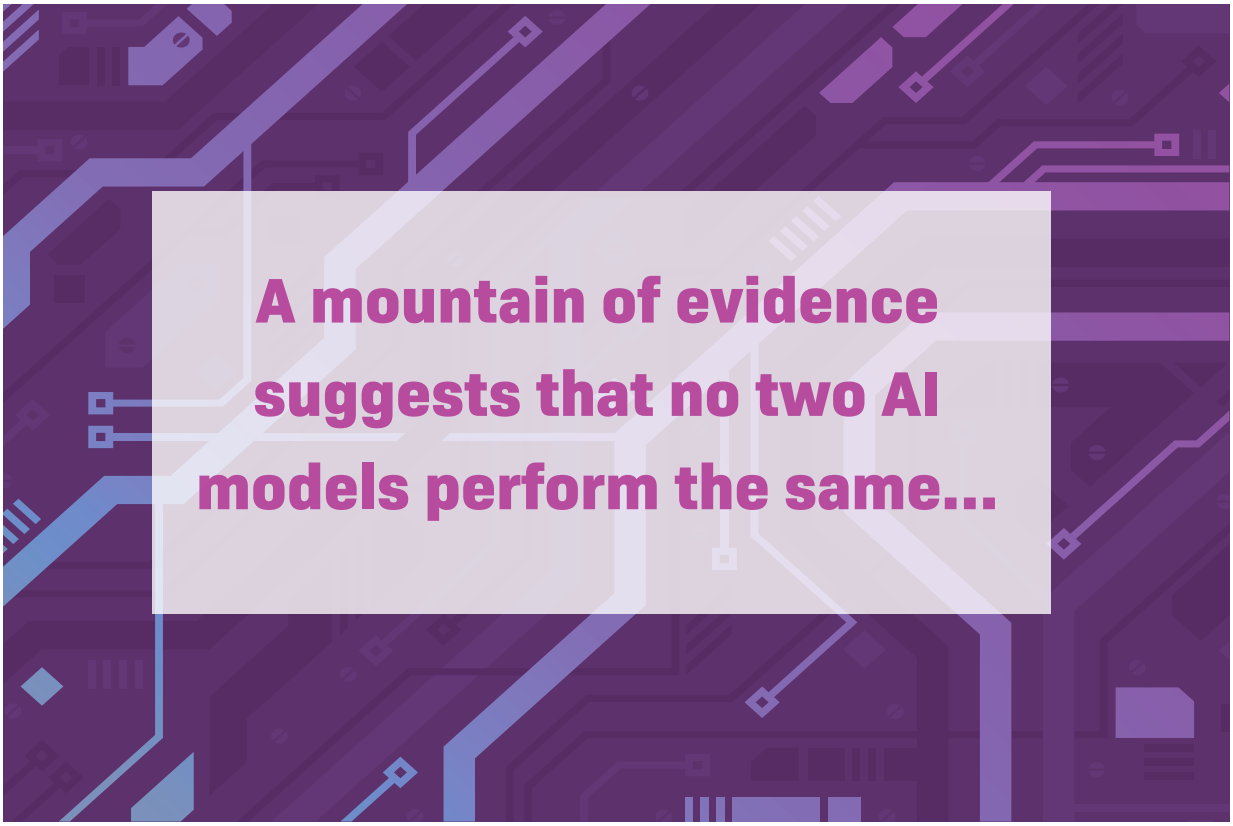
Since AI relies on labeled data inputs to model a latent space of knowledge, it is also obvious that no two AI systems trained on different input data sets could ever be identical. It is also inconceivably unlikely that any two models would have the same construct, layers, weights, and hyperparameters, let alone the same level of performance. Even the FDA's own guidance defines substantial equivalence as "no significant change in the … design or other features of the device from those of the predicate device," which clearly is not the case with AI systems.

Since substantial equivalence of AI systems cannot be reliably proven through technological characteristics, then what about safety and effectiveness? A mountain of evidence suggests that no two AI models perform the same,[5] individual models perform differently in different locations,[6] and are as impacted by performance drift and variability once deployed over diverse populations.[7] Therefore, effectiveness is also not a reliable indicator of equivalence.

Demonstrating equivalence then hinges on safety requirements; eg, do two models trained on different data sets have the same sub-stratification and hidden biases, known to cause safety issues? Almost certainly not.

Equally, if not more, pertinent, the device's intended use must also be the *same* as the predicate device in order for the company to claim equivalence. How then can one AI system intended to triage chest X-rays for pneumothorax be deemed the *same* as one intended for triaging CT head studies for brain aneurysms?[5] The FDA lumps both systems into one product code simply

©Anderson Publishing, Ltd. All rights reserved. Reproduction in whole or part without express written permission is strictly prohibited.



because they both triage, ignoring the significant clinical differences in risk between the use cases. Ultimately, this raises the question: is the 510(k) pathway sufficiently robust for this era of AI?

Neither European Union or United Kingdom medical device regulations rely on substantial equivalence as a factor in rendering marketing approval. Each device is audited on its own premarketing evidence plus robust post-market follow-up studies. While US manufacturers are likely to balk at the idea of providing additional evidence, that is simply the reality for device manufacturers here across the pond.

Perhaps the FDA eventually will come around, but as the saying goes, "regulations are written in blood."

Let's hope it doesn't come to that.

References

1) FDA.gov. The 510(k) program: evaluating substantial equivalence in premarket notifications. 2023. https://www.fda.gov/media/82395/download. Accessed Feb. 6, 2023.

2) Jacoby VL, et al. The FDA and the vaginal mesh controversy—further impetus to change the 510(K) pathway for medical device approval. *JAMA Intern Med.* 2016; 176(2): 277. doi:10.1001/jamainternmed.2015.7155. Accessed 6 Feb 2023.

3) Rosh J, et al. The 510(K) ancestry of transvaginal mesh. *JAMA Surgery*. 2021;156(8):701-702. doi: 10.1001/jamasurg.2021.0606. Accessed February 6, 2023.

4) Heneghan, C J, et al. Trials of transvaginal mesh devices for pelvic organ prolapse: a systematic database review of the US FDA approval process. *BMJ Open*. 2017; 7(12): e017125. doi:10.1136/bmjopen-2017-017125. Accessed February 6, 2023.

5) Zheng Q, et al. Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: a systematic review and meta-analysis. *Eclinical medicine*. 2021; 31:. 100669. Elsevier BV, doi:10.1016/j.eclinm.2020.100669. Accessed 6 Feb 2023.

6) Ehteshami BB, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017; 318(22): 2199. doi:10.1001/jama.2017.14585. Accessed February 6, 2023.

7) Rahmani K, et al. Assessing the effects of data drift on the performance of machine learning models used in clinical sepsis prediction. Pre-publication November 19, 2022. *International J of Med Informatics*. doi:10.1101/2022.06.06.22276062. Accessed 6 Feb 2023.

8) FDA Product Code QFM. FDA Report 2023. https://fda.report/Product-Code/QFM. Accessed February 6, 2023.