



AI adoption and workflow optimization following orchestration platform implementation and structured change management



Arsany Hakim¹✉, Christian Rummel^{1,2}, Christine Jacob³, Piotr Radojewski¹ & Roland Wiest^{1,4}

Artificial intelligence (AI) adoption in radiology remains limited by workflow integration challenges, lack of interoperability, and barriers to user engagement. We implemented a dual-track socio-technical strategy in a tertiary-care neuroradiology department, combining a vendor-neutral AI orchestration platform with a structured change management program, including user education, workflow standardization, documentation, feedback mechanisms, and governance. In an observational pre–post implementation study, AI adoption was evaluated by comparing a baseline period using a single-vendor solution (September–November 2022) with a post-implementation period following platform deployment (May–July 2025). Adoption was measured using the user-verification rate, defined as the proportion of AI outputs actively reviewed and either accepted or rejected by radiologists. The overall user-verification rate increased from 13.9% (318/2287) at baseline to 61.1% (1778/2910) after implementation, while full verification of examinations containing multiple AI tasks increased to 62.6%. Adoption varied among individual users, and trainees demonstrated higher verification rates when supervised by consultants with high AI adoption. During the post-implementation period, the platform processed 9480 AI tasks across 12 applications, demonstrating scalability across diverse clinical workflows. In conclusion, a combined orchestration platform and change management strategy was associated with substantially increased AI adoption and integration into routine neuroradiology workflow.

Radiology is undergoing a profound transformation with the introduction of AI, especially machine-learning and deep-learning models¹. These technologies augment imaging workflows. A growing number of commercial solutions are already available, offering functionalities such as triage of acute findings, assisted detection, and extraction of quantitative markers and imaging patterns in selected clinical settings relevant to diagnosis and management^{2,3}.

Despite this promise, several barriers limit the integration and adoption of AI into routine practice. One major barrier is technical integration, as AI tools may face resistance when not smoothly embedded into existing infrastructure, where interoperability with surrounding clinical systems is essential^{4,5}. Challenges extend beyond technical integration to include embedding AI into user workflows, limited training and awareness, and a

lack of structured frameworks for evaluating and incorporating AI outputs into reporting⁶ all of which hinder large-scale adoption and evidence generation.

These challenges are further compounded by a rapidly expanding AI marketplace, with a growing number of FDA-cleared and CE-marked solutions from diverse vendors. In this context, reliance on a single-vendor system is unlikely to be sustainable clinically (heterogeneous indications and evolving use cases) or operationally (multiple point-to-point interfaces, redundant integrations with the same systems, and fragmented workflows), underscoring the importance of vendor-neutral AI platforms. In principle, such platforms aim to consolidate technical and user-facing processes into a unified environment. Beyond integration, they also enable orchestration,

¹University Institute of Diagnostic and Interventional Neuroradiology, Bern University Hospital, Inselspital, University of Bern, Bern, Switzerland. ²Technische Hochschule Deggendorf, European Campus Rottal-Inn, Pfarrkirchen, Germany. ³University of Applied Sciences Northwestern Switzerland FHNW, Windisch, Switzerland. ⁴Department of Digital Medicine, University of Bern, Bern, Switzerland. ✉e-mail: arsany.hakim@insel.ch

Table 1 | Structured change management

User education
-Intensive training sessions on the platform and individual AI applications
-Refresher courses for the new and existing staff
-Regular scientific meetings on AI use, the installed AI applications, and clinical value
Standardization
-Structured workflow to streamline platform use
-Quality control procedures for AI-generated outputs
-Verification pathways adapted to application risk level (radiologist-in-the loop vs. on-the-loop)
-Standardized terminology and reporting for AI outputs
Feedback and support
-Dedicated internal expert support team for rollout and troubleshooting
-Continuous feedback loops addressing technical performance, output quality, and clinical utility
Documentation
-Central access to vendor-provided instructions and resources
-AI Playbook: consolidated reference including app-specific requirements, clinical indication, evaluation steps, and reporting guidance.
Governance and communication
-Clear definition of roles and responsibilities for evaluation, clinical integration and monitoring
-Regular coordination with vendors and hospital IT to ensure compliance and optimal integration
-Regular communication with users to provide updates and share best clinical practices

i.e., automated routing and processing of imaging studies to the appropriate AI applications, reducing manual steps and supporting workflows.

In parallel, radiology departments face mounting workloads, driven by the increasing volume and complexity of imaging examinations, alongside static or insufficient workforce growth⁷. Workflow streamlining through selective automation of non-interpretative tasks has therefore become indispensable for maintaining efficiency⁸. This optimization may contribute to system-wide resource allocation, helping address specialist shortages and improving care delivery at scale⁸.

These considerations highlight the need for a dual approach combining technical infrastructure and structured change-management, aligned with socio-technical theory^{9,10}.

We describe the implementation of this approach within an early adopter program of an orchestration platform at an academic center^{11,12}. Specifically, we report on the integration of a vendor-neutral AI orchestration platform, paired with a structured change-management plan (Table 1) for its integration in the clinical workflow, and evaluate its association with workflow optimization and AI adoption in a real-world implementation setting. We conducted an observational pre–post implementation study, comparing adoption rates during a baseline period with a single-vendor solution before platform integration with those observed after full deployment.

Results

Platform performance

During the study period (May–July 2025), a total of 9480 tasks from cross-sectional imaging were processed across 12 integrated AI applications (Table 2), demonstrating the scalability of the platform to handle large examination volumes in diverse clinical use cases. Furthermore, the same examination could be processed simultaneously by multiple applications, underscoring the platform’s capabilities for interoperability across apps (Supplementary Fig. 1). Data were automatically and temporarily transferred from PACS or modalities to the platform only for the duration of processing, ensuring that AI results were available in PACS/platform viewer at the time of reporting. For applications supporting longitudinal analysis, an automated query-retrieve mechanism enabled the retrieval and processing of prior studies.

Adoption

In the pre-implementation period (September–November 2022), 5142 MRI studies were performed, of which 2192 (42.6%) included at least one applicable AI task from the three AI applications available in both the pre- and post-implementation periods (aneurysm assessment, volumetry, and lesion analysis). Because individual exams could contain multiple tasks, this

Table 2 | Number of AI processing runs per application between May and July 2025

App	Modality	Processed tasks (May–July 2025)	Mean turnaround time per task (standard deviation)
CINA-ICH	CT	2789	02:34 (00:24)
CINA-LVO	CT	912	05:12 (01:19)
CINA-ASPECTS	CT	424	05:21 (01:11)
icobrain dm	MRI, CT	257	19:43 (02:25)
icobrain ms	MRI	536	32:41 (16:59)
icobrain tbi	CT, MRI	104	32:41 (16:59)
IB DELTA SUITE	MRI	1234	08:55 (01:18)
IB NEURO	MRI	284	11:43 (01:23)
IB DCE	MRI	30	32:18 (09:44)
mdbrain Aneurysm	MRI	2141	07:40 (03:10)
mdbrain Volumetry	MRI	182	10:07 (03:07)
mdbrain Lesion	MRI	587	17:04 (05:55)
Total	MRI, CT	9480	08:50 (09:19)

The table also reports the mean turnaround time per task (±standard deviation) until results were available in the viewer, encompassing exam reception, dormancy, transfer to cloud, processing, PACS query, retrieval of prior exams, and return to hospital network.

corresponded to a total of 2287 applicable tasks. Of these, 318 tasks (13.9%) were user-verified with AI outputs in PACS.

In the post-implementation period (May–July 2025), 6,039 MRI studies were performed (17.4% increase). Of these, 2688 (44.3%) contained applicable AI tasks from the same applications included in the pre-implementation analysis ($n = 2910$ tasks). Among these, 1778 tasks were user-verified, corresponding to an overall verification rate of 61.1% (95% CI: 59.3–62.9%), compared with 13.9% (95% CI: 12.5–15.4%) in 2022 ($p < 0.001$) representing an approximately four-fold increase in user-verification rate, our operational measure of adoption, compared to baseline (Fig. 1; Supplementary Table 1).

User-verification improved across all use cases (Fig. 1; supplementary Table 1) with the largest gains in applications for aneurysm and MS mimics.

Among the three applications available in both the pre- and post-implementation periods, the number of examinations with multiple AI-

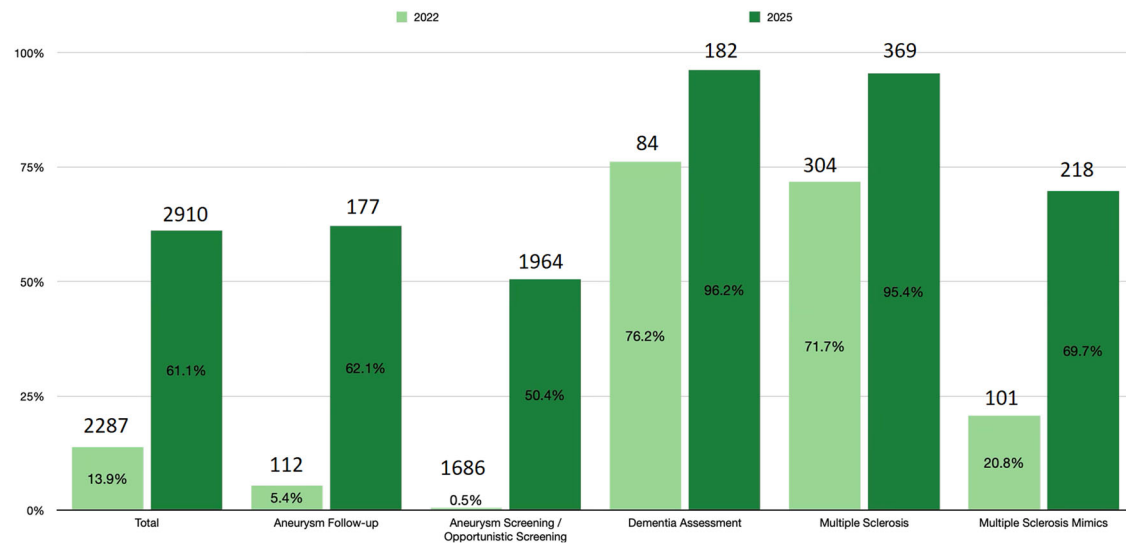
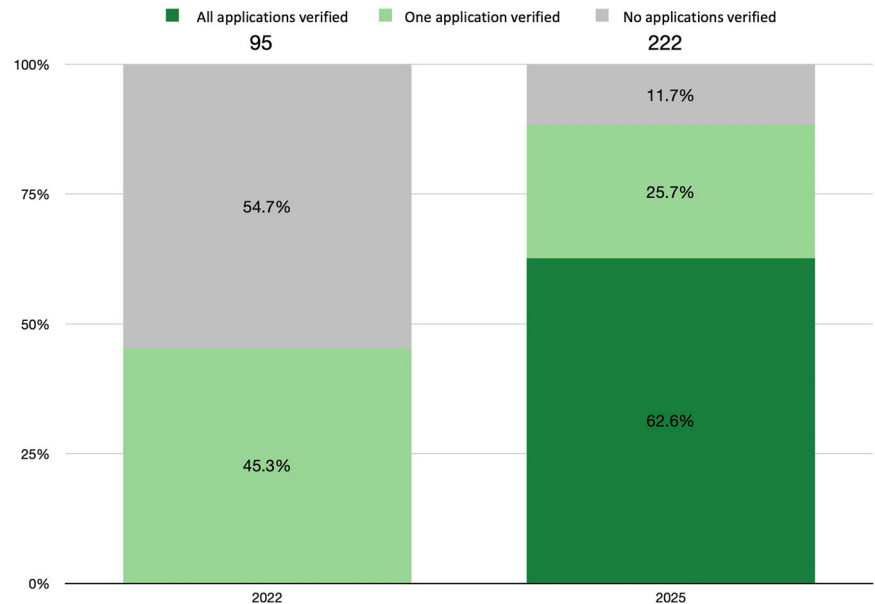


Fig. 1 | Verification rates of AI results in 2022 (before the platform implementation and change management, light green) and 2025 (after the implementation, dark green) across overall tasks and specific clinical use cases. The number of AI-applicable cases per task category is indicated above the bars.

Fig. 2 | Distribution of verification in multi-task MRI exams for 2022 versus 2025. Bars show the proportion of exams in which all one or no applicable AI tasks from the same exam were verified. The number of multitask exams is indicated above the bars.



applicable tasks per exam increased from 95 (2022) to 222 (2025). The proportion of examinations with full verification, i.e., all available AI outputs were verified, increased to 62.6% (139/222) (Fig. 2; Supplementary Table 1).

The weekly verification rate in 2025 started at 53.2% in week 1 and peaked at 71.2% in the final week, representing an increase of ~18 percentage points. Weekly verification rates varied during the post-implementation period, with a numerically increasing pattern that did not reach statistical significance ($p = 0.062$) (Supplementary Fig. 2).

Of 19 consultants active in 2025, 16 authored reports independently (without trainees) for at least a subset of exams with applicable AI tasks (Supplementary Table 2). Adoption patterns when reporting alone: 4 consultants were categorized as minimal adopters (<25% verification), 5 as intermediate adopters (25–75%), 5 as high adopters (75–90%), and 2 as very-high adopters (>90%). Among the 11 consultants with baseline data in 2022, 9 showed increased verification rates by 2025 (Supplementary Fig. 3 and Supplementary Table 3).

Of the four consultants categorized as minimal adopters, two had limited access to the scheduled training activities, and three missed the

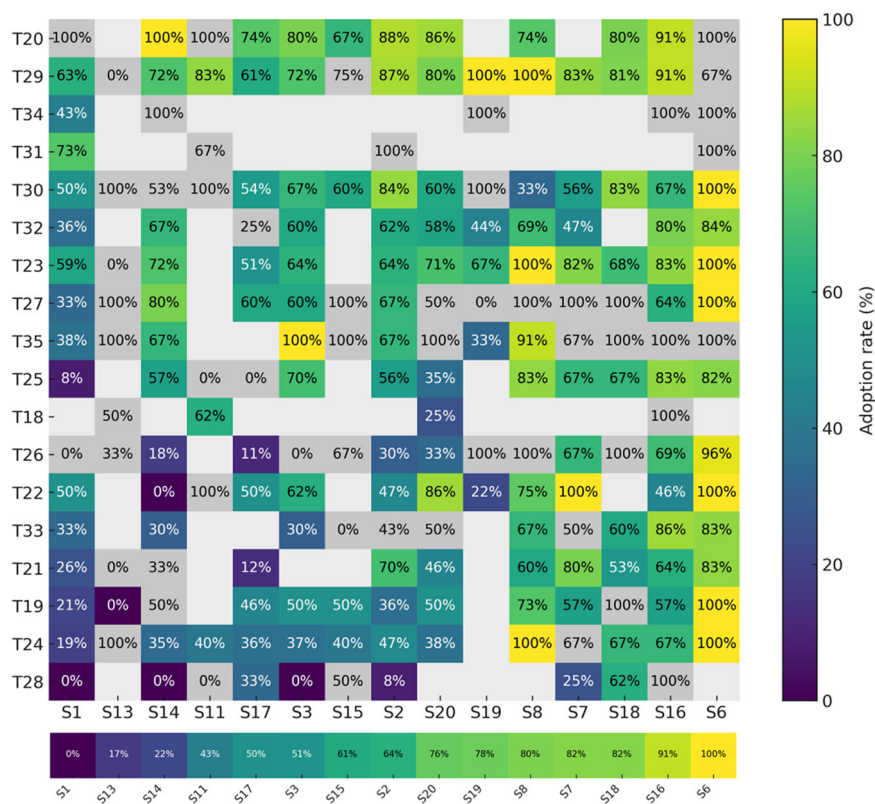
subsequent refresher courses. Two consultants in the minimal-adopters group contributed only occasionally to diagnostic reporting, limiting their integration into the daily reporting workflow with AI user-verification.

In 2025, 18 Trainees were actively reporting (Supplementary Table 2), without baseline data in 2022. Reports finalized under the supervision of very-high-adopter consultants consistently demonstrated higher user-verification rates than those supervised by minimal adopters (Supplementary Figs. 4 and 5).

Exploratory subgroup analyses demonstrated no statistically significant differences in overall verification rates according to training status, departmental duration, or AI-task exposure volume (Supplementary Table 4).

Consultant-trainee pairing revealed a distinct pattern: user-verification rates were consistently highest when reporting with very-high adopters, whereas minimal adopters were associated with low user-verification across nearly all trainees (Fig. 3). Notably, trainees displayed markedly different user-verification behavior depending on the supervising consultant,

Fig. 3 | Consultant-trainee dyad adoption heatmap. Each cell shows the adoption (verification) rate for that pairing, with color encoding the rate. Gray cells mark dyads with fewer than three applicable tasks (displayed for completeness but not colored in heatmap), and blank cells indicate no joint cases. The x-axis (consultants) is ordered by their independent adoption rate (without trainees), shown in the bottom strip. Higher rates cluster towards columns with higher independent adoption. *Note:* This visualization contains only the 15 consultants who supervised trainees in the study period.



highlighting the strong influence of consultant adoption style on trainees' practice.

Discussion

We present a dual-track socio-technical implementation strategy that combined technical integration with structured change management to support the adoption of AI tools and their integration in neuroradiology workflows within an early adopter program. The technical track involved the seamless integration of a vendor-neutral platform into existing infrastructure, enabling automation and orchestration of multiple applications both across different exams and within individual exams. The human track, a key element in our workflow, encompasses structured change management measures including user education, standardization, documentation, feedback, and governance¹³. The dual-track structure reflects a socio-technical understanding that technical, social, and organizational elements are deeply intertwined, mutually shaping one another, and that successful AI adoption requires moving beyond a technology-centric focus to fully account for the clinical context in which these systems operate, as emphasized in several contemporary implementation frameworks¹³. Collectively, this integrated approach was associated with a substantial increase in adoption, with overall verification rates rising more than four-fold and full verification achieved in more than half of multi-task examinations. Beyond individual AI solutions, this approach contributed to the development of a local AI ecosystem, in which technical infrastructure, standardized workflow, user training, and governance are aligned to support sustainable integration of AI into clinical practice. Importantly, adoption increased not only in previously underutilized applications, such as aneurysm screening, but also in applications that were already well adopted, such as assessment of brain atrophy and MS lesion segmentation. The observed increase in adoption likely reflects a combination of interacting factors, including technical integration, structured change management, and familiarity with AI within the clinical workflow. Automated routing of pseudo-anonymized data, centralized access, removal of separate logins, and timely result availability likely reduced workflow friction, whereas repeated exposure,

refresher training, internal case discussions, and visible clinical utility may have progressively strengthened user confidence and routine use over time. Although increasing familiarity with AI over time may have contributed, this explanation is unlikely to fully account for the findings, as the comparative analysis involved the same core applications, exposure during the transition period was not continuous, and all the trainees/fellows in the post-implementation period had no prior local AI experience. In exploratory subgroup analysis, no statistically significant differences were observed between groups, suggesting that adoption was not restricted to a single user subgroup and supports broad departmental integration. The upward weekly trend during the post-implementation period further supports that adoption was not static after rollout, but continued to evolve during routine clinical use. Given the real-world implementation setting, the individual contributions of technical and human factors cannot be isolated, as these components are inherently interdependent and co-evolve within socio-technical systems^{12–14}. The change-management components were selected pragmatically to address local barriers, including awareness, trust, workflow friction, and governance needs. Their effects were likely complementary rather than independent, and the most influential element may differ across users and institutional settings. Individual users likely respond to different implementation factors in distinct ways, suggesting that multimodal strategies may outperform single-factor approaches.

Our findings align with emerging literature indicating that successful clinical AI implementation depends not only on algorithm performance, but also on workflow integration, interoperability, governance, and user engagement^{9,14–17}. Prior reports have emphasized the importance of vendor-neutral or standards-based infrastructures to reduce fragmented point-to-point integrations and facilitate scalable deployment of multiple tools^{18–21}. Other implementation studies have highlighted barriers including technical complexity, limited training, workflow disruption, and uncertainty regarding governance^{14–16,22}. Our results extend this literature by demonstrating that, in a real-world setting, combining platform infrastructure with structured change management was associated with substantially higher routine AI adoption across multiple use cases^{14,23,24}.

However, adoption behavior was heterogeneous among users. There was an increase in adoption across almost every individual consultant between pre- and post-implementation. Four consultants were categorized as minimal adopters (<25% verification). Reduced adoption in this group was associated with either limited exposure to structured training or a primary clinical focus outside routine diagnostic reporting, underscoring the importance of inclusive team-building and targeted education to ensure all colleagues are engaged. Conversely, one consultant achieved near-complete adoption in consultant-only readings, illustrating that full integration of AI into reporting practice is feasible.

At the trainee level, verification behavior was strongly influenced by the supervising consultant. Trainees reporting with very-high adopters showed substantially higher verification rates compared to those supervised by minimal adopters, indicating that consultant behavior serves as a role model shaping trainee practice. Previous surveys and literature review identified limited mentorship and expert guidance as barriers to AI learning and appropriate clinical use. Al Mohammad et al. reported that lack of mentorship and guidance was the most frequently cited barrier to learning about AI²⁵, while Volin et al. emphasized the need for structured supervision by experienced users to support appropriate AI use among trainees²⁶. Our findings extend this evidence by showing that trainee verification behavior varied according to the supervising consultant's adoption pattern. This emphasizes the need to invest not only in trainee education but also in preparing consultants to model effective AI use, thereby embedding new technologies into departmental culture. Interestingly, two relatively new trainees in the department demonstrated a higher user-verification rate despite working frequently with minimal adopters, suggesting that, for some individuals, the simultaneous process of learning departmental workflows and AI use may facilitate adoption, though this was not consistent across all trainees. Importantly, user verification in our setting was not mandated by institutional policy, but was encouraged as part of quality-oriented reporting practices. No incentives or penalties were linked to tool usage. The observed variability in adoption across users, ranging from minimal to very high, further supports that verification reflects voluntary engagement rather than compliance with a predefined workflow requirement.

Beyond adoption, the results highlight the scalability of orchestration platforms: multiple applications can be seamlessly integrated both within and across examinations, supporting an expanding number and diversity of AI tools while maintaining workflow efficiency. Compared to single-application solutions, the platform enabled automation of multiple processes within and across exams. Other advantages of the platform include a unified interface for all applications, standardized verification and feedback mechanisms, and on-demand application testing.

Orchestration and automation reduce manual processing time, prevent workflow fragmentation, and improve scalability as imaging volumes grow. Such features help radiologists focus on interpretive and decision-making activities, thereby supporting improved departmental efficiency, patient throughput, and timeliness of clinical decision-making.

Several processes were essential for the change management strategy in the context of the observed increase in adoption. Standardized acquisition protocols established prior to implementation further facilitated automation and orchestration. A dedicated support team appointed from departmental staff provided trusted guidance and troubleshooting during rollout. Documentation was streamlined into an AI playbook, summarizing vendor instructions and local procedures in a user-friendly format embedded within the departmental playbook. In addition, internal case discussions and shared examples of diagnostic relevance fostered engagement and trust in AI outputs.

While this study demonstrated effective implementation alongside a substantial increase in adoption within an early adopter program at a single tertiary-care institution, orchestration platforms have the potential to support cross-institutional harmonization. Integration of heterogeneous real-time data sources, including imaging, histopathology, molecular diagnostics, and clinical records, could allow multimodal processing across specialties, with applications feeding results directly into workflows

spanning radiology, pathology, oncology, radiotherapy, and other departments. Such interoperability would enhance resource efficiency and support patient-centered care, and leverage individual data through the platform to serve as a foundation for precision medicine. And as the number of AI applications continues to expand across disciplines, scalability will be essential. Orchestration platforms offer a sustainable infrastructure to integrate increasing data volumes and diverse application types across institutional and cross-institutional workflows.

Our findings address integration and clinician adoption; however, a comprehensive evaluation of AI in clinical practice should consider multiple domains beyond adoption and workflow integration, including sustainability, cost, and governance, to fully capture real-world AI impact¹³.

Adoption should not be conflated with accuracy, as these represent distinct evaluation domains. In this study, both acceptance and rejection of an AI output constituted verification; therefore, the user-verification rate reflects documented engagement with AI outputs rather than the proportion of outputs accepted. Adoption reflects whether clinicians actively engage with a tool in practice, whereas accuracy relates to the correctness of its outputs, which may vary across applications and clinical use cases. Because the outcome measures in this study reflect user interaction with AI outputs (verification) rather than diagnostic correctness, direct conclusions regarding clinical decision-making or clinical impact cannot be drawn. Because the primary analyses were performed at the AI-task level, observations were not fully independent. Therefore, confidence intervals and P values should be interpreted as descriptive implementation metrics rather than strictly confirmatory inferential estimates. No correction for multiple testing was performed, and subgroup analyses should be interpreted as exploratory.

The study was limited to a relatively short evaluation period, limiting the ability to assess long-term adoption data. Subgroup analyses across individual users and applications were constrained by sample size, and workload data could not be systematically incorporated, as reporting does not always coincide with image acquisition. Consultant adoption rates were calculated from their independently reported cases, whereas trainee adoption rates were derived from consultant-trainee paired reporting, as trainees do not report independently.

Pre-implementation adoption could only be assessed for applications from a single vendor, as comparable baseline data were not available for other applications. Although the two chosen time periods are comparable app-wise and in workflow settings, the post-implementation period might be considered a natural continuation of the pre-implementation period as the first exposure to clinical AI. This is true for consultants active in both periods; however, it is not directly applicable to trainees, as all post-implementation trainees were new. Thus, the results reflect change management at both the individual and departmental levels. Accordingly, the findings should be interpreted as associations rather than implying a direct causal effect. In addition, transitional system modifications during the intervening period may have introduced temporary variability in adoption patterns, limiting the interpretability of intermediate data. Because the distribution of clinical use cases differed between periods, the overall verification rate should be interpreted alongside the application-specific analyses.

The single-center design represents another limitation, as institutional factors such as pre-existing AI culture, hardware, and software infrastructure may influence adoption, restricting generalizability to other settings. No formal economic evaluation was conducted, as the study primarily aimed to assess adoption and workflow optimization rather than cost-effectiveness.

Future research should evaluate the long-term sustainability of AI adoption and assess the generalizability of these findings across institutions with different clinical workflows, technical infrastructures, and levels of AI maturity. Larger prospective and multicenter studies could further investigate user-level determinants of adoption and examine how engagement with AI outputs translates into their use in routine reporting and clinical

practice. The economic implications of implementing and maintaining such systems should be evaluated as a separate domain.

In conclusion, in an era of rapidly expanding AI applications, orchestration platforms represent a sustainable strategy for workflow automation and interoperability. When combined with structured change management to enable successful integration into the clinical workflow, this dual-track socio-technical approach was associated with a substantial increase in AI adoption and integration into daily practice. These strategies may support improved efficiency and scalability across institutions and more integrated patient-centered care.

Methods

Study settings

The study was conducted in a tertiary-care neuroradiology department. Bern University Hospital comprises approximately 11,000 employees across 82 clinics, centers, and institutes and serves around 900,000 inpatients, 57,000 outpatients, and 55,000 emergency visits annually. The Institute of Diagnostic and Interventional Neuroradiology is embedded within an operative and administrative Neurocenter, alongside the departments of neurology and neurosurgery, and provides services for both inpatients and outpatients from internal and external referrals. Services encompass neurovascular imaging (acute stroke management, small vessel disease, aneurysms, vascular malformations), neuro-oncology (primary diagnosis and follow-up of brain tumors), traumatic brain injury, neuroinflammatory and demyelinating conditions (e.g. MS), neurodegeneration, epilepsy, including pre-surgical evaluation, and cerebrospinal fluid disorders. The institute employs 84 staff, including 19 board-certified neuroradiologists, and operates six clinical MRI scanners (four 3 T, one 1.5 T, and one intraoperative 1.5 T MRI), two research-dedicated scanners (3 and 7 T), a CT scanner, and three angiography systems. Additional facilities include two CT scanners and one MRI at affiliated peripheral hospitals, as well as a CT scanner in the emergency department shared with general radiology. Annually, the department reports ~9000 CTs, 23,000 MRIs, and 1600 angiographic procedures. A teleradiology network connects 19 external hospitals to provide both primary reporting and second opinions.

Clinical workflow and context

The high daily imaging volumes present challenges for timely, high-quality, standardized reporting. In this setting, AI applications are increasingly used to support workflow optimization through functions such as triage of acute cases, AI-assisted detection, and extraction of quantitative imaging biomarkers.

To ensure diagnostic consistency and longitudinal comparability, the department developed an institutional “Playbook” of clinical imaging SOPs, comprising harmonized protocols, to standardize and optimize practice. Most imaging examinations are reported by trainees (residents and fellows) under consultant supervision; consultants also serve as primary readers depending on workload.

Pre-platform baseline

Parallel to independent research on AI for medical imaging and the development of in-house solutions, the institute integrated the first commercial AI applications in 2021 (Fig. 4) lesion, volumetry, and aneurysm as an on-premises solution. To ensure oversight, the workflow was configured for manual PACS-to-DICOM node transfer for processing, with results returned to PACS after viewer verification.

Platform implementation

In 2023, given the increasing availability of commercial AI applications, data overload, and rising operational complexity, our institute joined an early-adopter program for a vendor-neutral AI orchestration platform (supplementary Table 5) (Fig. 4). This involved a contractual agreement between the vendor, the department of neuroradiology, and hospital informatics to implement the technology in a real-world setting. The project was initiated in March 2023 following hospital management approval and involved close collaboration among technical and medical teams, hospital IT services, PACS representatives, multiple app vendors, legal and administrative offices, clinicians, neuroradiologists, and technologists. The platform integrated multiple AI applications from different vendors and coordinated their execution within the clinical workflow, thereby harmonizing imaging data handling across modalities, PACS, and AI applications.

Technical deployment started in May 2023. An edge device was installed to allow for automated transfer of pseudo-anonymized studies

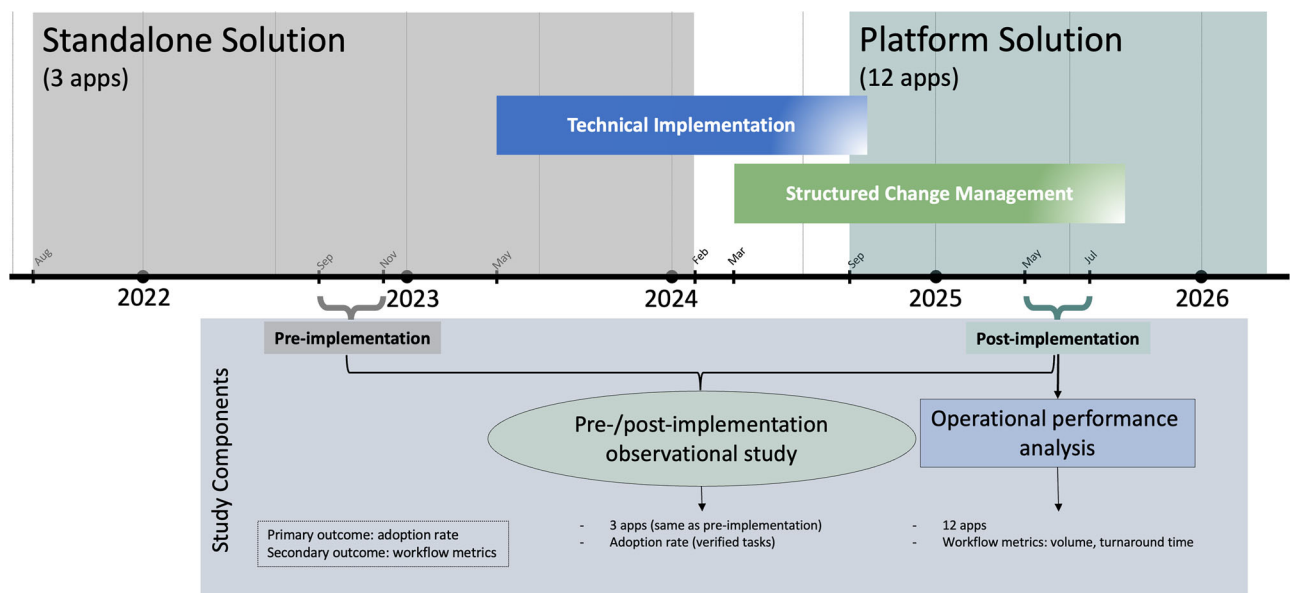


Fig. 4 | Implementation timeline and evaluation design. Upper panel: Timeline showing the pre-implementation period (left; standalone solution; 3 apps) and the post-implementation period (platform solution; 12 apps). Ribbons denote the two parallel tracks, technical implementation and structured change management, with change-management activities beginning before the initial clinical roll-out. Lower

panel: study components and analysis windows: pre-implementation (September–November 2022) and post-implementation (May–July 2025). The evaluation comprised (i) a pre-/post- implementation observational study assessing adoption rate and (ii) post-implementation operational performance analysis of workflow metrics.

from PACS to the AI applications. Patient data were pseudo-anonymized on the on-premises edge device prior to transfer, following the DICOM PS 3.15 Basic Application-Level Confidentiality Profile, ensuring removal or replacement of identifying attributes while maintaining data integrity for clinical use. Results were returned over an encrypted tunnel and re-identified only within the hospital network. Processing of the pseudo-anonymized data was performed in an EU-based cloud environment, including regional third-party clouds for selected applications. Data retention and deletion followed the application-specific contractual provisions, with data deleted after processing.

Primary integration was performed via PACS. Integration steps included: (i) configuration of additional DICOM nodes on CT and PACS for routing imaging studies to the edge device, (ii) implementation of a viewer-launcher button within PACS for direct access to platform results, (iii) a dedicated HL7 field to display AI processing status, and (iv) specific worklist-prioritization functionality. To support data throughput and redundancy, two virtual servers were deployed.

Based on clinical needs across multiple neuroradiology workflows, 12 applications from four vendors were integrated, covering acute triage, assisted detection, longitudinal monitoring, quantitative analysis, and advanced post-processing (Supplementary Table 5). Platform orchestration²¹ enabled automated routing of relevant imaging series according to predefined “relevancy rules”, highlighting the importance of standardized imaging protocols. To ensure optimal performance, vendor-recommended parameter optimization was implemented at the scanner level. Key terminology is summarized in the supplementary glossary.

Prior to deployment, staged testing was performed using retrospective cases to assess clinical utility, followed by prospective cases to evaluate automation reliability, clinical utility, and processing time. After iterative modifications (including viewer updates, improved automation, and query retrieve functionality for longitudinal analysis), the system was released for full clinical use in September 2024. After rollout, occasional short-duration downtime events occurred (fewer than 10 instances per year), typically related to server processes and resolved through server restart, resulting in only brief interruptions without lasting impact on clinical workflow. Examinations performed during these intervals could subsequently be reprocessed manually when clinically required.

Change management

In parallel with technical implementation, we implemented a structured change-management program (Table 1) to support adoption and integration into the routine workflow. This included tailored training, verification guidance, documentation, feedback loops, and governance measures. Verification refers to the end-user review of AI outputs, whereby radiologists actively confirm (accept) or discard (reject) results within the reporting workflow. Regular scientific meetings on AI, supported by internal case examples, were conducted. Essential information from vendors’ instructions for use, supplemented by locally defined evaluation and reporting steps, was distilled into “model-card-like” summaries²⁷ and consolidated into standardized “AI Playbook”. This was embedded into the existing departmental playbook to facilitate familiarity and adoption. Technical and AI output-based feedback channels were established. A dedicated team of expert users provided troubleshooting and clinical support during rollout (Fig. 5). The change management strategy was designed and implemented internally, tailored to local workflow requirements. The vendor’s involvement was limited to initial platform onboarding and did not influence the structure or execution of the change management process.

End-user workflow after platform implementation

Following platform implementation, eligible imaging examinations were automatically routed to relevant AI applications based on predefined relevancy rules. AI outputs were returned to the integrated platform viewer and linked to the patient study in the PACS environment, allowing radiologists to access results during routine reporting without separate manual data

transfer or multiple standalone logins. This replaced a fragmented multi-application workflow with a unified reporting experience.

A dedicated status field within PACS displayed the processing stage of each examination (e.g., initiated, ready for verification, completed), allowing radiologists to decide whether to begin reporting immediately or briefly await AI completion. In routine practice, processing speed was optimized such that AI results were typically available before the usual reporting started.

Urgent imaging findings generated by selected applications were flagged within PACS to alert radiologists and support timely case prioritization.

For most applications, output followed a human-in-the-loop workflow, requiring radiologist review and explicit verification before AI outputs were stored in the clinical environment. This step ensured expert oversight, reduced the risk of inappropriate AI outputs being displayed, and minimized the risk that preliminary or erroneous unverified output could be misinterpreted by patients or referring clinicians. For selected applications (perfusion processing), parametric maps were automatically transferred to PACS without a separate verification step, reflecting a human-on-the-loop model in which outputs were generated for direct use during image interpretation while remaining under radiologist supervision.

Access to the full AI viewer was granted to neuroradiologists trained in the interpretation of both imaging findings and AI outputs. Referring clinicians could access only outputs accepted by radiologists and stored within PACS, such as volumetric reports, atrophy patterns, or annotated image series. A schematic overview of the end-user workflow is provided in Supplementary Fig. 6, and representative examples of AI outputs are shown in Supplementary Fig. 7.

Study design and outcomes

Ethics approval was waived because the study evaluated a system-level operational strategy within existing departmental processes and did not constitute human-subject research under the Swiss Human Research Act (HRA). The Cantonal Ethics Committee Bern confirmed this assessment. Clinician interaction data were extracted from routine operational systems, analyzed in coded form, and reported only in a summarized or non-identifiable format.

This retrospective study focused on the technical performance of the orchestration platform and the adoption rate of AI tools in practice. We conducted a single-center evaluation of the intervention (orchestrator platform and structured change management) in routine neuroradiology workflow, following the Standards for Reporting Implementation Studies (StaRI)²⁸. The vendor had no role in study design, data analysis, interpretation of results, or manuscript preparation. The primary implementation outcome was AI adoption (overall and in examinations, including multiple tasks), operationalized as the user-verification rate, while secondary outcomes included platform operational metrics relevant to workflow integration, including processed-study volume and turnaround time.

Technical performance metrics were selected to evaluate whether the platform could operate reliably and at sufficient scale for integration into a high-throughput routine workflow. The platform’s monitoring and analytics dashboard was used to extract operational metrics, including processed-study counts and turnaround times for each integrated application. Processed studies were defined as examinations successfully analyzed by one or more integrated applications. Turnaround time was defined as the interval between image availability and completed AI result generation.

Adoption was measured as the user-verification rate: the proportion of AI outputs that received an explicit verification action by the reporting radiologist (accept or reject), out of all applicable AI tasks. Here, ‘verification’ refers to the radiologist’s interaction with the AI output, rather than its diagnostic correctness. Before the platform implementation, verification required completion of the manual workflow, including transfer of the examination for AI processing and subsequent verification of the results. After platform implementation, examinations were routed automatically, but verification still required the radiologist to open the dedicated platform

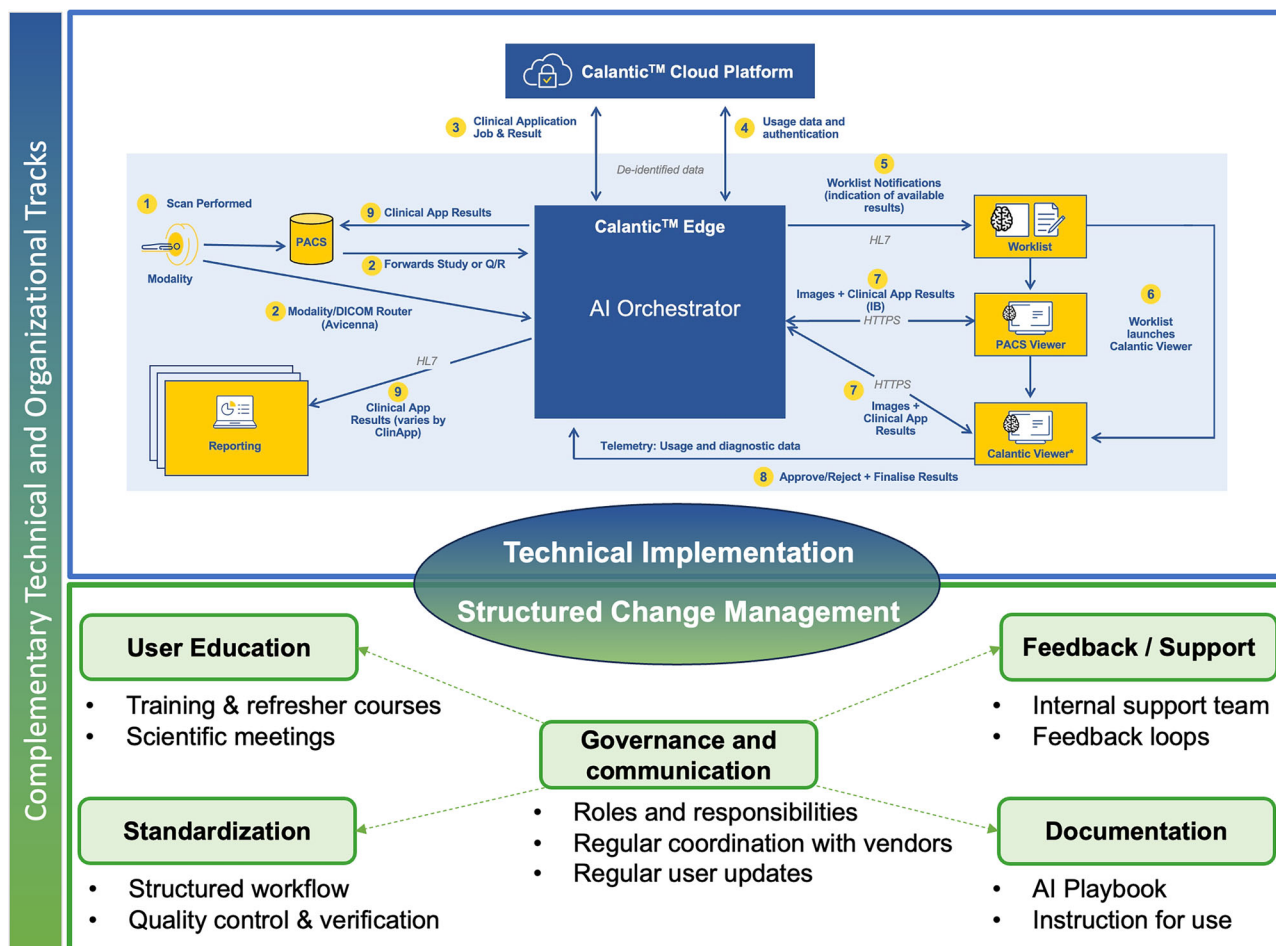


Fig. 5 | Integrated implementation strategy: combined illustration of technical and organizational approaches to AI adoption. *Top (blue):* technical implementation showing the workflow of the vendor-neutral orchestration platform and integration into institutional systems. Imaging data are transferred via edge to the cloud platform, where multiple commercial AI applications are orchestrated in parallel. AI results are automatically routed back to the viewer/PACS and made available to radiologists through a dedicated viewer button, HL7 communication

fields (processing/verification status), and a structured worklist. The system supports multi-application processing, automated workflow routing, and centralization of AI task verification. Verification in this study denotes end-user review of AI outputs whereby radiologists actively confirm (accept) or discard (reject) results within the reporting workflow. *Bottom (green):* summarizing the steps of structured change management.

viewer and explicitly accept or reject the AI output. Conversely, “no verification” indicated that the AI output was not consulted through the respective review workflow. Because each AI output required an explicit decision, the verification rate served as an objective and auditable indicator of active clinical engagement with the AI system rather than passive exposure. Both acceptance and rejection were counted as verification; therefore, the user verification rate measured documented consultation of AI outputs rather than the proportion of outputs accepted.

We conducted an observational pre-post evaluation study. AI adoption was evaluated by comparing two periods: a Pre-implementation baseline (September 1–November 30, 2022; 13 months after the single-vendor AI introduction at our institute), and a post-implementation (May 1–July 31, 2025; 8 months after full platform deployment, and 37 months after first clinical AI use at the institute) (Fig. 4). No formal sample-size calculation was performed; to reflect routine clinical activity, we included all consecutive eligible MRI examinations during the two defined periods. For the comparative adoption analysis, we restricted the dataset to the three applications that were available in both pre- and post-implementation periods, as other applications, including CT-based tasks, were newly introduced after platform implementation and therefore lacked comparable baseline data. The performance of the three applications included in the adoption analysis has been previously validated in dedicated studies^{29–31}. MRI studies were

identified via PACS search by exam date. Eligible studies were identified according to the clinical indication, and predefined application-specific criteria (Supplementary Table 6). From these, five clinical use cases were defined: (1) follow-up of intracranial aneurysms, (2) primary and opportunistic screening for intracranial aneurysms, (3) evaluation of brain atrophy in neurodegenerative disorders (dementia and movement disorders), (4) screening and longitudinal monitoring of multiple sclerosis (including radiologically isolated syndrome), and (5) lesion detection and quantification in multiple sclerosis mimics (e.g. other inflammatory conditions, small vessel disease, sarcoidosis). Only exams performed on clinical scanners were included; research scans and acquisitions from peripheral hospitals were not included. All PACS searches and eligibility assessments were performed by a senior consultant neuroradiologist responsible for MR protocol design and optimization within the department.

Statistics and analysis

All analyses were performed independently by the authors. AI-applicable imaging examinations were exported to a structured table containing examination date, reader(s), trainees, and supervising consultant, application(s) triggered per exam, number of tasks per exam, and the verification outcome for each task.

For this study, overall adoption was defined as the user-verification rate (verified AI results/total applicable AI tasks). Multitask exams were defined as exams with more than one applicable task, and were summarized according to the proportion with no verification, single-application verification, and all-applications verification. Verification proportions were summarized using Wilson 95% confidence intervals, and comparisons between 2022 and 2025 verification rates were performed using χ^2 tests. For MS lesion analyses, automatic volumetry released from the same case was counted as part of a single task.

The weekly verification rate in the post-implementation period was modeled using least-squares fitting of a linear model.

Consultant adoption was calculated from consultant-only reports. Trainee adoption was calculated from all reports in which the trainee contributed and a supervising consultant finalized the case (consultant-trainee pairs). For both groups, adoption was categorized as: minimal (<25%), intermediate (25–75%), high (75–90%), and very high (>90%).

Exploratory subgroup analyses were performed to evaluate adoption patterns according to training status, departmental duration, and AI-task exposure volume. Because verification rates were not normally distributed, nonparametric tests were used (Mann–Whitney *U* test and Kruskal–Wallis test, as appropriate).

Data availability

The datasets generated and/or analyzed during the current study contain patient-level clinical and operational workflow data and are not publicly available due to institutional, ethical, and legal data protection restrictions. Requests for access to de-identified data may be considered by the corresponding author on reasonable request and subject to institutional approval and applicable data protection regulations.

Code availability

No custom code or scripts were generated or used for the analyses reported in this study. Statistical analyses were performed using standard statistical methods.

Received: 6 December 2025; Accepted: 10 September 2026;

Published online: 22 September 2026

References

- Brady, A. P. et al. Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. a Multi-Society Statement from the ACR, CAR, ESR, RANZCR and RSNA. *Radiol. Artif. Intell.* **6**, (2024).
- van Leeuwen, K. G., Schalekamp, S., Rutten, M. J. C. M., van Ginneken, B. & de Rooij, M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *Eur. Radiol.* **31**, 3797–3804 (2021).
- Antonissen, N. et al. Artificial intelligence in radiology: 173 commercially available products and their scientific evidence. *Eur. Radiol.* <https://doi.org/10.1007/s00330-025-11830-8> (2025).
- Sosna, J., Joskowicz, L. & Saban, M. Navigating the AI landscape in medical imaging: a critical analysis of technologies, implementation, and implications. *Radiology* **315**, (2025).
- Korfatis, P. et al. Implementing artificial intelligence algorithms in the radiology workflow: challenges and considerations. *Mayo Clin. Proc.: Digit. Health* **3**, 100188 (2025).
- Eltawil, F. A., Atalla, M., Boulous, E., Amirabadi, A. & Tyrrell, P. N. Analyzing barriers and enablers for the acceptance of artificial intelligence innovations into radiology practice: a scoping review. *Tomography* **9**, 1443–1455 (2023).
- Afshari Mirak, S., Tirumani, S. H., Ramaiya, N. & Mohamed, I. The growing nationwide radiologist shortage: current opportunities and ongoing challenges for international medical graduate radiologists. *Radiology* **314**, (2025).
- Jing, A. B., Garg, N., Zhang, J. & Brown, J. J. AI solutions to the radiology workforce shortage. *npj Health Syst.* **2**, 20 (2025).
- Wenderott, K., Krups, J., Weigl, M. & Wooldridge, A. R. Facilitators and barriers to implementing AI in routine medical imaging: systematic review and qualitative analysis. *J. Med. Internet Res.* **27**, e63649 (2025).
- Wenderott, K., Krups, J., Luetkens, J. A. & Weigl, M. Radiologists' perspectives on the workflow integration of an artificial intelligence-based computer-aided detection system: a qualitative study. *Appl. Ergon.* **117**, 104243 (2024).
- Kim, B., Romeijn, S., van Buchem, M., Mehrizi, M. H. R. & Grootjans, W. A holistic approach to implementing artificial intelligence in radiology. *Insights Imaging* **15**, 22 (2024).
- Salwei, M. E. & Carayon, P. A sociotechnical systems framework for the application of artificial intelligence in health care delivery. *J. Cogn. Eng. Decis. Mak.* **16**, 194–206 (2022).
- Jacob, C. et al. AI for IMPACTS framework for evaluating the long-term real-world impacts of AI-powered clinician tools: systematic review and narrative synthesis. *J. Med. Internet Res.* **27**, e67485 (2025).
- Naicker, S. et al. Implementing an artificial intelligence decision support system in radiology: prospective qualitative evaluation study using the nonadoption abandonment scale-up, spread, and sustainability (NASSS) framework. *J. Med. Internet Res.* **28**, e80342 (2026).
- Strohm, L., Hehakaya, C., Ranschaert, E. R., Boon, W. P. C. & Moors, E. H. M. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. *Eur. Radiol.* **30**, 5525–5532 (2020).
- Lawrence, R. et al. Artificial intelligence for diagnostics in radiology practice: a rapid systematic scoping review. *EClinicalMedicine* **83**, 103228 (2025).
- Verwey, J., Zwart, B., IJzerman, M., Visser, J. J. & Sülz, S. Factors influencing AI acceptance in radiology: a systematic review across the radiology workflow. *Eur. Radiol.* **36**, 3687–3696 (2025).
- Sohn, J. H. et al. An open-source, vendor agnostic hardware and software pipeline for integration of artificial intelligence in radiology workflow. *J. Digit. Imaging* **33**, 1041–1046 (2020).
- Tejani, A. S., Cook, T. S., Hussain, M., Sippel Schmidt, T. & O'Donnell, K. P. Integrating and adopting AI in the radiology workflow: a primer for standards and integrating the healthcare enterprise (IHE) profiles. *Radiology* **311**, (2024).
- de Almeida, J. G. et al. Medical machine learning operations: a framework to facilitate clinical AI development and deployment in radiology. *Eur. Radiol.* **35**, 6828–6841 (2025).
- Cohen, R. Y. & Sodickson, A. D. An orchestration platform that puts radiologists in the driver's seat of AI innovation: a methodological approach. *J. Digit. Imaging* **36**, 700–714 (2022).
- Recht, M. P. et al. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. *Eur. Radiol.* **30**, 3576–3584 (2020).
- van Leeuwen, K. G., de Rooij, M., Schalekamp, S., van Ginneken, B. & Rutten, M. J. C. M. Clinical use of artificial intelligence products for radiology in the Netherlands between 2020 and 2022. *Eur. Radiol.* **34**, 348–354 (2023).
- Shah, C., Ghodasara, S., Chen, D. & Chen, P.-H. Does it work, help, and stay? A framework for implementing artificial intelligence tools in radiology. *J. Am. College Radiol.* **23**, 378–388 (2026).
- Al Mohammad, B., Aldaradkeh, A., Gharaibeh, M. & Reed, W. Assessing radiologists' and radiographers' perceptions on artificial intelligence integration: opportunities and challenges. *Br. J. Radiol.* **97**, 763–769 (2024).
- Volin, J., van Assen, M., Bala, W., Safdar, N. & Balthazar, P. Artificial intelligence and its effect on radiology residency education: current

- challenges, opportunities, and future directions. *J. Am. College Radiol.* **22**, 1264–1270 (2025).
27. Mitchell, M. et al. Model cards for model reporting. In *Proc. Conference on Fairness, Accountability, and Transparency* 220–229 (ACM, 2019). <https://doi.org/10.1145/3287560.3287596>.
28. Pinnock, H. et al. Standards for Reporting Implementation Studies (StaRI) statement. *BMJ* i6795 (2017) <https://doi.org/10.1136/bmj.i6795>.
29. Lehnen, N. C. et al. Automated detection of cerebral aneurysms on TOF-MRA using a deep learning approach: an external validation study. *Am. J. Neuroradiol.* **43**, 1700–1705 (2022).
30. Rudolph, J. et al. Artificial intelligence-based rapid brain volumetry substantially improves differential diagnosis in dementia. *Alzheimer's Dement.: Diagn. Assess. Dis. Monit.* **16**, (2024).
31. Peters, S. et al. AI-based assessment of longitudinal multiple sclerosis MRI: strengths and weaknesses in clinical practice. *Eur. J. Radiol.* **183**, 111941 (2025).

Acknowledgements

The institution participated in a contracted early-adopter program for the Calantic platform (Bayer AG), which included technical support for implementation. No external research funding was received for this implementation study. The vendor had no role in study design, data analysis, interpretation of results, or manuscript preparation. This work was made possible through close collaboration between Bern University Hospital and industry partners. We gratefully acknowledge the Calantic Bayer team under the leadership of Lucia Fernandez Lopez, including Edalecio Simoes, Rudi Wolfs, Mike Zinzen, Tobias Pink, Christian Lienerth, Juan Gutierrez Alliende, and Tomas Lopes da Fonseca, for their substantial support during platform implementation and integration. We especially thank Edouard Cugni for his invaluable technical support throughout the project.

Author contributions

A.H.: Conceptualization, data curation, formal analysis, methodology, project administration, resources, visualization, and writing-original draft. C.R.: Analysis, verifying the data, validation, visualization, and writing-review and editing. C.J.: Writing-review and editing. P.R.: Resources, project administration, and writing-review and editing. R.W.: Conceptualization, project administration, resources, and writing-review and editing.

Competing interests

A.H. reports honoraria for presentations and support for attending meetings from Bayer AG. C.R. declares no competing interests. C.J. declares no

competing interests. P.R. declares no competing interests. R.W. reports honoraria for presentations from Bayer AG, consulting fees from Lilly, Servier, and Eisai, participation in the Teaching Course Subcommittee of the European Academy of Neurology, and institutional research funding from the Swiss National Science Foundation. The authors declare no other competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-026-03278-x>.

Correspondence and requests for materials should be addressed to Arsany Hakim.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.