Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 142 (2026) 451–456

www.elsevier.com/locate/procedia

36th CIRP Design Conference (CIRP Design 2026)

Phantom validation framework: a tool for systematic and reproducible medical phantom assessment

Marie Wegner[a,b]*, Elisabetta Gargioni[b], Dieter Krause[a]

[a]Hamburg University of Technology, Institute of Product Development and Mechanical Engineering Design, Denickestr. 17, 21073 Hamburg, Germany
[b]Department of Radiotherapy and Radio-Oncology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

* Corresponding author. Tel.: +49-040-30601-4687; fax: +49-040-42878-2296. *E-mail address:* marie.wegner@tuhh.de

Abstract

Phantoms are physical models designed to mimic medical tissues and structures in medical imaging. They are used in medical research, quality assurance, and training. Despite their increasing application and widespread use, standardized frameworks for the verification and validation of phantoms remain insufficiently established, primarily due to the interdisciplinary challenges involved. This paper addresses this gap by identifying and categorizing essential validation aspects for medical imaging phantoms, derived from an extensive literature review. Building upon these categories—application validity, physiological validity, anatomical validity, user validity, economic validity, and concurrent validity—a structured validation matrix is proposed to guide users through a comprehensive interdisciplinary evaluation process. The matrix incorporates both objective measurements, such as imaging performance benchmarks and quantitative anatomical comparisons (e.g., Hounsfield Values, Dice coefficients, respectively), and subjective assessments from expert and user feedback. Emphasis is placed on aligning validation procedures with the specific purpose and intended use of each phantom, ensuring reproducibility through rigorous documentation of test conditions and outcomes. By integrating multidisciplinary criteria into a systematic validation framework, this work aims to enhance the reliability and usability of phantoms, providing a methodological tool embedded in the design process, which supports their development and application in advancing medical imaging technologies.

Keywords: phantom; medical imaging; validation; verification; framework; design method

1. Introduction

Physical models designed to replicate anatomical structures and tissue properties for medical imaging are referred to as phantoms. They are commonly used for clinical training, quality assurance, and research [1, 2]. By providing controlled and reproducible conditions, they serve as essential tools in the development, evaluation, and optimization of imaging technologies [3]. However, despite their increasing use across modalities and research, a standardized and transparent approach to phantom validation remains insufficiently established.

Validation often represents the final step in the design and development process—a critical point at which it is assessed whether the product meets stakeholder requirements and delivers the intended performance. To ensure reproducibility, traceability, and reliability, this step must be structured and clearly aligned with the intended use of the product [4]. Importantly, validation activities initiated early in development are key to guiding design decisions and improving outcomes [5, 6]. In engineering design, validation and verification are distinct yet complementary processes [4, 7]. Verification refers to confirming—by objective evidence—that the specified requirements are being implemented correctly [5, 7, 8]. It answers the question: “Have we built the system correctly?”

10.1016/j.procir.2026.05.290

[9]. Verification is performed at various system levels and development stages [8], ensuring that technical specifications are met at all times throughout the development. Validation, on the other hand, examines whether a method or system actually measures or performs what it is intended to [7, 8]. In engineering terms, it answers: "Have we built the correct system?" [4, 9]. It assesses the degree to which the final product satisfies stakeholder needs, and is typically conducted with their involvement. Furthermore, since product objectives often evolve during development, validation requires continuous alignment and re-evaluation to reflect updated insights [5]. This distinction is essential: a product may pass verification but fail validation, or vice versa. For example, a training phantom for ultrasound-guided needle insertion may contain the correct internal structures at the right anatomical locations, and its dimensions match the design specifications, thus verification is successful. However, if users find that the needle does not advance realistically through the tissue layers, or the ultrasound image lacks sufficient contrast to identify key features, the validation fails—because the model does not serve its intended training purpose. Conversely, if the material properties were not implemented exactly as specified, but users still report a highly realistic training experience, validation may be considered successful even though verification has failed.

Phantoms, particularly those used for medical training or imaging calibration, require both evaluations: they must meet technical criteria (e.g., accurate Hounsfield units, geometric fidelity), but also provide realistic tactile feedback or visual representation meaningful to the end-user. Validation must therefore go beyond metrics, incorporating qualitative and user-centered insights into model performance. In practice, however, the boundary between verification (ensuring that specified requirements are implemented correctly) and validation (ensuring that the product fulfills its intended purpose) is often blurred—especially in research-driven or prototype contexts. For the sake of clarity and consistency, the term validation will be used throughout this paper as a general umbrella term, encompassing both the verification of technical specifications and the evaluation of application-specific usability. This integrated perspective reflects the demands of phantom development, where both objective performance and user relevance must be assessed in parallel.

The diversity of phantom types and their modality-specific requirements make generalizable validation protocols difficult to define. Each imaging system—CT, MRI, ultrasound—demands distinct materials, physical behavior, and anatomical representations. Additionally, current literature often presents isolated, context-specific validations, frequently without sufficient methodological detail for reproduction, comparison or adaptation. This lack of standardization hinders both comparative assessments and the broader advancement of phantom-based research and development.

To address these issues, this work presents a comprehensive literature analysis on phantom validation approaches, identifies recurring criteria, and proposes a structured validation matrix. This matrix supports interdisciplinary, purpose-aligned, and traceable verification and validation.

2. Methodological Approach

To get an overview of phantom validation and verification approaches found in the literature and to gather generic categories that need to be addressed in phantom evaluation, a systematic literature review was conducted. The literature search was performed using Scopus and PubMed using the PRISMA statement [10] with the terms shown in Table 1.

Table 1. Search terms used for the literature assessment.

phantom*	AND	validat* OR verificat*	AND	method* OR approach* OR procedure* OR support* OR criteri*	AND	imag*
		(title)				(title-abs-key)

To ensure a focused and manageable result, the search was primarily restricted to titles. Broader searches including abstracts and keywords resulted an unmanageably high number of results (e.g., 14,390 in Scopus as of June 19, 2024), many of which lacked generic relevance to phantom validation procedures. These publications often focused on a single, specific phantom application rather than discussing validation frameworks. To account for potentially relevant research overlooked by the restricted search terms, the most highly cited papers, as well as research found using a snowball approach, were also screened, together with the search string results.

After removing duplicates, non-English/German records, and posters from the literature set, a total of 128 records were screened based on their titles and abstracts (see Fig. 1). Only seven publications were identified as meeting all the desired assessment criteria. These criteria provided a systematic approach to phantom validation, addressing both validation and requirements verification, as well as the consideration of transferability.

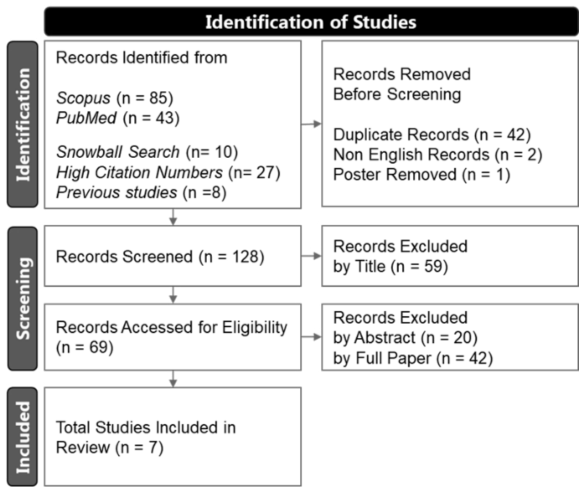

Fig. 1. PRISMA diagram [10] illustrating the performed systematic review of phantom validation and verification.

Across the literature, phantom validation is addressed in diverse yet often fragmented ways, reflecting the lack of a standardized framework. Many studies, such as those by Fricke [11], Bowen et al. [12], and Merken et al. [13], rely heavily on application-specific validation procedures, often blending verification of technical requirements with limited attempts at broader validation. Fricke, for example, proposes a recursive validation cycle and considers anatomical and physiological validity, emphasizes practical applicability, involving clinicians and using a questionnaire. Similarly, Bowen et al. and Malekzadeh et al. [14] focus on comparing phantom measurements with reference values (e.g., relaxation times or calibration results), but neglect criteria such as reusability, reproducibility, or user involvement. Studies involving user feedback, such as Merken et al. and Fricke, offer more holistic perspectives but remain modality-specific and lack transferability. Illi et al. [15], through their systematic review, highlight the heterogeneity of validation practices in additive manufacturing (AM) cardiovascular phantoms and advocate for standardized checkpoints—yet they stop short of defining a concrete framework. The proposal of a structured "4C" validation concept by Hacker et al. [16] underscores the need for coordinated effort, but remains largely theoretical and focused on biophotonic phantoms. The "4Cs" refer to: characterization, collaboration, communication, and commitment. Meanwhile, works like Hatamikia et al. [17] illustrate the frequent fusion of phantom validation with algorithm or application validation. Although geometric and physiological properties are assessed, these remain closely tied to the test scenario rather than to general phantom quality.

Overall, the literature reveals a clear trend: most studies conduct task-specific verification under the label of validation, with minimal emphasis on transferability, standardization, or comprehensive user evaluation methods. However, several key categories can be derived from the literature to generate a holistic framework for phantom validation.

3. Results

3.1. Validation Categories

In total, seven validity categories for phantoms were identified from the literature, two of which specifically address phantom application: imaging validity and therapy validity. The remaining categories encompass the validity of following categories: physiological, anatomical, economic, user and concurrent, each reflecting different phantom requirements and relevance in phantom design and evaluation.

These categories are further defined by a range of assessment criteria and aspects—such as: visual assessment, haptics, geometrical mapping, costs, durability, or user satisfaction—that, together, capture both the technical and practical requirements of phantoms. An overview of these categories and their associated criteria is illustrated in Fig. 2. In the following, each category will be explained in detail.

3.1.1. Imaging Validity

Imaging validity evaluates how well a phantom performs in medical imaging and should include both expert **visual assessment** and objective **property measurements**. Since phantoms are defined by their use in imaging, this assessment is always required. Visual evaluation examines material appearance, tissue contrast within and between structures, and the absence of artifacts. It can be done directly by users or through structured questionnaires in which imaging experts assess factors such as noise, tissue representation, and organ contrast.

Objective imaging property assessments focus on quantifying material-specific imaging parameters and their reproducibility. This includes the uniformity of values, as well as comparisons against predicted values derived from material tests, relevant literature, patient data, or reference phantoms. Reported properties may include, for example, CT numbers, relaxation times, linear attenuation coefficients, or the speed of sound propagation, depending on the intended imaging modality [18]. Importantly, while imaging property validation is the most common approach in phantom design literature, parameters are often not documented in a comprehensive way. Details such as acquisition protocols (e.g., CT tube voltage) or reconstruction settings are frequently omitted, making it difficult to reproduce and compare different studies [18]. A complete documentation of conditions is therefore essential to ensure that results can be replicated and that future iterations of a phantom can be benchmarked against a reliable reference.

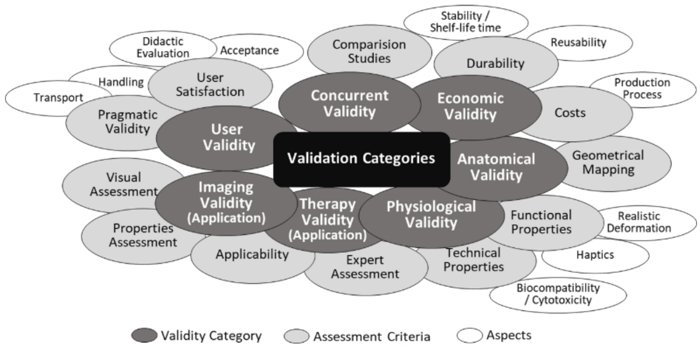

Fig. 2. Phantom validation categories and criteria identified.

3.1.2. Therapy Validity

Therapy validity concerns the applicability of the phantom in image-guided therapy applications, like radiotherapy, brachytherapy, catheterization, or surgical procedures. Validation should include quantitative measurements and expert evaluation, focusing on the phantom's performance within the therapeutic workflow. **Applicability** measurement may include therapy-specific assessments such as dosimetry, covering absorbed dose distribution, dose homogeneity and reproducibility in radiotherapy applications. Further **expert assessment** could involve questionnaires with therapy specialists to evaluate the phantom's suitability, usability and realism within the clinical workflow. For instance, this assessment could include the usability of instruments in the phantom or the realistic handling. The importance of subjective criteria depends on the intended use: in training phantoms for catheterization, realistic haptics may outweigh precise material properties, whereas radiotherapy dosimetry phantoms require high dosimetric accuracy and reproducibility.

3.1.3. Physiological Validity

While therapy validity assesses whether a phantom can be applied effectively within a therapeutic workflow (e.g., radiotherapy, surgery, catheterization), physiological validity addresses whether a phantom is capable of reproducing the physiological requirements desired for its intended application.

Physiological parameters refer to characteristics that describe functions and processes occurring within the body, such as deformation of soft tissues, organ motion, fluid flow, heart rate, or body temperature. If a phantom is designed to reproduce such parameters, they should be validated separately to ensure their fulfillment. For example, a deformable pelvic phantom may be required to simulate the deformation pattern of a human bladder during filling and emptying cycles [19].

Physiological validity is divided into technical and functional properties. **Technical properties** can be examined through measurable parameters obtained via testing, such as flow rate, density, elasticity modulus, tensile strength, and cytotoxicity. These provide a quantitative basis for assessing whether the phantom materials meet the physiological requirements of the simulated tissue or organ. **Functional properties** require a combination of objective measurements and expert assessments to evaluate more complex aspects (e.g., realistic organ deformation, motion, elasticity of a tissue, or pulse sequence). Depending on the application, structured expert questionnaires can be used to capture the realism and usability of these features.

3.1.4. Anatomical Validity

Anatomical validity is concerned with verifying that the phantom accurately represents the geometry and structural design of the targeted anatomy. This involves ensuring that dimensions, shapes, and spatial relationships of structures are realistic reproduced. Validation can be performed through a combination of quantitative measurements and expert evaluation. The required level of anatomical realism depends on the phantom's purpose. Simple geometric or calibration phantoms may only need dimensional accuracy and reproducible basic shapes. In contrast, anthropomorphic or patient-specific phantoms require higher fidelity, including fine structural detail and correct spatial relationships to represent clinically relevant anatomy.

On the quantitative side, comparisons can be made against reference datasets, such as patient imaging or anatomical atlases. Common similarity metrics include the structural similarity index measure, Dice similarity coefficient, and Hausdorff distance, which provide numerical indicators of how closely the phantom corresponds to real anatomical data [20]. For more complex phantoms, regional fidelity—such as alignment of bone contours or organ boundaries—may also be analyzed. Complementing these metrics, anatomical experts can provide qualitative assessments, for example through structured questionnaires or scoring systems. Their evaluation focuses on whether key anatomical features are correctly represented and recognizable, whether proportions appear realistic. By combining quantitative comparisons with expert-driven feedback, anatomical validity ensures that a phantom not only has realistic dimensions but also delivers a functionally recognizable and clinically meaningful representation of the desired human anatomy.

3.1.5. User Validity

User validity addresses both the acceptance and satisfaction of the end-users (e.g., clinicians, trainees, technicians) and the pragmatic considerations related to the phantom's handling and usability in practice. This category emphasizes not only whether the phantom functions correctly but also whether it is perceived as useful, realistic, and practical by those who use it in training, research, or clinical workflows. **User satisfaction** is commonly assessed through questionnaires, interviews, or structured evaluation sessions with trainees and clinical experts. Key aspects may include the skill transferability, the didactic effectiveness, the training performance, the realism of difficulty, and the overall acceptance. **Pragmatic validity** refers to the usability of the phantom and is equally important. This includes evaluations of the transportability and the handling of the phantom. User feedback on these pragmatic aspects is typically collected via questionnaires, usability testing, or observation.

In summary, user validity emphasizes aligning a phantom with the expectations, needs, and practical constraints of its users. Even if a phantom performs well anatomically or in imaging, limited user acceptance or feasibility reduces its long-term value in training or therapy. Although crucial, user validity is rarely addressed explicitly, largely because it depends on subjective feedback, is difficult to quantify, and requires longitudinal studies with diverse users. As a result, many phantoms are technically validated but lack systematic evidence of acceptance or practical utility.

3.1.6. Economic Validity

Economic validity is another often underreported aspect in phantom studies. It involves a systematic evaluation of production process, material and manufacturing costs, and long-term durability of the phantom. The manufacturing process is assessed to verify feasibility, reproducibility, and cost-effectiveness. This includes evaluating whether the chosen production methods (e.g., additive manufacturing, molding) can be reproduced consistently. Economic considerations also involve comparing alternative materials and fabrication techniques in terms of price-performance ratio, supply availability, and ease of processing.

Durability evaluation should address the phantom's reusability, its long-term stability, and shelf lifetime. This could involve performing mechanical and material stability tests over time. For gel-like mixtures, microbiological stability checks (e.g., imprint tests with microbial indicators) can be used to analyze the shelf lifetime of a phantom.

3.1.7. Concurrent Validity

The final category, concurrent validity, entails measuring and comparing the phantom's performance in relation to other phantoms, whether commercial products or alternative prototypes found in the literature. This comparison can involve both quantitative and qualitative assessments, such as matching imaging properties, anatomical accuracy, or training outcomes against established references.

Concurrent validity is particularly valuable for new phantom design, as it can highlight improvements, such as enhanced anatomical realism, reduced cost, better usability, or improved durability. These comparisons may be performed through side-by-side imaging studies, parallel training evaluations, or structured expert assessments. By positioning a new phantom against existing solutions, concurrent validity ensures it is assessed against meaningful benchmarks and helps determine whether it is a viable alternative, a complementary tool, or a significant advancement for its intended application.

4. Evaluation (Validation Matrix)

To translate the introduced validation categories into a practical evaluation tool, a structured validation matrix was developed. This matrix can be used as a general checklist for phantoms, enabling both qualitative and quantitative assessment within a systematic framework. Each validation category is assigned a score on an ordinal 1–5 Likert scale (ranging from very poor to very good). Quantitative metrics (such as organ dimensions in millimeters or HU values of biological tissue) are converted into this ordinal format, allowing them to be integrated into the overall evaluation process. To emphasize the relative importance of different aspects, a weighted point-scoring system is applied. Each criterion receives a weighting factor between 0 and 100, with the total sum across all criteria equaling 100, ensuring that essential parameters are prioritized over desirable requirements. The weightings for each category can be determined through discussions within the stakeholder team, following a requirements-prioritization approach based on the relevance of each validation aspect to the intended application. Individual validation categories are composed of multiple aspects; each aspect is scored separately, and these scores are then combined - either by averaging or weighted aggregation - to produce a comprehensive score for the overall category.

In typical phantom development workflows, the validation matrix is applied at several key stages. During requirements definition, it supports the selection of relevant categories and the assignment of application-specific weightings. During prototyping, it guides which tests and measurements should be conducted and ensures that critical aspects (e.g., imaging appearance, dosimetry, geometric fidelity) are explicitly evaluated. During final validation, the matrix serves as a gate criterion before release, providing a transparent and reproducible summary of whether the phantom meets the predefined performance thresholds. The roles typically involved in this process include the project lead, clinical expert, medical physicist, and an independent reviewer. Resource and scheduling needs align with standard testing stages already present in phantom development, ensuring that integration of the matrix does not impose substantial additional overhead.

To provide an example of the matrix application, the validation matrix was applied to GALE, a CT and radiotherapy phantom of the lower extremities [21]. The results are presented in Table 2. Imaging validity was determined by the phantom's overall appearance in CT imaging and its HU parameters. The overall appearance consisted of evaluations of soft-tissue contrast, bone-marrow contrast, compact-bone contrast, and the absence of artifacts. Each of these aspects was scored individually, and the resulting values were averaged to yield a comprehensive score for the CT-appearance category. For GALE, imaging parameters, dosimetry evaluation, and positioning accuracy received the highest impact factors, as the phantom is primarily used for quality assurance. HU values were particularly emphasized, with ratings based on uniformity, reproducibility, and acceptable deviation ranges benchmarked against literature values and a commercial phantom. For example, for CT-number accuracy, a deviation below 10% corresponds to a score of 5.

Table 2. Application of the Phantom Validation Matrix to GALE, a CT phantom of the lower extremities (GALE is described in [21]).

Validation Category	Factor	Score
Imaging Validity		
Overall Appearance in CT	20	4.5
Soft Tissue Contrast		5
Bone Marrow Contrast		3
Compact Bone Contrast		5
Free of Artefacts		5
Overall HU Parameters	20	4.7
Soft Tissue HU Parameters		5
Bone Marrow HU		4
Compact Bone HU		5
Therapy Validity	20	4.5
Appropriate Radiotherapy Planning and Execution		5
Dosimetry Evaluation with Radiochromic Film		5
Dosimetry Evaluation with Ionization chamber		5
Dosimetry Evaluation with OSL		5
Physiological Validity	10	5
Positioning in Flat and Bent Positions		5
Compatible with Patient Pillows used for the Positioning		5
Anatomical Validity	10	5
Shape comparison		5
Overall Dimensions		5
Connection to ART-Phantom		5
User Validity	10	4.7
User Satisfaction		5
Handling		4
Transportability		5
Economic Validity	5	4.7
Evaluation of Long-term Stability		4
Reusability		5
Production Process and Cost		5
Concurrent Validity	5	5
Compatible with ART Phantom		5
Overall Validation Rating	100	4.7

5. Conclusion

This work develops a structured general guideline with clearly defined categories to support the validation of medical phantoms. These categories serve as a practical framework, outlining both general assessment criteria and subjective evaluation aspects. By ensuring consistency, this approach enhances the thoroughness of the validation process and provides a robust foundation for evaluating phantoms across diverse applications. The validation process relies on input from a broad range of experts and user groups, since developers are often not directly familiar with the clinical or training context. In particular, subjective assessments—such as the evaluation of usability in training scenarios—must be included to ensure acceptance and practical relevance.

Based on the analysis of existing literature and practice, the phantom validation is divided into the following main categories: imaging, therapy, physiological, anatomical, user, economic, and concurrent validity. Each category is composed of several sub-criteria. From these a validation matrix as both a decision-support and documentation tool, defining the scope, methods, and criteria for validation and providing a consistent structure for carrying out the process is derived. At this stage, test setups—including imaging devices and relevant parameters—must be specified, together with the roles of experts responsible for conducting and evaluating the tests. A timeline with clear milestones ensures transparency and supports efficient execution. The assessment must be aligned with the requirements defined earlier in the development process, which provide concrete benchmarks such as acceptable HU deviations or mechanical stability thresholds. Depending on the intended use case, additional stability or durability tests may be necessary to assess long-term performance.

The validation matrix therefore fulfills several important functions. First, it reduces subjectivity, enhances reproducibility, and creates transparent records for quality assurance and research dissemination. It also facilitates comparisons between prototypes and commercial phantoms, supports internal decision-making, and enables external benchmarking. By integrating quantitative measurements with expert feedback, the matrix ensures comprehensive evaluation across technical, physiological, economic, and training-related aspects.

To illustrate the application of this approach, the validation matrix was applied to GALE, a CT and radiotherapy phantom of the lower extremities. In this example, imaging validity was assessed through multiple sub-criteria, such as soft tissue contrast, bone marrow contrast, compact bone contrast, and the absence of artifacts, which were individually scored on the ordinal scale and then averaged into a comprehensive category result. This example demonstrates how the matrix can be applied in practice, translating abstract criteria into structured, reproducible validation outcomes.

The key benefit of this methodological approach lies in its completeness: by systematically considering all relevant elements, the validation process becomes more reliable, reproducible, and transferable across different phantom types and applications. Nonetheless, some limitations remain. No framework can capture all possible aspects for every phantom application, and certain criteria may not be listed or applicable in all contexts. Moreover, overall numeric scores retain a degree of subjectivity, influenced by expert judgment and applied weightings. Despite these limitations, the framework is highly valuable for comparing phantoms and identifying design weaknesses, such as insufficient long-term stability or suboptimal handling. In this sense, it not only supports validation but also serves as a tool for iterative development and optimization.

Future work will focus on strengthening the robustness and broader applicability of the matrix. Planned studies will assess inter-rater reliability, reproducibility of scores, and sensitivity to weighting variations. Pilot applications across diverse phantom types will evaluate generalizability. These efforts aim to consolidate the Framework as a reliable and widely applicable tool for methodical, interdisciplinary, and reproducible phantom design and validation.

Acknowledgements

Marie Wegner is supported by the Life-Science Add-on Fellowship of the Joachim Herz Foundation.

References

[1] DeWerd, L.A., Kissick, M., 2014. The Phantoms of Medical and Health Physics. Springer, New York.
[2] Wegner, M., Gargioni, E., Krause, D., 2023. Classification of phantoms for medical imaging, Procedia CIRP, 119, p. 1140.
[3] Valladares, A., Beyer, T., Rausch, I., 2020. Physical imaging phantoms for simulation of tumor heterogeneity in PET, CT, and MRI: An overview of existing designs. Med Phys, 47, p. 2023.
[4] Albers, A., Behrendt, M., Klingler, S., Matros, K., 2016. Verifikation und Validierung im Produktentstehungsprozess, in Handbuch Produktentwicklung, Carl Hanser Verlag GmbH & Co. KG, München.
[5] Pahl, G., Beitz, W., Feldhusen, J., Grote. K.H., 2007. Engineering Design: A systematic approach. Springer, London.
[6] Albers, A., Reinemann, J., Hirschter, T., Fahl, J. et al., 2019. Validation-Driven Design in the Early Phase of Product Development, Procedia CIRP, 84, p. 630.
[7] Maropoulos, P.G., Ceglarek, D., 2010. Design verification and validation in product lifecycle, CIRP Annals, 59, p. 740.
[8] VDI/VDE, 2021. VDI-Guideline 2206: Development of mechatronic and cyber-physical systems, Verein Deutscher Ingenieure (VDI).
[9] Gräßler, I., Oleff, C., 2022. Systems engineering: Verstehen und industriell umsetzen. Springer Vieweg, Berlin, Heidelberg.
[10] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I. et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev, 10, p. 89.
[11] Fricke, C., 2013. Ein anthropomorphes Phantom zur Evaluation eines chirurgischen Assistenzsystems mit intraoperativer Bildgebung, Leipzig.
[12] Bowen, D.L., Touchet, T.J., Maitland, D.J., McDougall, M.P., 2023. Technical note: The design and validation of a multi-modality lung phantom. Med Phys.
[13] Merken, K., Monnens, J., Marshall, N., Johan, N. et al., 2023. Development and validation of a 3D anthropomorphic phantom for dental CBCT imaging research. Med Phys 50, p. 6714.
[14] Malekzadeh, M., Abbasi-Rad, S., Shahgholi, M., Naghdi, P. et al., 2019. Design and Validation of Synchronous QCT Calibration Phantom: Practical Methodology. J Med Imaging Radiat Sci 50, p. 157.
[15] Illi, J., Bernhard, B., Nguyen, C., Pilgrim, T. et al., 2022. Translating Imaging Into 3D Printed Cardiovascular Phantoms: A Systematic Review of Applications, Technologies, and Validation. JACC Basic Transl Sci 7.
[16] Hacker, L., Wabnitz, H., Pifferi, A., Pfefer, T.J. et al., 2022. Criteria for the design of tissue-mimicking phantoms for the standardization of biophotonic instrumentation. Nat Biomed Eng 6, p. 541.
[17] Hatamikia, S., Gulyas, I., Birkfellner, W., Kronreif, G. et al., 2023. Realistic 3D printed CT imaging tumor phantoms for validation of image processing algorithms, Physica Medica, 105, p. 102512.
[18] Wegner, M., McMillan, N., Krause, D., 2025. Supporting medical phantom design through a comprehensive design catalogue for phantom materials, Proceedings of the Design Society, 5, p. 2931.
[19] Wegner, M., Schmiech, J., Sobirey, E., Krause, D. et al., 2024. Requirement analysis in medical phantom development: a survey tool approach with an illustrative example of a multimodal deformable pelvic phantom, Front Phys, 12.
[20] Lin, B.Z.H., Zhang, E.T., Ng, H., Tan, M.Y.L. et al., 2025. Design and additive manufacture of patient-specific head phantom for radiotherapy, Materials & Design, 252.
[21] Wegner, M., Alizadeh Azbari, N., Krause, D., Gargioni, E., 2025. Development of an anthropomorphic phantom of the lower extremities for feasibility studies and verification of total-body irradiation, Physica Medica, 136, p. 1120