ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO REAL WORLD DATA AND DIGITAL ONCOLOGY

REVIEW

Real-world radiology data for artificial intelligence-driven cancer support systems and biomarker development

D. Navarro-Garcia[1], A. Marcos[1], R. Beets-Tan[2], L. Blomqvist[3], Z. Bodalal[2], D. Deandreis[4], M. Crispin-Ortuzar[5], F. Gallagher[5], T. Giandini[6], M. J. Graves[5], N. Lassau[4], K. Maier-Hein[7], A. Prelaj[6], P. Schader[7], H.-P. Schlemmer[7], O. Sedlaczek[7], M. Vaiani[6] & R. Perez-Lopez[1*]

[1]Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain; [2]Netherlands Cancer Institute (NKI), Amsterdam, The Netherlands; [3]Department of Nuclear Medicine/Hospital Physics, Karolinska University Hospital, Stockholm, Sweden; [4]Gustave Roussy Cancer Center, UMR 1281 INSERM, CEA CNRS, Université Paris-Saclay, Paris, France; [5]University of Cambridge, Cambridge, UK; [6]Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; [7]The German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ), Heidelberg, Germany

Available online xxx

The integration of artificial intelligence (AI) and real-world data (RWD) opens up a new paradigm for exploiting radiology data to develop advanced diagnostic and therapeutic support systems. This review explores the advantages and challenges of utilizing vast digital image datasets from routine clinical practice and computational AI capabilities to enhance cancer patient care. Particularly, the application of AI to radiology data has shown promise in developing tools that automate clinical processes, such as tumor detection, while also identifying novel biomarkers in cancer for potential treatment support. Deep learning models, crucial for this transformation, require substantial data, making RWD a valuable resource for accelerating assay development. RWD offer diverse, extensive data reflecting real-world clinical practices, complementing clinical trial data and providing a broader understanding of patient populations and treatment responses. However, challenges such as data access, variability in quality, and processing complexities must be addressed. Standardizing data processing protocols and feature extraction methods is essential to ensure reproducibility and clinical applicability. Moreover, building trust among clinicians, patients, and regulatory bodies is crucial for successful implementation. This review highlights the potential of AI to analyze RWD imaging data and radiology reports, extracting relevant information and enhancing biomarker discovery. To facilitate practical use, we offer tools to address the main challenges associated with utilizing real-world imaging data, such as key aspects of image access and data processing.

Key words: real-world data, radiology, artificial intelligence, oncology, biomarkers

INTRODUCTION

The availability of vast digital image datasets and the computational capabilities of artificial intelligence (AI) are revolutionizing the field of radiology.[1,2] Images are not only a tool for diagnosis by visual inspection, but are also rich sources of data for digital AI-based models. These advancements have not only led to novel disease detection assays, particularly useful in cancer screening programs,[3-5] and computer-aided tools for organ segmentation for radiotherapy planning,[6] but have also opened the door to developing novel biomarkers for improved patient and tumor characterization.[7,8] Such progress has resulted in promising prognostic, predictive, and response biomarkers based on image patterns, which, despite not yet being implemented in clinical practice, are opening new opportunities for research and hold significant potential for future utility.

Traditionally, handcrafted radiomic features have been integrated with classical machine learning models to enhance clinical tasks. This process involves manually selecting specific features for analysis. In contrast, deep learning employs advanced neural networks to automatically extract complex patterns from raw data. Recent advancements in deep learning have demonstrated that it often outperforms traditional methods, broadening the potential applications of radiomics and improving patient care in the diagnosis and treatment of diseases.[9,10] However, deep learning models need large volumes of data to effectively identify patterns associated with specific diseases or clinical outcomes. Utilizing real-world data (RWD) can accelerate the development of deep learning-based

*Correspondence to: Dr Raquel Perez-Lopez, Vall d'Hebron Institute of Oncology (VHIO), 11-13 Saturnino Calleja, 08035 Barcelona, Spain. Tel: +34 932543450 ext. 8682
E-mail: rperez@vhio.net (R. Perez-Lopez).

assays, as digitalization of medical images and their associated reports is common in clinical practice. These data are typically collected in electronic health records (EHRs), which enable the seamless integration of patient information from diverse sources.

In conjunction with vision language models for imaging data, large language models are emerging as promising tools for the comprehensive analysis of standard medical records.[11] Currently, the integration of medical records and imaging data using these models is a highly active research area due to the significant enhancements in performance that text data can provide. By combining RWD from medical records and imaging data, these models have the potential to offer deeper insights into the complex relationships between clinical data and radiological findings.[12-14]

However, despite the many positive aspects, the use of RWD presents significant challenges. While RWD offer a rich, diverse, and extensive source of data that mirror clinical practice, they also come with variability in data quality and completeness. This variability showcases the need for robust methods for data processing and analysis to ensure reliability. Moreover, collecting, standardizing, and integrating data from various sources can be complex and time-consuming. Addressing these challenges is crucial to fully harness the potential of RWD for generating generalizable and applicable insights in real-world settings.

In this review, we provide an overview of the advantages and challenges associated with exploiting real-world image data and implementing AI-aided imaging tools and radiomics-based biomarkers in clinical scenarios. Additionally, we show sources of real-world image data, describe the cornerstone of image data processing, and describe ways to enhance trust in this new paradigm.

RADIOMICS EVOLUTION AND THE RISE OF FOUNDATION MODELS IN MEDICAL AI

Radiology began in 1895 with Röntgen's discovery of X-rays, the first technology to non-invasively visualize the human body (Figure 1). The 20th century brought major advancements, including ultrasound (1950s), computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) (1970s), revolutionizing cancer imaging. Importantly, the digitization of medical imaging improved accessibility, enabled the collection of large datasets, and facilitated data-driven insights. However, for decades, radiological assessment remained largely dependent on visual interpretation.

A major shift came in the early 2010s with the emergence of radiomics, driven by advances in computational power and statistical methods. First introduced by Lambin et al. in 2012, radiomics extracts quantitative features, such as volume, intensity, shape, and texture, from medical images, enabling statistical analysis to uncover clinically relevant patterns.[15] Recognizing the need for standardization, Gillies et al. published radiomics guidelines in 2016, establishing a framework for reproducibility and large-scale validation.[16]

Deep learning further transformed radiomics, moving beyond handcrafted feature extraction to models that learn directly from raw images. Convolutional neural networks, such as U-Net and ResNet, have become essential tools for tumor detection and segmentation.[17,18] Additionally, deep learning has enabled multimodal AI, integrating radiology with pathology and genomics to enhance prognostic predictions and treatment response assessments.[19] Despite the initial skepticism, AI is now embedded in clinical practice, with more than 700 Food and Drug Administration-approved AI-enabled radiology devices as of December 2024, many actively used for lesion detection and radiotherapy planning.[20]

The latest breakthrough is the rise of foundation models, which are reshaping AI across multiple fields, including natural language processing and medical imaging.[21] These large-scale neural networks, pretrained on vast unannotated datasets, can be fine-tuned with smaller labeled datasets, making AI adoption more efficient and accessible.[22] Models like GPT and DALL-E exemplify their adaptability. In radiomics, foundation models streamline research and clinical workflows, enabling deep learning on small patient cohorts and overcoming previous machine learning limitations.[23,24] Their ability to integrate multimodal data, such as radiology reports and diverse imaging modalities, enhances RWD utilization, positioning them as a transformative force in medical imaging.[25]

RWD IN ONCOLOGY: DEVELOPING BIOMARKERS BEYOND CLINICAL TRIALS

While clinical trials remain the gold standard for evaluating the efficacy and safety of new treatments as well as associated biomarkers, RWD have emerged as a valuable source, providing insights into the practical application and performance of these therapies and tools outside the controlled trial setting.[26]

RWD encompass data collected from various sources, such as EHRs, disease registries, and patient-reported outcomes. These data can offer a comprehensive understanding of patient characteristics, treatment patterns, and long-term outcomes.[27,28] Particularly, the integration of RWD into the development of biomarkers in oncology holds significant promise as it can be complementary to clinical trial data.

There are several differences in how RWD are generated and used when compared with the data generated in clinical trials (Figure 2). One of their clear advantages is the possibility to study larger and more diverse patient populations than those typically included in clinical trials. Trial populations are carefully defined through inclusion and exclusion criteria to minimize bias. This careful design makes extrapolating findings to less controlled populations a significant challenge. Moreover, clinical trials often underrepresent certain populations, as participants typically have fewer comorbidities, better performance status, and are younger.

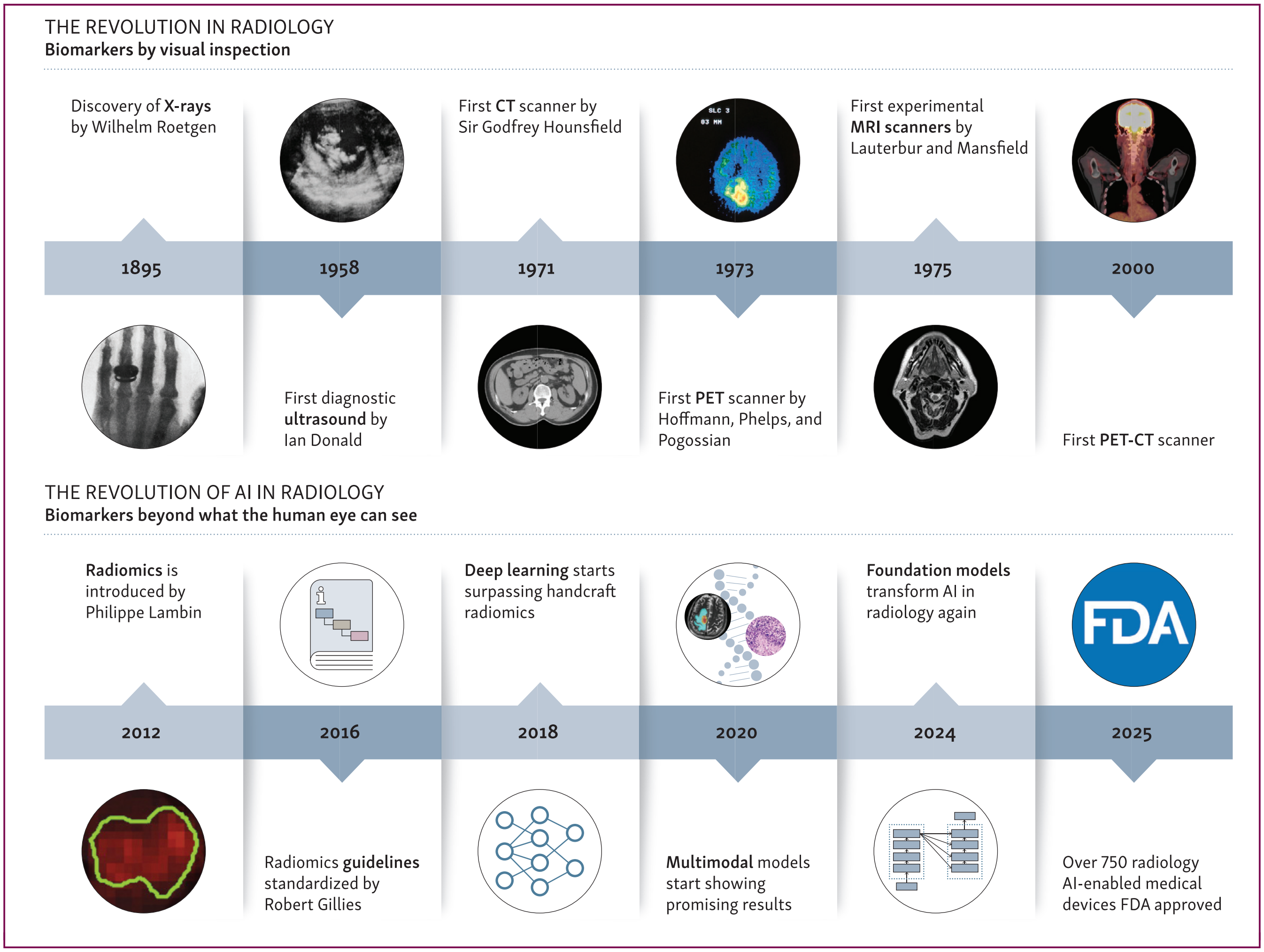

Figure 1. Evolution of AI biomarkers in radiology. Timeline of key breakthroughs in radiology and AI-driven advancements is shown. The top section illustrates the evolution of medical imaging modalities. The bottom section highlights the AI revolution in radiology, from the emergence of handcrafted radiomics to deep learning, multimodal AI, and foundation models.

AI, artificial intelligence; CT, computed tomography; FDA, Food and Drug Administration; MRI, magnetic resonance imaging; PET, positron emission tomography.

As a result, a significant portion of the real-world population is not adequately represented in clinical trial data.[29] Consequently, AI-aided tools developed in this context may lack reproducibility or be technically difficult to implement in routine clinical practice.

RWD are extracted directly from broader populations, with much less controls. RWD datasets typically include many centers and data generation protocols coming from diverse sources. Therefore, although the research workflow with the data can be similar, their less controlled origin makes them more suitable for the identification of biomarkers that are more representative of real-world populations and may be more generalizable. Therefore, beyond simply increasing the volume of data, RWD provide complementary information to clinical trial data, which is highly valuable for the development of deep learning-based tools and for demonstrating their true applicability in clinical practice. Importantly, the diversity inherent in RWD can help AI models become more robust under varied patient demographics and conditions. Moreover, the capability of AI models for continuous learning allows for the adaptation and improvement of biomarker models over time. As more data become available, AI systems can update and refine their algorithms, leading to more accurate and reliable biomarkers.

CHALLENGES IN AI-ENHANCED BIOMARKER DEVELOPMENT USING RWD

AI algorithms can process vast amounts of heterogeneous data, identifying subtle associations and interactions that could lead to the discovery of new biomarkers and the development of tools to automate clinical routine tasks.[30] Since AI requires large amounts of data, the use of RWD for this purpose is opening new possibilities but also comes with associated challenges (Table 1).

A primary concern with deep learning models is their reliance on large datasets for effective training. Leveraging RWD from diverse sources poses a promising strategy to enhance the performance of these data-intensive models.

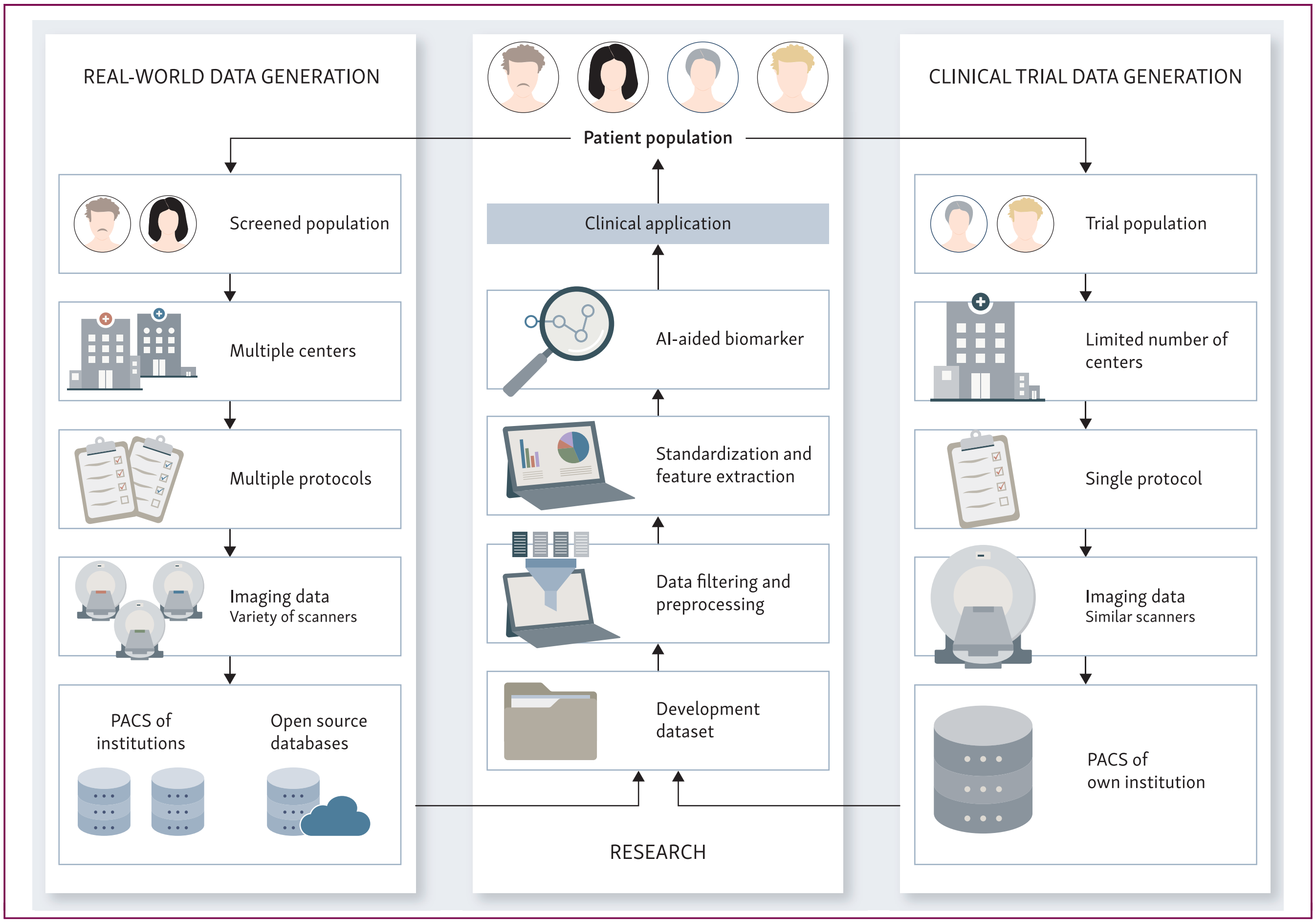

Figure 2. Comparison of imaging data generation processes in RWD and clinical trials. Key differences in data generation of real-world data as compared with clinical trials are shown. It also illustrates how data can be used for radiomics-based biomarker discovery and how this biomarker is then applied to the general patient population.
AI, artificial intelligence; PACS, picture archiving and communication systems; RWD, real-world data.

However, the high variability of the data can compromise its quality and consistency. Importantly, variations in data collection methods, missing data, and the influence of local health care systems can introduce biases and inconsistencies, complicating the development and validation processes.[31] Particularly, in the context of radiology, there may be variations in how images are captured or annotated according to different clinical practices. Thus, implementing data preprocessing and quality control measures to some degree is crucial to mitigate these challenges and ensure the reliability of samples for model development and application.

Table 1. Opportunities and challenges in AI-based biomarker development using RWD

Opportunities
Higher volumes of data: Removing the need of a controlled population enables access to bigger datasets
Real-world applicability: Diverse representation of the population
Generalizability: Due to the heterogeneous nature of data, findings on RWD have more potential to be generalizable to wider populations
Cost efficiency: Due to its retrospective nature, temporal and monetary costs are reduced compared with clinical trials
Challenges
Data quality: Variability across sources generates big differences in data quality, making it challenging to apply standard processing techniques
Signal dilution: Certain patterns seen in controlled populations might be lost in real-world datasets
Reproducibility: Due to data heterogeneity, results might not be reproducible in more controlled populations
Privacy and security: Use of many sources raises extra difficulties to ensure adequate and secure treatment of the data

Main opportunities and challenges in the development of new AI-based biomarkers using RWD are presented.
AI, artificial intelligence; RWD, real-world data.

At the same time, a major limitation in existing AI studies from controlled clinical trials is the lack of generalizability, where algorithms cannot be validated on independent datasets that are disparate from the training sets. In this regard, access to large and diverse RWD datasets facilitates the learning process and improves generalizability, which is crucial for developing tools and biomarkers that are applicable across different populations and clinical settings. However, this can also be detrimental in certain scenarios where patterns specific to subpopulations might get lost in the diversity, causing signal dilution. This issue becomes more pronounced when exploring advanced imaging techniques, such as tracking immune cells with radiotracers, or quantifying specific tissue properties using MRI.[32-34] These require highly detailed, quantitative data, typically obtained from dedicated, prospective clinical trials, rather than from observational RWD. Thus, while RWD are invaluable for improving generalizability, they fall short when the goal is to gain a detailed understanding of specific biological processes or disease characteristics. Therefore, the choice between controlled clinical trial data and diverse RWD should be guided by the specific objectives of the study.

Of significant importance are ethical considerations, such as ensuring patient privacy. In the case of medical image data, this requires proper anonymization by removing any personally identifiable information, such as names or dates of birth. Special attention should be given to head scans or individuals with distinct anatomical characteristics, as these images may allow for patient identification even without explicit personal data. In such cases, advanced techniques, such as pixel-based modifications or blurring of identifying features, should be applied to further protect privacy. Decentralized AI model development is another approach to minimize ethical and privacy concerns. By training models across multiple sites without sharing sensitive data, the risk of breaches is significantly reduced, making this approach particularly suited for multicentric studies.[35]

ADDRESSING CHALLENGES IN RADIOMICS-BASED BIOMARKER DEVELOPMENT FROM REAL-WORLD IMAGE DATA

Real-world image data offer immense potential for radiomics-based biomarker development, yet their effective use is challenging due to several factors. Addressing these challenges through the implementation of robust methods, techniques, and protocols is essential for the development and validation of reliable and clinically useful radiomics-based biomarkers (Figure 3).

Signal dilution

The use of RWD presents a significant challenge due to their inherent heterogeneity. Factors such as patient demographics, type of treatment, cancer type, and comorbidities contribute to a diverse dataset in which specific patterns and characteristics of certain subpopulations may be diluted or lost. This 'signal dilution' problem can be critical when the hidden patterns are key for clinical decisions that improve patient outcomes. To mitigate this issue, it is crucial to implement strategies that involve stratifying subpopulations during experimental designs. Such stratification helps in isolating the effects and responses unique to distinct groups, enhancing the clarity and applicability of research findings. Additionally, validating these findings in external cohorts is essential, ideally using prospective studies that accurately represent the subpopulations of interest.

Privacy and security

Addressing the issues of privacy and security in the use of RWD is critical to maintaining the integrity and trustworthiness of health care research. Privacy concerns primarily revolve around the anonymization of patient data to prevent re-identification, securing patient consent for data usage, ensuring compliance with regulations, and adhering to principles of fair data use to avoid biases and ensure equity. Each of these steps requires meticulous attention to detail to protect patient identities while still allowing for the valuable insights that RWD can offer. On the security front, it is essential to implement robust measures for secure data storage and management, which help prevent unauthorized access and potential data leakage. Moreover, federated or swarm learning platforms, where only model features and weights are shared rather than the raw data, offer a viable solution for enhancing data privacy.[36] In the era of foundation models, where embeddings serve as key representations of the population, data privacy concerns are further mitigated, as the original data are not directly exposed.[37]

Acquisition protocols

An important source of variability in imaging RWD arises from the use of different scanner manufacturers, image acquisition settings, inter- and intraoperator variability, and technical considerations such as the use of contrast agents or the choice of reconstruction algorithm. Standardization of image acquisition and processing is crucial for the reproducibility and generalizability of radiomics-based biomarkers.[38,39] Lack of standardization can lead to significant variability in radiomic features, limiting their clinical adoption. Initiatives like the Quantitative Imaging Biomarkers Alliance (QIBA), the European Imaging Biomarker Alliance (EIBALL), and Imaging Biomarker Standardization Initiative (IBSI) aim to establish consensus protocols for imaging,[40-43] and adhering to these standardized imaging protocols can help reduce variability.

Image quality

The quality of real-world image data can vary significantly due to differences in equipment, acquisition protocols, and operator expertise. Variability in image resolution, contrast, and noise levels, as well as the presence of artifacts, can introduce inconsistencies that affect the accuracy and reliability of radiomic features extracted from the images. Common artifacts, such as motion artifacts in MRI or beam-hardening artifacts in CT, can distort the image data, challenging feature extraction and leading to erroneous biomarker identification. Moreover, in clinical practice, images may be truncated, incomplete, or may not cover the full extent of the region of interest due to patient movement or other technical issues. To address this, algorithms have been designed to minimize the effects of artifacts, which in recent years have improved significantly due to the use of deep learning.

Matching resolutions between images is essential for the correct processing of images in algorithms. Since RWD can come in very different resolutions, it is important to ensure that techniques such as interpolation or downsampling are used for a proper match. Besides, RWD images can also be available with different contrast, which distorts the

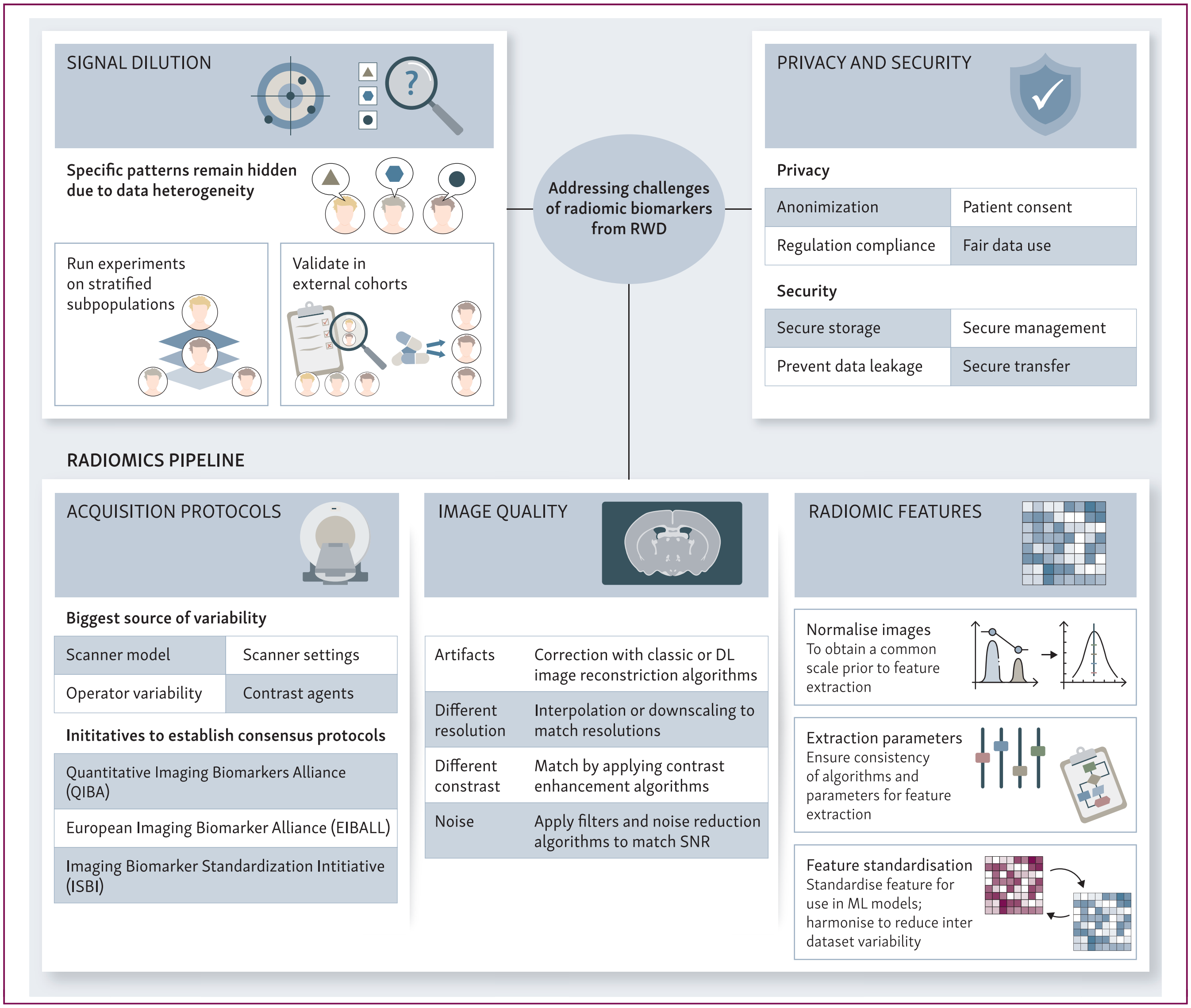

Figure 3. Addressing challenges of radiomics-based biomarkers from RWD. An overview on how to address the main challenges of radiomic-based biomarker development from RWD, including signal dilution, privacy and security, and those related to the radiomics pipelines (acquisition protocols, image quality, and radiomic features), is presented.
AI, artificial intelligence; DL, deep learning; EIBALL, European Imaging Biomarker Alliance; IBSI, Image Biomarker Standardization Initiative; ML, machine learning; QIBA, Quantitative Imaging Biomarkers Alliance; RWD, real-world data; SNR, signal-to-noise ratio.

processing of the images in the algorithms. Ensuring contrast match between images by applying contrast enhancement algorithms is therefore also a factor to take into consideration.

Finally, real-world image data are often subject to various sources of signal noise, which can obscure meaningful patterns and features from image data. Effective noise reduction techniques are essential to enhance the signal-to-noise ratio and improve the reliability of radiomic features. Some filtering and denoising algorithms, such as Gaussian smoothing, anisotropic diffusion, and wavelet denoising, can be used to mitigate this issue.

AI processing

The use of imaging RWD in AI poses many challenges to be addressed specific to its processing. For AI models to perform optimally, a careful preprocessing must be carried out to standardize image quality across different data sources. This involves applying the above-mentioned image quality methods, and to normalize the images to a common scale to mitigate variations due to different acquisition settings.[39]

Ensuring consistency in feature extraction is also crucial. When doing handcrafted radiomics, consistency in the extraction algorithms and parameters is key. Open source libraries such as PyRadiomics[44] provide standardized pipelines to ensure robustness across studies. Besides, to enter prediction models, these extracted features must be standardized and harmonized to accommodate variations inherent in RWD. By contrast, deep learning approaches automate many of these processes within the neural network architecture. This automation can potentially reduce the complexity of manual preprocessing and feature standardization, although some preprocessing steps may also be needed.

RADIOMIC-BASED BIOMARKERS: BUILDING TRUST AND INTEGRATION INTO HEALTH CARE SYSTEMS

One of the primary concerns surrounding the implementation of AI-aided imaging tools in cancer care is the need to build trust among health care professionals, patients, and regulatory bodies. Clinicians, who have relied on their own visual assessments and expertise for years, may be hesitant to trust complex algorithms that they do not fully understand and cannot easily explain to their patients. To address this, it is crucial that AI developers work closely with oncologists, radiologists, nuclear medicine specialists, and other key stakeholders to ensure the robustness, accuracy, and reliability of the imaging biomarkers produced by these AI systems.[45] In addition, hospital-based physicists and biomedical engineers play a pivotal role in bridging this gap between AI tools and clinical practice (Figure 4).

Explainability is also a key factor in enhancing trust and facilitating the integration of AI-aided imaging biomarkers. The interpretability of AI algorithms, particularly in the context of deep learning models with complex, multilayered architectures, is a significant challenge. Clinicians and patients require a clear understanding of how these AI systems arrive at their conclusions, as this transparency is essential for building confidence and fostering a collaborative decision-making process. One of the most promising tools for deep learning explainability is attention maps, which provide color maps explaining how much the model has used each part of the image to make the decision. Along with informative graphs about the certainties and uncertainties of the predictions, these are essential tools to achieve a transparent and explainable AI reports integrated into clinical practice. Other approaches, such as Shapley Additive Explanations values, provide individual feature importance scores, while biological explanations aim to correlate AI predictions with tissue-based markers. Importantly, all AI tools must be subjected to strict controls to ensure their reliability. Technical validations from independent experts are necessary to identify potential pitfalls in the pipeline and the explainability methods. AI experts can provide insights into efficient code implementation, while clinicians can give hands-on feedback to further improve the tools. In addition, clinical validation is also needed, sometimes in the form of clinical trial testing, to ensure that

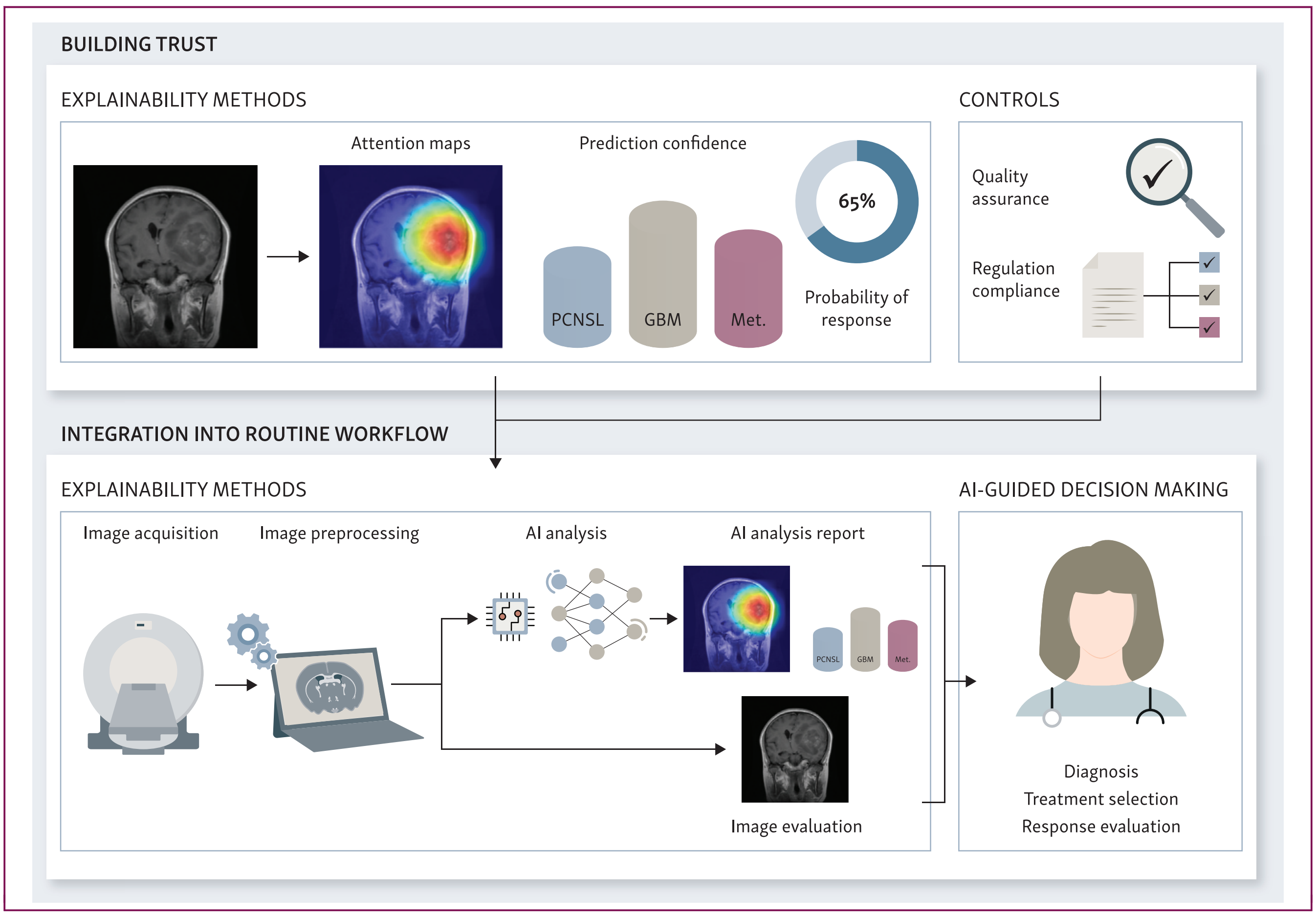

Figure 4. Building trust and integration of radiomic-based biomarkers into health care systems. Process of integrating radiomic-based biomarkers into health care systems, emphasizing building trust and routine workflow integration. The top section focuses on building trust through explainability methods and robust controls, while the bottom section describes the integration of the methods into the routine workflow.
AI, artificial intelligence; GBM, glioblastoma multiforme; Met., metastasis; PCNSL, primary central nervous system lymphoma; Prob., probability.

the models and tools work properly on the target populations. Finally, there must be regulatory controls to ensure that the technologies used follow current and future international and internal AI regulations. As the scientific community continues to explore the potential of AI-aided tools, it will be crucial to address these considerations to ensure the successful and ethical integration of these technologies into real-world workflows in clinical practice.

Moreover, integrating AI-aided imaging biomarkers into the radiology department workflow requires a strategic approach to ensure seamless adoption. This strategy includes training radiologists and support staff on the use and interpretation of AI-generated data and incorporating the tools into existing imaging software platforms. It also requires the development of robust and seamless interfaces between the AI tools and the various imaging and information systems used in the department, integration of the AI algorithms into existing picture archiving and communication systems (PACS) and radiology information systems (RIS). AI tools should not create additional tasks for physicians but should instead simplify and enhance their daily routines, as if these tools are not seamlessly integrated and easy to use, they will never truly be implemented in clinical practice.

SOURCES OF REAL-WORLD RADIOLOGY DATA FOR AI-AIDED TOOL DEVELOPMENT AND BIOMARKER DISCOVERY

Real-world image data represent an invaluable resource for the development and validation of computer-aided support systems in oncology. Derived from routine clinical practice, these datasets contribute to advancements in tumor detection and the segmentation of organs and adjacent structures, which are essential for radiotherapy planning. Additionally, these data support the development of tools for identifying imaging biomarkers, aiding potentially therapeutic decisions, and improving cancer patient management.

Some notable sources of real-world medical image data include:

- Hospital PACS: Hospitals maintain extensive PACS that store a wide variety of imaging data, including CT, MRI, and PET scans. These systems capture images from routine clinical workflows and provide a rich source of real-world imaging data. This information can also be linked and accessible through the EHR systems, providing a holistic view of patient health.
- National and international cancer registries (Table 2): These registries, created by governmental and international institutions, collect and curate large volumes of cancer-related data, including diagnostic imaging, treatment regimens, and patient outcomes collected as part of routine clinical care.
- Collaborative research networks (Table 2): Initiatives that provide publicly accessible, well-annotated imaging datasets, often linked to clinical data. These resources facilitate collaborative research and the validation of radiomics-based biomarkers.
- Datathon dataset portals: Beyond cancer-specific data portals, some generic dataset hosts and computational challenge portals include datasets of cancer imaging, such as:
- Kaggle
- Medical Image Datathon
- Grand Challenge
- The Medical Image Computing and Computer Assisted Intervention Society (MICCAI) registered challenges
- Google Dataset Search
- Dataportal Asia

DISCUSSION

AI-aided tool development typically requires large, diverse datasets representative of the populations in which they will be applied, yet the scarcity of extensive multicentric databases remains a major limitation for their development and true applicability in clinical scenarios. The integration of AI with RWD offers a significant potential for computer-aided tools and biomarker development, including uses in automatic organ delineation, tumor detection, and personalized decision making. RWD offer diverse, extensive data reflecting real-world clinical practices, complementing clinical trial data and providing a broader understanding of patient populations. Nevertheless, although the use of RWD presents significant advantages, it also poses challenges. This review highlights the immense potential of AI to analyze vast datasets of medical imaging and radiology reports, extracting relevant data and enhancing biomarker discovery while addressing as well the challenges associated with using RWD and providing suggestions for minimizing these hurdles.

Real-world image data, including diverse medical imaging modalities acquired across various clinical settings, represent a rich source of information for computer-aided tools and biomarker discovery. These datasets reflect the complexities and variabilities of actual clinical practice, providing a more comprehensive understanding of patient characteristics, treatment patterns, and outcomes. This is crucial for developing assays that are generalizable and applicable across different populations and clinical settings.

Despite the advantages, the use of RWD introduces several methodological challenges. Data quality is a primary concern, as variability in equipment, acquisition protocols, and operator expertise can lead to inconsistencies that affect the accuracy and reliability of radiomic features. Standardization of image acquisition, processing, and feature extraction is critical to ensure reproducibility and clinical applicability. Additionally, common artifacts, such as motion artifacts in MRI or beam-hardening artifacts in CT, can further impact feature extraction. Effective noise reduction techniques are also essential to enhance the signal-to-noise ratio and improve the reliability of extracted features. Initiatives like the QIBA, the EIBALL, and the IBSI provide guidelines for harmonizing these processes.[40-43] Adhering to standardized imaging protocols and using

Table 2. National and international cancer registries and collaborative research networks

Platform	Data type	Link	Description	# Cases	Data availability
National Cancer Institute (NCI)	CT, MRI, pathology	https://datascience.cancer.gov/resources/nci-data-catalog	US Cancer Institute with several data collections available to download beyond imaging		Public
The Cancer Genome Atlas (TCGA)	Pathology	https://portal.gdc.cancer.gov/	NCI cancer genomics data collection hosting several datasets, some linked to TCIA datasets with pathology imaging data available	35 982	Public
Lung Imaging Database Consortium (LIDC)	CT	https://imaging.cancer.gov/informatics/lidc_idri.htm	NCI initiative to support developing consensus guidelines for a spiral CT lung image resource	1010	Public
Clinical Proteomic Tumor Analysis Consortium (CPTAC)	CT, MRI, pathology	https://proteomics.cancer.gov/programs/cptac	NCI initiative to accelerate the understanding of the molecular basis of cancer through the application of large-scale proteome and genome analysis, or proteogenomics	1790	Public
Human Tumor Atlas Network (HTAN)	CT, MRI, pathology	https://humantumoratlas.org/	NCI-funded initiative to construct three-dimensional atlases of the dynamic cellular, morphological, and molecular features of human cancers	2147	Public
National Lung Screening Trial (NLST)	CT, pathology	https://cdas.cancer.gov/nlst/	NCI trial to determine whether screening for lung cancer with low-dose helical CT reduces mortality from lung cancer in high-risk individuals relative to screening with chest radiography	26 254 CT 451 Pathology	Under request
American College of Radiology Imaging Network (ACRIN)	CT, MRI, pathology	https://www.acr.org/Research/Clinical-Research/ACRIN-Legacy-Trials	NCI cooperative group devoted to conducting trials directed toward evaluating the applications of diagnostic imaging and image-guided treatment to cancer	1222	Public
Quantitative Imaging Network (QIN)	CT, MRI, pathology	https://imaging.cancer.gov/programs_resources/specialized_initiatives/qin/about/default.htm	NCI network for development and clinical validation of quantitative imaging tools and methods for the measurement or prediction of tumor response to therapies in clinical trial settings	571	Public
Cancer Moonshot Biobank (CMB)	CT, MRI, pathology	https://moonshotbiobank.cancer.gov/	NCI initiative to support investigations into drug resistance and sensitivity, collecting data from voluntary participants who donate samples during treatment	370	Public—clinical data under request
Patient-Derived Models Repository (PDMR)	MRI, pathology	https://pdmr.cancer.gov/	NCI repository of clinically annotated patient-derived pre-clinical models to support public—private partnerships and academic drug discovery through an accessible database	198	Public
National Clinical Imaging Research Registry (ANCIRR)	CT, MRI	https://www.acr.org/Research/Clinical-Research/National-Clinical-Imaging-Research-Registry	American College of Radiology collection of registries that focus on collecting and curating imaging data from registered partners	Unreported	Registered partners
UK Biobank	MRI	https://www.ukbiobank.ac.uk/	Large-scale British biomedical database containing genetic, health, and imaging data from patients treated in the UK public health care system	2853	Registered partners
ACR National Clinical Imaging Research Registry	CT, MRI	https://www.acr.org/Research/Clinical-Research/National-Clinical-Imaging-Research-Registry	US collection of radiology registries, some of them of cancer patients	Unreported	Registered partners
The Cancer Imaging Archive (TCIA) and Cancer Research Data Commons (CRDC)	CT, MRI, pathology	TCIA: https://www.cancerimagingarchive.net/ CRDC: https://datacommons.cancer.gov/explore	US services that centralize medical imaging data from several US institutions for their query and retrieval	54 951	Public
WORC database	CT, MRI	https://github.com/MStarmans91/WORCDatabase	Dataset from the paper Starmans et al.[46]	930	Public
EuCanImage (EUCAIM)	—	https://cancerimage.eu/	Pan-European infrastructure to host cancer images for research. Still under development	—	—

consistent algorithms for feature extraction (including both the use of handcrafted and foundation model embeddings) can mitigate variability and enhance the robustness of AI-aided imaging-based tools and radiomic biomarkers.

The clinical validation of AI-enabled biomarkers is another critical challenge. Using independent and diverse RWD sources for validation is key to confirming the generalizability and reliability of AI models. Beyond this, building trust among clinicians, patients, and regulatory bodies is essential for the successful implementation of AI-aided imaging biomarkers. In this context, the explainability of AI models is crucial, as clinicians need to understand how these systems arrive at their conclusions. Developing transparent and explainable AI algorithms can facilitate this understanding, fostering confidence and acceptance among health care professionals. Moreover, integrating AI-aided biomarkers into routine clinical workflows requires careful planning and collaboration with health care professionals. Training radiologists and support staff on the use and interpretation of AI-generated data, incorporating AI tools into existing imaging software platforms, and establishing protocols for review and validation are essential steps. Ensuring seamless integration with existing systems, such as PACS and RIS, can help streamline the adoption process and minimize additional workload for clinicians.

In conclusion, the transformative potential of AI and RWD in radiology is clear, but addressing challenges such as data standardization and quality checks to an appropriate extent, validation, trustworthiness, and ease of use is critical for translating this potential into clinical practice.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the members of the Cancer Core Europe (CCE) Imaging Task Force for their valuable contributions to the discussions and collaborative efforts that have informed the development of this review.

FUNDING

None declared.

DISCLOSURE

The authors have declared no conflicts of interest.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

ChatGPT was used solely for grammar and language editing in this manuscript. No AI tools were used to generate or modify the scientific content.

REFERENCES

1. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18(8):500-510.
2. Rajpurkar P, Lungren MP. The current and future state of AI interpretation of medical images. *N Engl J Med*. 2023;388(21):1981-1990.
3. Frazer HML, Peña-Solorzano CA, Kwok CF, et al. Comparison of AI-integrated pathways with human-AI interaction in population mammographic screening for breast cancer. *Nat Commun*. 2024;15(1):7525.
4. Ng AY, Oberije CJG, Ambrózay É, et al. Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. *Nat Med*. 2023;29(12):3044-3049.
5. Gilbert FJ, Astley SM, Gillan MG, et al. Single reading with computer-aided detection for screening mammography. *N Engl J Med*. 2008;359(16):1675-1684.
6. Shi F, Hu W, Wu J, et al. Deep learning empowered volume delineation of whole-body organs-at-risk for accelerated radiotherapy. *Nat Commun*. 2022;13(1):6566.
7. Bera K, Braman N, Gupta A, Velcheti V, Madabhushi A. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol*. 2022;19(2):132-146.
8. Garcia-Ruiz A, Pons-Escoda A, Grussu F, et al. An accessible deep learning tool for voxel-wise classification of brain malignancies from perfusion MRI. *Cell Rep Med*. 2024;5(3):101464.
9. Astaraki M, Yang G, Zakko Y, Toma-Dasu I, Smedby Ö, Wang C. A comparative study of radiomics and deep-learning based methods for pulmonary nodule malignancy prediction in low dose CT images. *Front Oncol*. 2021;11:737368.
10. Demircioğlu A. Are deep models in radiomics performing better than generic models? A systematic review. *Eur Radiol Exp*. 2023;7(1):11.
11. Van Veen D, Van Uden C, Blankemeier L, et al. Adapted large language models can outperform medical experts in clinical text summarization. *Nat Med*. 2024;30(4):1134-1142.
12. Zhang S, Xu Y, Usuyama N, et al. BiomedCLIP: a multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs. *arXiv [cs.CV]*. 2023. https://doi.org/10.48550/arXiv.2303.00915.
13. Blankemeier L, Cohen JP, Kumar A, et al. Merlin: a vision language foundation model for 3D computed tomography. *arXiv [cs.CV]*. 2024. https://doi.org/10.48550/arXiv.2406.06512.
14. Hamamci IE, Er S, Almas F, et al. Developing generalist foundation models from a multimodal dataset for 3D computed tomography. *arXiv [cs.CV]*. 2024. https://doi.org/10.48550/arXiv.2403.17834.
15. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48(4):441-446.
16. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* 2016;278(2):563-577.
17. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *arXiv [cs.CV]*. 2015. https://doi.org/10.48550/arXiv.1505.04597.
18. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *arXiv [cs.CV]*. 2015. https://doi.org/10.48550/arXiv.1512.03385.
19. Vanguri RS, Luo J, Aukerman AT, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer.* 2022;3(10):1151-1164.
20. Center for Devices, Radiological Health. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. U.S. Food and Drug Administration. FDA. 2024. Available at https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices. Accessed February 1, 2025.
21. Zhang S, Metaxas D. On the challenges and perspectives of foundation models for medical image analysis. *Med Image Anal*. 2024;91:102996.
22. Awais M, Naseer M, Khan S, et al. Foundation models defining a new era in vision: a survey and outlook. *IEEE Trans Pattern Anal Mach Intell.* 2025;47(4):2245-2264.
23. Ramesh A, Pavlov M, Goh G, et al. Zero-shot text-to-image generation. *arXiv [cs.CV]*. 2021. https://doi.org/10.48550/arXiv.2102.12092.
24. Bommasani R, Hudson DA, Adeli E, et al. On the opportunities and risks of foundation models. *arXiv [cs.LG]*. 2021. https://doi.org/10.48550/arXiv.2108.07258.
25. AlSaad R, Abd-Alrazaq A, Boughorbel S, et al. Multimodal large language models in health care: applications, challenges, and future outlook. *J Med Internet Res*. 2024;26:e59505.
26. Baumfeld Andre E, Carrington N, Siami FS, et al. The current landscape and emerging applications for real-world data in diagnostics and clinical decision support and its impact on regulatory decision making. *Clin Pharmacol Ther.* 2022;112(6):1172-1182.

27. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W, on behalf of GetReal Work Package 1. What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health*. 2017;20(7):858-865.
28. Togo K, Yonemoto N. Real world data and data science in medical research: present and future. *Jpn J Stat Data Sci*. 2022;5(2):769-781.
29. Liu R, Wang L, Rizzo S, et al. Systematic analysis of off-label and off-guideline cancer therapy usage in a real-world cohort of 165,912 US patients. *Cell Rep Med*. 2024;5(3):101444.
30. Perez-Lopez R, Ghaffari Laleh N, Mahmood F, Kather JN. A guide to artificial intelligence for cancer researchers. *Nat Rev Cancer*. 2024;24(6):427-441.
31. Naidoo P, Bouharati C, Rambiritch V, et al. Real-world evidence and product development: opportunities, challenges and risk mitigation. *Wien Klin Wochenschr*. 2021;133(15-16):840-846.
32. Schwenck J, Sonanini D, Seyfried D, et al. In vivo imaging of $CD8^{+}$ T cells in metastatic cancer patients: first clinical experience with simultaneous [^{89}Zr]Zr-Df-IAB22M2C PET/MRI. *Theranostics*. 2023;13(8):2408-2423.
33. Lau D, McLean MA, Priest AN, et al. Multiparametric MRI of early tumor response to immune checkpoint blockade in metastatic melanoma. *J Immunother Cancer*. 2021;9(9):e003125.
34. Garcia-Ruiz A, Macarro C, Zacchi F, et al. Whole-body magnetic resonance imaging as a treatment response biomarker in castration-resistant prostate cancer with bone metastases: the iPROMET clinical trial. *Eur Urol*. 2024;86(3):272-274.
35. Truhn D, Tayebi Arasteh S, Saldanha OL, et al. Encrypted federated learning for secure decentralized collaboration in cancer image analysis. *Med Image Anal*. 2024;92(103059):103059.
36. Teo ZL, Jin L, Li S, et al. Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. *Cell Rep Med*. 2024;5(2):101419.
37. Denner S, Zimmerer D, Bounias D, et al. Leveraging foundation models for content-based medical image retrieval in radiology. *arXiv [cs.CV]*. 2024. https://doi.org/10.48550/arXiv.2403.06567.
38. Hagiwara A, Fujita S, Ohno Y, Aoki S. Variability and standardization of quantitative imaging: monoparametric to multiparametric quantification, radiomics, and artificial intelligence. *Invest Radiol*. 2020;55(9):601-616.
39. Ligero M, Jordi-Ollero O, Bernatowicz K, et al. Minimizing acquisition-related radiomics variability by image resampling and batch effect correction to allow for large-scale data analysis. *Eur Radiol*. 2021;31(3):1460-1470.
40. Shukla-Dave A, Obuchowski NA, Chenevert TL, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging*. 2019;49(7):e101-e121.
41. deSouza NM, Achten E, Alberich-Bayarri A, et al. Validated imaging biomarkers as decision-making tools in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European Society of Radiology (ESR). *Insights Imaging*. 2019;10(1):87.
42. Whybra P, Zwanenburg A, Andrearczyk V, et al. The Image Biomarker Standardization Initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights. *Radiology*. 2024;310(2):e231319.
43. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295(2):328-338.
44. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77(21):e104-e107.
45. Sanchez M, Alford K, Krishna V, Huynh TM, et al. AI-clinician collaboration via disagreement prediction: a decision pipeline and retrospective analysis of real-world radiologist-AI interactions. *Cell Rep Med*. 2023;4(10):101207.
46. Starmans MPA, van der Voort SR, Phil T, et al. Reproducible radiomics through automated machine learning validated on twelve clinical applications. *arXiv [eess.IV]*. 2022. https://doi.org/10.48550/arXiv.2108.08618.