

# Medical devices must be carefully validated

The safety and performance of medical devices should be validated in the conditions and the environment that would most benefit patients.

When medical technologies are used to diagnose, monitor, treat or alleviate a medical condition, they can increase the longevity or quality of life of the patient, and also ease pressure on the healthcare system. For example, phenotyping tissue is the predominant method for diagnosing cancers, yet DNA microarrays are increasingly seen as valuable hospital equipment for such diagnostic tests. These systems can analyse the expression levels of thousands of genes simultaneously, improving diagnostic precision and speeding up laboratory turnaround times. For new technology, to ensure that it fulfils the requirements for the intended use and user, the validation criteria that must be met on the path to clinical implementation and before potential market approval have to be rigorously designed. To justify the utility of any biomedical innovation, it should be demonstrated to be technically superior to an appropriate gold standard, or at least non-inferior to the gold standard yet more cost-effective or efficient. Most importantly, a new technology should be proven to be safe and produce consistent results. Strong evidence of validation minimizes the risk to patients and provides reasonable safeguards against failure in clinical practice.

Ideally, validation of a device in a clinical trial that compares the performance of the device against the standard medical procedure in clinical practice should provide the strongest evidence for the device's safety and efficacy. Although a randomized controlled trial (RCT) is widely accepted as the gold standard for validating pharmaceutical products and advanced therapies, the appropriateness of this pathway is less obvious for medical devices. For example, because the effectiveness of devices often depends on the user's skills and knowledge, RCTs may not necessarily take into account variations in the skills of the surgeon (for an intraoperative or implanted device) or the patient (for a worn device, for instance). Inadequate device efficacy could thus be a consequence of ineffective use. Also, although in a trial of an implanted device sham surgery might appear to be an obvious control group, it can be unethical to offer sham intervention, as the risks to the patient are typically less justifiable than in trials of drugs; instead, when possible, the control arm should be

selected on the basis of existing standards of care. Randomization and blinding, which are designed to minimize bias, can also be harder or impossible to implement in medical-device trials.

Four Articles in this issue navigated these long-standing validation challenges to provide preliminary evidence of safety or efficacy for new devices. [Robert MacLaren and colleagues](#) conducted a first-in-man study of robotic-assisted intraocular surgery. The researchers demonstrate the feasibility and safety of the robotic device for the peeling of retinal membranes and for the injection of a therapeutic under the retina. In this trial, patients were randomized to receive either surgery from the robotic-assisted surgeon or manual surgery (the standard of care). The degree of precision and accuracy executed by the clinician-guided robot exceeded human limits, although the robotic surgery was slower than manual surgery. Safety was enhanced, as evidenced by the fewer retinal touches and consequent microhaemorrhages reported (these surrogates of risk might, however, not be sufficient for the assessment of safety in larger clinical trials of the robotic device, as noted by [Peter Gehlbach](#) in an associated News & Views article). As it was impossible for the surgeons to be blinded to the procedure, blinding the assessment of the outcome (which was carried out by different investigators to those doing the surgery) was considered the best alternative.

In a global-health context, technology can be most effective when employed at the point of care (POC). For example, although the best method for detecting lymphoma is a conventional pathology work-up obtained from core biopsies, this is not always possible in low-resource settings either because of the limited number of pathologists or the lack of access to refrigeration for reagent storage or of electricity to run the equipment. [Ralph Weissleder and colleagues](#) now report a low-cost POC device that uses contrast-enhanced microholography and deep learning to accurately detect aggressive lymphomas in a prospective clinical trial of patients referred for aspiration and biopsy of enlarged lymph nodes. The technology is designed to address the limited pathology resources in many low- and middle-income countries by enabling the diagnosis to



A device for nucleic acid amplification being powered with sunlight in a Ugandan clinic. Figure reproduced from the Article by [Erickson and colleagues](#), Springer Nature Ltd.

be obtained from a fine-needle aspirate analysed at the same location it was collected. Despite the fact that the device was not tested in the target population, proof of principle of its feasibility as a POC technology for lymphoma diagnosis was shown in 40 patients referred for suspected lymphoma, and compared against clinical pathology and cytology data. The integrated cassette-based device incorporates lyophilized antibodies that showed effectiveness under different storage conditions, which was used as a surrogate test for environments absent of cold-chain resources. The researchers also show that the use of chromogens to detect intracellular markers is as accurate as flow cytometry, thus enabling the visualization of the test results without the need for expensive microscopy.

To establish wider utility at the POC, device validation should be carried out within the target population, and in the most appropriate environment and use conditions. In this respect, [David Erickson and colleagues](#) deployed, in settings with limited access to stable electricity, a portable device (pictured) running a nucleic acid amplification test that uses a phase-change material to store energy from diverse sources — such as sunlight, electricity or an open flame — and to maintain the constant temperature needed for the test to diagnose a viral infection. The researchers validated the device in two Ugandan health clinics,

and report that it achieved comparable performance to a commercial quantitative polymerase chain reaction (PCR) machine for patients suspected of Kaposi's sarcoma (a rare cancer that characteristically affects the skin and the mouth). Compared with battery-powered machines for nucleic acid amplification, the device is an order of magnitude smaller. The miniaturization was made possible by the use of the phase-change material instead of the typically bulky thermocycler, and by amplification chemistry that obviates the need for thermal cycling.

Miniaturization also makes possible new types of medical device. By incorporating a piezoelectric composite in a soft silicone

elastomer, [Sheng Xu and colleagues](#) developed a conformal ultrasonic device in the shape of a flexible skin patch that can accurately monitor both central blood pressure (that is, the pressure in the aorta) and peripheral blood pressure from a variety of anatomical locations. The researchers also provide a first validation of the accuracy of the wearable patch when located above blood vessels in the foot, neck, arm and wrist of a single healthy human subject, by comparing its performance against non-invasive tonometry (the true gold standard for central-pressure waveforms involves instead an invasive catheter). The patch led to lower measurement uncertainty, higher precision and greater accuracy. Unlike

tonometry, which requires the operator to hold on tightly to the device, the patch enables continuous long-term monitoring of blood pressure, even during motion.

The devices discussed here will all require further validation in either larger patient cohorts or with other relevant comparators. Yet a first validation step for feasibility and improved performance, ideally against the clinical standard, is needed for new technology to be credible. How to best select performance (surrogate) markers and control arms is, however, not always straightforward. □

Published online: 11 September 2018  
<https://doi.org/10.1038/s41551-018-0302-2>