

ALTERNATIVE APPROACH FOR DEVICE SUBSTITUTABILITY: HUMAN FACTORS STUDIES FOR GENERIC AND BIOSIMILAR PRODUCTS

Carrie O'Donel, Henri Akouka and Mark DeStefano of Teva Pharmaceuticals, along with Leslie Sanchez-Torres, discuss the shortcomings of current comparative use approaches to determining substitutability of devices for generics and biosimilars and propose an alternative approach to human factors studies based on long-standing methodologies that has the potential to unlock innovation and improve the user experience for patients.

"IT SEEMS COUNTERPRODUCTIVE TO REPLICATE A DESIGN THAT DOES NOT REFLECT CURRENT BEST PRACTICES, ESPECIALLY WHEN THERE ARE OPPORTUNITIES TO IMPROVE THE USER EXPERIENCE AND REDUCE RISKS WITHOUT CREATING AN INCREASE IN USE ERRORS."

As developers of generic and biosimilar combination products, Teva frequently encounters questions around device substitutability – specifically, whether to replicate the reference device exactly or take the opportunity to improve upon it. These discussions often centre on the trade-offs between strict replication and thoughtful innovation, especially when the original device may not be optimal for patient use.

Simply copying a reference device – regardless of its complexity or usability – can result in products that are more difficult to manufacture, less reliable and prone to the same user errors already identified in the original design. This approach can unintentionally carry forward known issues, rather than address them.

It seems counterproductive to replicate a design that does not reflect current best practices, especially when there are opportunities to improve the user experience and reduce risks without creating an increase in use errors. The idea that patients cannot adapt to improved device interfaces is outdated. In fact, this assumption can slow down the delivery of more affordable medicines to patients who need them.

That is why exploring a shift away from the current comparative use human factors (CUHF) model is warranted. Instead, this article proposes using more traditional human factors methods, such as early usability testing and summative validation principles, to demonstrate substitutability. To that end, the authors embarked on a device usability study designed using ANSI/AAMI HE75, Faulkner and US FDA guidance principles to show that device substitutability across various device types was possible without an increase in user

errors. This approach better supports innovation, patient safety and faster access to high-quality, cost-effective therapies.

STREAMLINING HUMAN FACTORS FOR DEVICE SUBSTITUTABILITY

The FDA's recommended CUHF study methodology guidance for demonstrating device substitutability in generic and biosimilar combination products brings notable challenges to device development efforts. The study method complexity, inconsistencies and time-intensive nature often overshadow the benefits it aims to deliver. Considering current human factors practices, this approach may not be the optimal path for proving therapeutic equivalence.

The fundamental goal of substitutability is clear – to enable patients to use alternative devices for generic medications without an unacceptable increase in use errors, thereby expanding access to affordable treatments. However, duplicating device design and functionality features – a strategy often chosen for ease of FDA approval – can delay patient access and stifle innovation and improved standards of care.

CUHF studies rely on "clinical non-inferiority" methods to demonstrate sameness, but it is worth asking – can more traditional human factors methodologies deliver the same results with less complexity, faster timelines and more opportunities to optimise usability and safety? The study outlined here dives deeper into this question, presenting real-world insights from actual device testing, under real-world conditions, using alternative yet well-established methodologies.

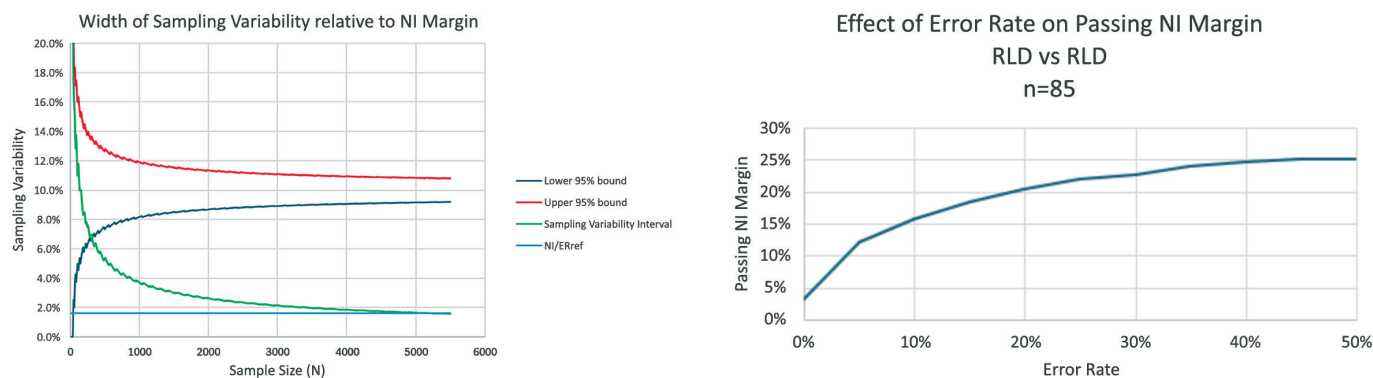


Figure 1: Contribution of sample size to sampling variability.

Developers of 505(j) generics and 351(k) biosimilars have a shared mission – to provide patients with affordable treatments quickly. To achieve this, human factors assessments are essential – not just for compliance, but for minimising use errors, reducing user risk and improving the overall user experience where possible, while aligning with established standards, such as IEC 62366-1:2015/AMD1:2020.

Abbreviated approval pathways created under the Hatch-Waxman amendments prioritise therapeutic equivalence, requiring generic drugs to match the clinical effect and safety profile of their reference-listed drug (RLD). Similarly, biosimilars seeking interchangeability must demonstrate that they are “highly similar” to their reference drug product without compromising safety or usability. Human factors play a critical role in ensuring that the substitution of these products does not lead to unacceptable use errors or additional training or intervention requirements – a pivotal concern for both regulators and developers.

RETHINKING COMPARATIVE USE GUIDANCE

The goal of developing a substitutable drug-device combination product (DDCP) – and the FDA’s evaluation of such – is not “sameness” of user interface, rather, it is safe and effective use without healthcare provider intervention or patient/user training. While

the comparative use guidance emphasises user interface “sameness” to minimise critical risks, this approach can pose barriers to innovation, propagate antiquated and problematic device designs and limit usability enhancements that are beneficial to patients.

CUHF study designs employ a clinical non-inferiority model, which is better suited for objective measures and placebo controls. In contrast, human factors and the sources of use errors involve subjective complexities that are difficult to address through statistical comparisons alone. For example, negative transfer – where prior mental models interfere with understanding new devices – can introduce use errors that are not detectable through traditional objective methods.

By applying rigid clinical frameworks to human factors validation, there is a risk of complicating approval processes for safe, effective products and delaying patient access to affordable treatments. For example, a human factors study requires 85 users to pass a non-inferiority (NI) margin of 16% at 95% power – a large sample size. Sampling variability often dominates the NI margin, leading to a high false negative rate. Adapting continuous data methods to small sample sizes requires a larger NI margin or more participants, as shown in Figure 1, which illustrates the relationship of sampling variability to the NI margin, as well as the ratio of the NI margin to the reference error rate (NI/ERref).

Alternative methodologies, such as the human factors validation study presented here, offer a promising solution. Grounded in human factors best practices such as IEC 62366-1 and ISO 14971, this approach emphasises use specifications for devices targeting users familiar with reference product interfaces. It includes use-related risk assessments (URRAs) to identify tasks prone to negative transfer and evaluates resulting use errors without relying on placebo-controlled statistical analyses. This streamlined design aligns with FDA human factors expertise while addressing mental model challenges in users experienced with the reference product (RP).

This alternative study design also considers how use errors evolve over time. Errors not caused by negative transfer are often resolved through repeated use, though the tolerance for learning curves depends on the treatment’s safety profile. Emergency-use products demand greater scrutiny and lower tolerance for learning through experience, while chronic treatment products may allow for some user adaptation. By leveraging expanded risk mitigation strategies that go beyond minimising design differences, this innovative human factors validation framework paves the way for safer, more accessible combination products, empowering developers, regulators and patients alike.

INNOVATIVE STUDY DESIGN FOR SUBSTITUTABLE DDCPs

The proposed human factors study design for substitutable DDCPs aligns with the framework outlined in the FDA’s Complete Submission Guidance. This study focused on participants experienced

“HUMAN FACTORS AND THE SOURCES OF USE ERRORS INVOLVE SUBJECTIVE COMPLEXITIES THAT ARE DIFFICULT TO ADDRESS THROUGH STATISTICAL COMPARISONS ALONE.”

with four-step autoinjectors, three-step autoinjectors or prefilled syringes (PFSs) for various conditions, assessing their ability to complete a dose using a two-step autoinjector, as shown in Table 1, without prior training or intervention during first-time use and simulated realistic user scenarios to validate DDCP usability.

A key feature of the study scenario involved representing the real-world situation where a pharmacist would substitute the RP for the DDCP without the knowledge of the user. Unlike the CUHF study model, which primarily compares the RP and DDCP in controlled settings, this study design emphasises patient-centric usability. Study recruitment criteria mirrored the intended RP product use specification, ensuring that participants had established prior experience with the RP interface. To replicate home-use environments, the study setting was designed for familiarity and comfort, reflecting standard human factors assessment practices.

PROACTIVE RISK ASSESSMENT AND TASK ANALYSIS

A risk evaluation was conducted on the device using use failure mode and effects analysis (uFMEA) and threshold analysis (TA) methodologies. TA identified “other design differences” in task analysis and physical comparison, underscoring the practical application of human factors validation as per comparative use guidance. A labelling comparison was not conducted since the focus of these analyses was on the functional usability aspects of the device user interface.

The URRA pinpointed critical tasks prone to errors stemming from design differences. The URRA should consider all conceivable use errors and mitigate them as much as possible. However, this is also accomplished through improving the user interface of the substitutable DDCP, such as by reducing the number of critical tasks, as was done in this study. The critical tasks were defined for this study in Table 2, each representing actions essential to dose completion.

These tasks were evaluated for potential use errors caused by negative transfer. Additionally, all user challenges observed






Description	Physical Presentation*	Example
3-step Autoinjector	Rounded body, button on back	
4-Step Autoinjector 1	Rounded body, button on back, locking mechanism, twist-off cap	
4-Step Autoinjector 2	Rounded body, two caps to remove prior to activation, button on back	
PFS with NSS	1 mL long syringe, passive NSS	
2-step Autoinjector (proposed substitutable DDCP)	Rounded body with squared edges, single cap	

Table 1: Physical attributes of study reference products (*attributes that would typically be considered as other differences).

Step	Evaluation Type	Success Criteria
Prepare/visually inspect the autoinjector for physical damages, medication colour/quality	Observation of performance	Inspects the device contents for quality
Inject/uncap the autoinjector	Observation of performance	Pulls the cap off the device
Inject/place the autoinjector against the skin	Observation of performance	Places the device against the skin at a 90° angle
Inject/press and hold down the autoinjector against the skin until a click is heard	Observation of performance	Firmly presses the device down against the skin
	Root cause investigation/post-test interview	Reports: Hearing a click, feeling device actuation, seeing the plunger start movement, indication that the injection has begun
Inject/continue to hold down the autoinjector until a second click is heard and a blue indicator is seen in the fill window	Observation of performance	Delivers a full dose
	Root cause investigation/post-test interview	Reports: Hearing a second click, feeling end of injection, seeing the blue indicator in the viewing window. Or waiting 15 sec after start of injection, indication that the injection is complete.

Table 2: Two-step autoinjector critical tasks with success criteria.

$$(Design\ Verification)\ n = \frac{\ln(1 - C)}{\ln(R)} = \frac{\ln(1 - 0.95)}{\ln(0.95)} = 59$$

$$(Human\ Factors\ Validation)\ n = \frac{\ln(1 - C)}{\ln(R)} = \frac{\ln(1 - 0.95)}{\ln(0.86)} = 20$$

Figure 2: Attribute design verification equations.

during the simulated use scenario were analysed to identify root causes, regardless of their classification as critical tasks.

Determining an appropriate sample size is critical to ensure robust human factors validation. Attribute design verification equations and methodologies, such as those outlined by Faulkner, guided the sample size selection (Figure 2). The sample size

for this study was selected to optimise for observation of use errors based on the risk profile of the product under evaluation. For example, a sample size of 59 with zero observed failures ensures 95% reliability with 95% confidence. Alternative reliability targets were recommended based on each product’s benefit-risk profile, with examples provided in Table 3.

For maintenance treatments, a success rate of 82–86% can meet benefit-risk thresholds even with an observed error rate of 14–18% (95% confidence level) using 15–20 subjects successfully completing the dose. Here, understanding whether primary endpoint failures are attributable to negative transfer is essential, and the subjective nature of use errors and varied root causes underline the importance of targeted risk mitigation.

This study used a sample size of 20 users per group, aligning with industry standards, to scale human factors validation proportional to the product’s safety profile. The goal was to evaluate use errors until an increased sample size yielded diminishing returns, ensuring validation remained practical and efficient.

STUDY RESULTS ON
DEVICE SUBSTITUTABILITY

The findings from this four-part study revealed promising insights into device substitutability. Experienced users of all RP devices demonstrated a successful transition to alternative devices incorporating “other design differences” with minimal use errors and no need for prior training or intervention, when the user’s RP was substituted with an alternate device by the pharmacist. These results highlight the feasibility of designing devices that


	Design Verification		Design Validation	
Primary Endpoint	Confidence/Reliability	Attribute Sample Size	Error rate sensitivity	HF Sample Size for primary endpoint
Complete Dose (maintenance/preventive use)	95%/95%	59	14-18%	15-20
Needle Safety Activation and Override	95%/99%	299	6-14%	20-50

Table 3: Sample size examples for objective and subjective measures.

PHARMA'S GO-TO SOURCE

FOR DRUG DELIVERY INDUSTRY INFORMATION & INTELLIGENCE

To subscribe is simple, just scan the QR code



DrugDelivery • Issue 178

31

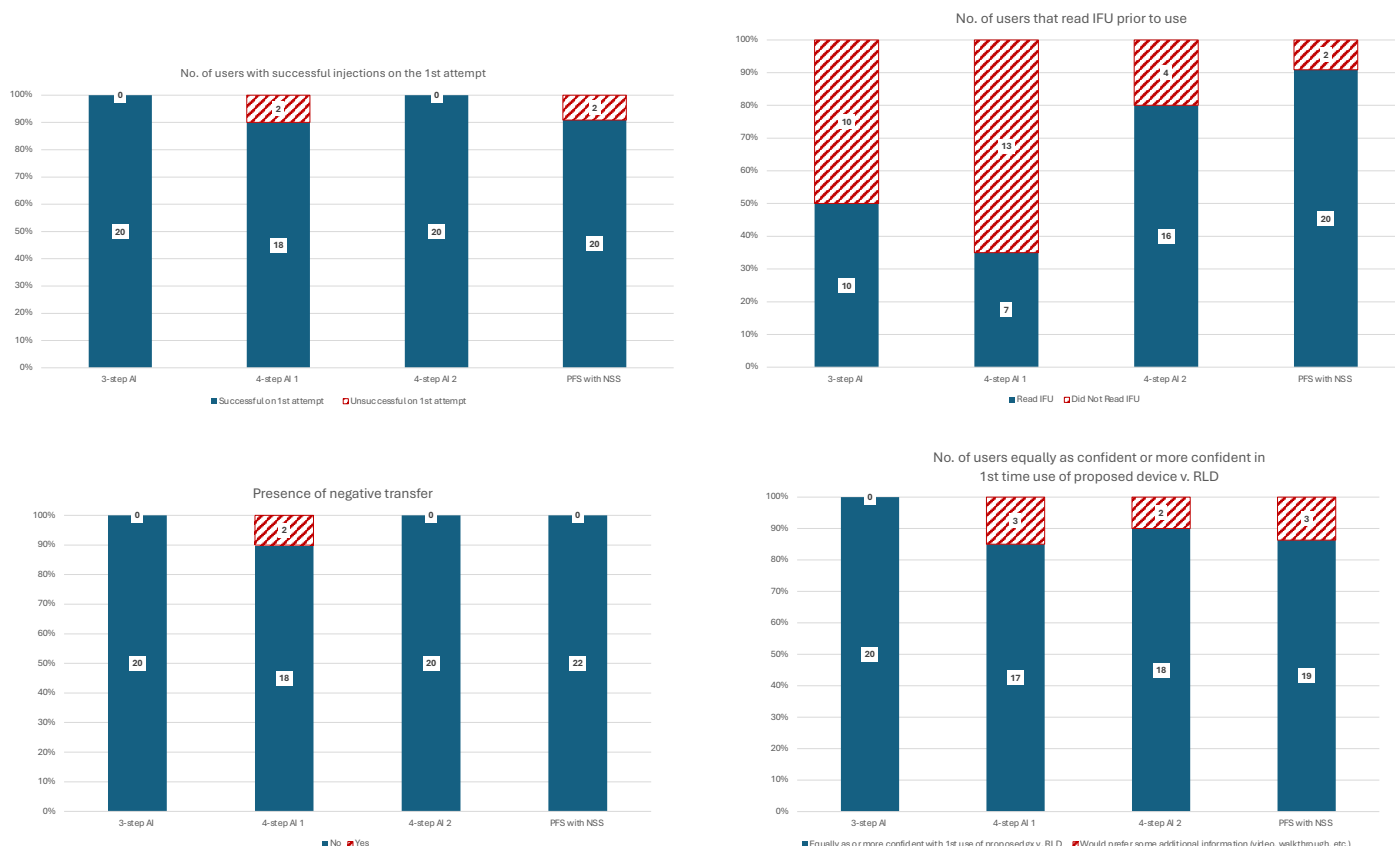


Figure 3: Summary results.

maintain usability while offering design flexibility. The data summary is illustrated in Figure 3, emphasising key metrics that validate this methodology as a practical alternative to CUHF studies.

The results are quite clear. Unlike the CUHF study model, which emphasises error rate comparisons and often struggles with statistical complexities surrounding subjective outcomes, this study took a more pragmatic approach. By shifting the focus to identify sources of negative transfer and defining acceptable error rates tailored to the product's safety profile, it provides a more targeted and efficient framework for human factors evaluations.

Of note, the two observed failures with the four-step autoinjector were attributed to mismatched injection times during sample preparation, with the injection time of the substitute device being two times longer than that of the RP. Both users lifted the two-step autoinjector from their skin too early, assuming that the injection duration would match that of their RP or be faster due to expectations tied to new technology. One user achieved

success on a third attempt during follow-up. Crucially, these injection time discrepancies are an artefact of the study setup and sample device creation and would not exist in a generic or biosimilar product application, as injection times are standardised across DDCP devices relative to the RP.

For PFSs with needle-safety systems, unsuccessful dose completions were linked to accidental early activation and confusion regarding the instructions for use, specifically about maintaining pressure during injection. These issues underscore the value of improved labelling to enhance usability and ensure treatment consistency without necessitating a clinical investigation.

ENSURING SAFETY, EFFICACY AND BIOEQUIVALENCE FOR SUBSTITUTABLE DDCPs

For substitutable DDCPs, the primary objective is clear – to deliver the same clinical effect and safety profile as the RP. Achieving this requires bioequivalence, which hinges on administering the full

therapeutic dose without error. This four-part study highlighted that RP-experienced users can successfully transition to an alternative device featuring “other design differences,” achieving the primary endpoint without the need for additional training. The results demonstrate that substitutable devices can maintain safety, efficacy and bioequivalence even when alternative designs are introduced. Importantly, root cause analysis and URRAs confirmed that design differences did not contribute to primary endpoint failures or negative transfer.

This study supports the viability of standard human factors validation methodologies for substitution and bioequivalence evaluations. By leveraging best practices from IEC 62366, ISO 14971 and FDA draft guidance, the proposed approach focuses on usability improvements and error mitigation specific to pharmacy substitution scenarios. Rather than mirroring reference designs exactly, it emphasises innovation and user-centric design to enhance safety and usability beyond the RP.

"AS INJECTION TECHNOLOGIES CONTINUE TO EVOLVE, PRIORITISING THE USER EXPERIENCE IS NOT JUST BENEFICIAL – IT IS ESSENTIAL."

As injection technologies continue to evolve, prioritising the user experience is not just beneficial – it is essential. Generic and biosimilar DDCP developers can innovate, striving for “state of the art” devices, while adhering to regulatory frameworks, such as the EU MDR. By assessing risk based on negative transfer and its impact on primary endpoints, developers can uphold high standards of care while streamlining market introduction. This flexible strategy accelerates access to affordable, substitutable DDCPs, meeting patient needs efficiently and effectively.

ABOUT THE COMPANY

Teva Pharmaceuticals is an innovative biopharmaceutical company, enabled by a generics business. From in-house innovation to strategic partnerships, Teva is persistent in the creation of innovative medications, generic medicines and biologics to increase the accessibility and affordability of existing medicines. The Teva Combination Product and Device R&D group is the internal Teva entity responsible for all device development for Teva’s combination product portfolio.

BIBLIOGRAPHY

1. “ANSI/AAMI HE75:2009/(R)2018; Human factors engineering – Design of medical devices”. AAMI, 2021.
2. Brunet-Manquat L et al, “Pre-ANDA strategy and Human Factors activities to de-risk pharmaceutical companies ANDA submission of drug-device combination products: case study of a formative Comparative Use Human Factors study”. *Expert Opin Drug*



Carrie O'Donel

Carrie O'Donel is Associate Director, MS&T, External Manufacturing at Teva Pharmaceuticals. She obtained her Bachelor of Arts in Physics from State University of New York, Buffalo State University (NY, US) in 2013, and her Master of Science in Mechanical Engineering from Villanova University (PA, US) in 2017.

T: +1 484 319 8798

E: carrie.odonel@tevapharm.com



Henri Akouka

Henri Akouka is Director of Biopharmaceutical Combination Products at Teva Branded Pharmaceutical Products R&D. He received his Bachelor of Science in Mechanical Engineering from Norwich University (Northfield, VT, US) in 1999 and has been developing combination products since 2007.

T: +1 732 647 6028

E: henri.akouka@tevapharm.com



Mark DeStefano

Mark DeStefano is Director, Global Technology Innovation, Combination Products and Device Research & Development at Teva Pharmaceuticals. He obtained his Bachelor of Science in Mechanical Engineering from Drexel University (Philadelphia, PA, US) in 1985 and has been developing medical devices since 1990.

T: +1 267 472 6984

E: mark.destefano@tevapharm.com

Teva Branded Pharmaceuticals R&D

145 Brandywine Parkway, West Chester, PA 19380, United States
www.tevapharm.com



Leslie Sanchez-Torres

Leslie Sanchez-Torres is a former employee of Combination Products and Device Research & Development at Teva Pharmaceuticals. She obtained her Bachelor of Science in Biomedical Engineering degree from Purdue University, West Lafayette (IN, US) in 2018.

3. *Deliv*, 2024, Vol 21(5), pp 767–778.
3. “IEC 62366-1:2015/AMD1:2020 Amendment 1 - Medical devices - Part 1: Application of usability engineering to medical devices”. IEC, 2020.
4. “ISO 14971:2019 Medical devices — Application of risk management to medical devices”. ISO, 2019.
5. Faulkner L, “Beyond the five-user assumption: Benefits of increased sample sizes in usability testing”. *Behav Res Methods*, 2003, Vol 35, pp 379–383.
6. “How People Learn”. National Research Council, 2000.