DR. SERGEY MOROZOV

WEEKLY AI INTELLIGENCE REPORT

Radiology and Medical Imaging

Week 35 | 23 to 29 August 2026

Dr. Sergey Morozov | drsergeymorozov.com

Compiled from publicly available sources (indexed peer-reviewed literature, official press releases, regulatory databases). Does not constitute medical, regulatory, or financial advice.

8	5	3	3
Peer-reviewed papers	Industry / FDA items	Media highlights	Key opinion leaders

GOVERNANCE TAKEAWAY OF THE WEEK

This week's evidence cluster is operational AI, the tools that never reach your AI committee because nobody calls them clinical: a no-show model that decides which patients get outreach along a deprivation gradient, a camera system screening MR-unsafe wheelchairs, an LLM routing extraspinal findings to specialists. Pick one operational tool that is live or planned in your department this month and write its acceptance test around the harmful error direction: the high-risk patient not flagged, the unsafe chair called safe, the finding downgraded out of the follow-up queue. The wheelchair study shows the method: report the harmful direction on its own denominator (8.3% of MR-unsafe instances missed in the worst-case model, invisible inside a 0.702 mAP), then decide.

SECTION 0

EXECUTIVE SUMMARY

Key trends, week of 23 to 29 August 2026

[OPERATIONS] NO-SHOW PREDICTION IS AN EQUITY DECISION WITH A 57.7 MILLION DOLLAR PRICE TAG

Tippareddy C et al. (JACR) built a no-show model on 32,776 radiology appointments: no-show rates climb from 15.0% to 31.8% along the deprivation gradient, the top risk decile holds 25.7% of all no-shows at a number needed to intervene of 1.8, and the yearly cost of missed appointments was estimated at 57.7 million dollars. Whoever deploys this class of model is choosing which patients get outreach.

[SAFETY AI] THE ACCEPTANCE NUMBER IS THE HARMFUL ERROR DIRECTION, NOT THE DETECTION SCORE

Suzuki T et al. (JIIM) put a YOLOv8n model on MRI anteroom cameras to catch MR-unsafe wheelchairs. Headline detection looked fine (mAP@0.5:0.95 of 0.702), but the number that decides deployability is different: up to 8.3% of MR-unsafe wheelchairs called safe in the worst-case model, end-to-end sensitivity 81.5%. The authors label it feasibility work, and their failure-mode accounting is the template to copy.

[WORKFLOW LLM] REPORT FOLLOW-UP AUTOMATION CLEARS A HIGH BAR, WITH ONE ERROR DIRECTION TO WATCH

Ong W et al. (JMIR) validated a privacy-preserving LLM that finds and routes extraspinal findings in 400 MRI spine reports: 99.8% detection, all 12 urgent findings caught, agreement matching two human readers. The residual risk sits in the 2% of findings downgraded to lower significance categories, the direction that silently removes findings from follow-up queues.

[TRAINING DATA] WHAT IS IN THE TRAINING SET DECIDES WHERE THE MODEL WORKS

Zielke J et al. (AJNR) added post-treatment scans to training for pediatric diffuse midline glioma segmentation and then tested externally on a prospective trial: whole-tumor Dice 0.90 vs 0.81 for the BraTS-PEDs 2024 winner, with the biggest gap on post-treatment scans. Training-set composition, not architecture, was the difference.

[WATCHPOINT] OPERATIONAL AI NEEDS THE SAME GOVERNANCE AS DIAGNOSTIC AI

Three of this week's strongest results are not diagnostic algorithms: a scheduling model that decides who gets outreach, a camera system that guards the MRI suite, and an LLM that routes findings to specialists. None of them will pass through the committee that reviews your CAD tools unless you send them there. The common instruction: define the harmful error direction for each (high-risk patient not flagged, unsafe chair called safe, finding downgraded out of follow-up), test on that number, and monitor it after go-live.

SECTION 1

PEER-REVIEWED PAPERS

Source: PubMed | Verified indexing | No preprints | 23 to 29 August 2026

Multimodal Ultrasound-Based Decision Support for Bethesda IV Thyroid Nodules: Integration of Clinical, Radiomics, Deep Learning, and Topological Features.

Ding H, et al. | ***Journal of imaging informatics in medicine*** | *2026-08-25*

Before you accept any multimodal fusion claim, look at what survives correction for multiple comparisons. In this single-center study of 243 Bethesda IV thyroid nodules, adding radiomics, deep learning and topological features to clinical information produced the highest observed AUC (0.836 for the full fusion model), yet no pairwise AUC difference among the fusion models remained statistically significant after Holm-Bonferroni correction. The authors are unusually honest: reclassification improved, discrimination did not, and they position the tool as an interpretable probability estimate for case-level assessment, not a stand-alone treatment decision.

Key metrics: 243 patients (138 malignant, 105 benign), 7:3 patient-level split. Validation AUCs: full fusion 0.836, clinical + radiomics + deep learning 0.829, clinical + radiomics 0.828, clinical + deep learning 0.805; no pairwise difference significant after Holm-Bonferroni. Continuous NRI vs clinical baseline 0.750 (95% CI 0.310 to 1.174, p = 0.002); IDI 0.029 (95% CI -0.050 to 0.105, p = 0.478).

Clinical relevance: Bethesda IV is exactly the indeterminate category where surgeons need a better preoperative probability, and this framework shows a structured way to combine what is already collected. The result itself is modest and single-center, and the authors state plainly that prospective multicenter external validation across institutions and ultrasound systems is required before clinical implementation.

PMID 42642692 **DOI** 10.1007/s10278-026-02208-3

Radiology No-Show Calculator: A Social Determinants of Health-Enriched Machine Learning Prediction Model with Financial Analysis.

Tippareddy C, et al. | ***Journal of the American College of Radiology*** | *2026-08-27*

Before you fund a no-show intervention, quantify who is actually missing appointments. Across 32,776 outpatient radiology appointments, no-show rates ran from 15.0% in the least deprived neighborhoods to 31.8% in the most deprived, and a 16-predictor random forest enriched with social determinants of health concentrated a quarter of all no-shows in the top risk decile, with a number needed to intervene of 1.8. The estimated yearly cost of missed appointments at this academic system: 57.7 million dollars.

Key metrics: 32,776 appointments from 9,998 patients (1:1 case-control design, population no-show prevalence 3.9%). Random forest with 16 predictors including Area Deprivation Index: AUROC 0.760. No-show rate 15.0% (least deprived) to 31.8% (most deprived). Top 10% of risk captured 25.7% of all no-shows, NNI 1.8. Estimated yearly cost burden 57.7 million USD. Probabilities recalibrated to population prevalence for deployment.

Clinical relevance: This is operational AI with direct equity consequences: the model's strongest signal is neighborhood deprivation, so any deployment decision is also a decision about which patients get outreach. So what for your AI committee: acceptance testing for scheduling models must include calibration to your own prevalence and performance by deprivation subgroup, and monitoring must track whether targeted outreach actually closes the gradient. Under EU AI Act Article 10, the data governance file should record the social determinants used and why.

PMID 42660476 **DOI** 10.1016/j.jacr.2026.08.015

Feasibility of AI-Based Wheelchair Classification for MRI Anteroom Safety Management: A Single-Centre Deep Learning Object-Detection Study.

Suzuki T, et al. | ***Journal of imaging informatics in medicine*** | *2026-08-26*

Read this as a template for the acceptance question that matters in safety AI: not the headline detection score but the harmful error direction, on the denominator that matters. A lightweight YOLOv8n model watching an MRI anteroom through surveillance cameras classified wheelchairs as MR-safe or MR-unsafe; in the worst-case model, 8.3% of MR-unsafe wheelchairs were misidentified as safe, and end-to-end sensitivity for MR-unsafe chairs was 81.5%. The authors call it what it is: a single-center feasibility study with a characterized failure-mode profile, hypothesis-generating rather than deployment-ready.

Key metrics: 1,730 images over 5 working days from two cameras (training 1,251, validation 118, test 361). 640 px models: mAP@0.5:0.95 of 0.702 ± 0.012, 41.4 FPS CoreML inference on Apple M4. Critical misclassification (MR-unsafe called MR-safe) up to 8.3% against the 72 MR-unsafe instances, 0.21% to 1.28% (mean 0.64%) of all 468 annotated instances. End-to-end sensitivity for MR-unsafe wheelchairs 81.5% (range 76.4 to 86.1%).

Clinical relevance: MR-unsafe wheelchairs in the scanner room are a real hazard, and camera-based screening is an attractive backstop. So what for your AI committee: this changes how you write acceptance tests for any safety-screening AI. Specify the clinically harmful error direction (unsafe called safe), report it on the clinically relevant denominator, and set the go-live threshold there, not on mAP. Under EU AI Act Article 15 (accuracy and robustness), the worst-case model across seeds, not the average, is the number that belongs in the risk file.

PMID 42649334 **DOI** 10.1007/s10278-026-02219-0

EndoVLM: A Vision-Language Assistant for Gastrointestinal Endoscopy.

Wang S, et al. | ***Journal of imaging informatics in medicine*** | *2026-08-26*

If your endoscopy service is tracking vision-language assistants, the useful number here is the external one. EndoVLM replaces the usual ViT encoder with a hierarchical ConvNeXt backbone to compress high-resolution endoscopy frames into fewer, denser tokens, and adds a three-stage fine-tuning schedule. On the Kvasir-VQA benchmark it edges past the strongest public baseline; on external validation against Gastrovision, accuracy rises from 0.4760 to 0.5799 over the two-stage baseline, which is a real gain and still far from clinical territory.

Key metrics: Kvasir-VQA: ROUGE-1 0.7326 ± 0.0038 and BLEU 0.5103 ± 0.0051, improving on the strongest public baseline by +0.0166 ROUGE-1 and +0.0323 BLEU. External validation on Gastrovision: accuracy 0.4760 to 0.5799 and macro-F1 0.1109 to 0.1268 versus the representative two-stage baseline.

Clinical relevance: The engineering direction is sensible: hierarchical convolutional encoders cut token overhead on high-resolution frames, which matters for deployment cost. The external numbers speak for themselves: a specialized, carefully adapted model still answers barely more than half of external questions correctly, with a macro-F1 that signals weak performance on rarer findings. Research-stage work, presented on its merits.

PMID 42649331 **DOI** 10.1007/s10278-026-02067-y

An Imaging Informatics Workflow for Angiography-Derived Coronary Physiology Profiling and Virtual PCI Simulation.

Georgy K, et al. | ***Journal of imaging informatics in medicine*** | *2026-08-26*

This is what workflow integration looks like when it is taken seriously: AngioAI-QFR chains automated stenosis localisation, lumen segmentation, geometric reconstruction, physiology estimation and virtual PCI simulation into one pipeline over routine coronary angiography. Against invasive FFR in 100 vessels it reached an AUROC of 0.943 for FFR of 0.80 or below, ran fully automatically in 93% of vessels, and returned results in a median of 41 seconds, fast enough to matter inside a procedure.

Key metrics: 100 vessels from 92 patients, invasive FFR as reference. Detection precision 0.966, mAP@50 0.973, segmentation IoU 0.757, Dice 0.861. Agreement with FFR: Pearson r = 0.917, MAE 0.035, RMSE 0.044, bias -0.004, 95% limits of agreement -0.091 to 0.083. Diagnostic performance for FFR at or below 0.80: AUROC 0.943, sensitivity 0.881, specificity 0.897. Fully automatic in 93% of vessels; median time to result 41 s.

Clinical relevance: Angiography-derived physiology has been vessel-level and fragmented; the per-millimetre flow profiling here separates focal from diffuse disease and feeds an interactive virtual PCI step with immediate physiology recomputation. Retrospective, single evaluation cohort, and the authors list prospective multicentre validation as the next requirement. On clinical merit, one of the more complete imaging informatics builds this year.

PMID 42649327 **DOI** 10.1007/s10278-026-02209-2

Improved Deep Learning Segmentation of Pediatric Diffuse Midline Gliomas After Treatment.

Zielke J, et al. | ***AJNR. American journal of neuroradiology*** | *2026-08-26*

The design decision, not the architecture, is the story: adding post-treatment scans to the training set is what made this pediatric diffuse midline glioma segmentation model work on the images that matter for response assessment. Externally validated on 88 studies from the PNOC007 prospective trial, DMGtracker beat the BraTS-PEDs 2024 winning model on whole-tumor segmentation, and the advantage was largest exactly where longitudinal trials live: post-treatment scans.

Key metrics: Training: 140 institutional pre- and post-treatment studies plus 261 BraTS-PEDs 2024 pre-treatment studies; internal set 153 scans (59 post-treatment) from 74 patients. External validation on 88 PNOC007 studies (49 patients): whole-tumor DSC 0.90 (IQR 0.72 to 0.95) vs 0.81 (0.66 to 0.90) for the BraTS-PEDs 2024 winner; RVD 9.6% vs 16.8%. Post-treatment scans (n = 50): DSC 0.90 vs 0.80, RVD 9.7% vs 19.8%, $p < 0.001$ for both. Clinically acceptable segmentation (DSC > 0.80) in 64.0% vs 52.0% of post-treatment cases.

Clinical relevance: Volumetric response assessment in DMG trials depends on segmentation that survives radiation effects, cavities and vaccine-era imaging changes, and challenge-winning models trained on pre-treatment data degrade exactly there. A model externally tested on prospective trial imaging, with the gap quantified per treatment phase, is the right evidence shape for longitudinal use.

PMID 42648875 **DOI** 10.3174/ajnr.A9609

Automating the Management of Extraspinal Findings in Magnetic Resonance Imaging Spine Studies Using a Privacy-Preserving Large Language Model: Retrospective Validation Study.

Ong W, et al. | ***Journal of medical Internet research*** | *2026-08-26*

If report follow-up management is on your automation list, this is the level of validation to demand before go-live. A privacy-preserving large language model processed 400 MRI spine reports to find, classify and route extraspinal findings: it detected 99.8% of findings, caught all 12 clinically urgent ones, and matched two human readers on overall agreement. The failure mode is the interesting part: 8 of 401 findings were downgraded into lower significance categories, the error direction that quietly hides findings from follow-up.

Key metrics: 400 MRI spine reports from 395 patients; 401 extraspinal findings, most commonly renal (31.9%), gynecological (27.4%), endocrine (12%). Detection 400/401 (99.8%); 12/12 urgent findings identified; 8/401 (2%) misclassified into lower C-RADS categories, 4/401 (1%) into higher. New vs preexisting findings 98.5% accuracy; specialty referral decision 96.3%. Gwet kappa 0.959 (95% CI 0.937 to 0.981) vs human readers at 0.943 and 0.934.

Clinical relevance: Extraspinal findings are a known source of missed follow-up, and LLM triage of existing report text is a low-friction place to deploy language models. So what for your AI committee: this changes acceptance testing and human oversight design for report-workflow LLMs. Score the downgrade direction separately, because a finding classified one category too low disappears from the referral queue without anyone noticing, and keep a human on referral routing per EU AI Act Article 14 until your local error audit says otherwise.

PMID 42647045 **DOI** 10.2196/86251

Gait-based diagnostic network for localization and pathological characterization of spine and pelvis diseases.

Liu F, et al. | ***Medical image analysis*** | *2026-08-24*

A different input signal for an old triage problem: instead of static imaging, this group trained a hierarchical spatiotemporal transformer on clinical gait videos from 419 patients to localize spine and pelvis disease and then characterize it. The first diagnostic level, distinguishing cervical/thoracic from lumbar/pelvis disease, reached 76% accuracy and an AUC of 85%, with attention maps that match established pathophysiology.

Key metrics: 419 patients with clinical gait videos. First diagnostic level (cervical/thoracic vs lumbar/pelvis): accuracy 76%, AUC 85%. Second level classifies tumor vs non-tumor subtypes across regions; the abstract reports no numeric performance for this level. Interpretability via spatiotemporal attention visualizations.

Clinical relevance: Gait carries real diagnostic signal that static imaging ignores, and a camera is cheaper than any scanner, which is why this direction keeps attracting work for primary screening and triage. At 76% accuracy on the first branching decision and a single-center dataset, this is a feasibility argument, not a screening tool. Worth watching for where video-based triage goes next.

PMID 42664663 **DOI** 10.1016/j.media.2026.104260

SECTION 2

INDUSTRY & REGULATION

Sources: Official registers | Press releases | FDA databases

[FDA] ROPCA receives FDA 510(k) clearance for Arthur, an autonomous musculoskeletal ultrasound robot, and its Diana AI analysis software

ROPCA | 26 August 2026

Cleared for: standardized ultrasound image acquisition of finger, hand and wrist joints by the Arthur robot under healthcare-professional supervision, with Diana software analyzing the images and supporting automated report generation for arthritis assessment. Evidence level: the announcement cites CE marking since 2022 and European clinical experience of more than 5,400 scans and 83,500 joints across six countries; no new peer-reviewed US validation is referenced. Questions for the vendor: which joint findings Diana is actually cleared to report in the US and with what performance; how acquisition quality is verified when no sonographer holds the probe; what the supervising professional is expected to check before a report is released; failure and abort rates per scan in routine European use. Regardless of clearance, log robot-acquired studies separately in your QA so acquisition failures and repeat rates are visible from day one.

Source: Imaging Technology News

[MARKET] Independent government-funded FUSION trial: Heartflow FFRCT cuts unnecessary invasive angiography by 44% at one year

Heartflow | 28 August 2026

This is the evidence shape to demand from every imaging AI vendor: an independent randomized trial, funded by the Dutch National Health Care Institute with no industry involvement in design or analysis, across twelve Dutch hospitals and multiple scanner vendors, published simultaneously in JACC. In 528 patients with stable chest pain and obstructive stenosis on CCTA, adding FFRCT reduced unnecessary invasive coronary angiography by 44% at one year (22% vs 39%, $p < 0.001$) while revascularization rates stayed identical (20% in both arms). Evidence level: peer-reviewed RCT, the highest tier seen in this digest for a commercial imaging AI product this year. The buyer question flips here: if a vendor claims comparable workflow impact for their CT-FFR or triage product, ask where their independent randomized evidence is.

Source: Heartflow / GlobeNewswire

[PARTNERSHIP] SimonMed and Optellum wire lung nodule malignancy risk scoring into an enterprise-wide workflow

SimonMed / Optellum | 25 August 2026

SimonMed, one of the largest US outpatient imaging providers, will route eligible incidentally detected pulmonary nodules from chest CT into Optellum's Lung Cancer Prediction AI, on top of its existing Infervision-powered volumetric sizing and growth tracking. What the tool is cleared for: Optellum LCP is FDA-cleared as decision support for malignancy risk prediction of eligible incidental nodules, not for autonomous diagnosis; every score is reviewed by a radiologist before entering the report. Evidence level: vendor materials plus the product's original clearance evidence; the announcement adds deployment scale, not new clinical evidence. Questions to ask before replicating this: which nodules are ineligible and how often eligibility filtering fails; how the LCP score changes downstream referral behavior; who audits score-to-outcome concordance at enterprise scale. Note the reimbursement driver: Optellum cites Medicare payment policy naming its product.

Source: SimonMed press release

[PARTNERSHIP] Aiforia and Stratipath integrate breast cancer risk stratification into one digital pathology platform

Aiforia / Stratipath | 25 August 2026

Helsinki-based Aiforia and Stockholm-based Stratipath announced on 25 August that Stratipath Breast, a CE-marked image-based breast cancer risk stratification tool, will be integrated into the Aiforia Clinical Platform, with joint marketing across Europe. What it is: an integration agreement, not a finished deployment; no launch date, no financial terms, and no statement yet on whether the combined workflow preserves Stratipath Breast's existing CE marking or requires recertification. Evidence level: Stratipath Breast carries peer-reviewed evidence; the integration itself carries none yet. Questions for the vendors: regulatory status of the combined workflow under IVDR; whether validation was performed on the integrated pipeline end to end; data flow and logging responsibilities when two vendors sit in one diagnostic chain.

Source: Next AI Press

[PARTNERSHIP] Philips and Imricor launch a commercial interventional MR lab for cardiac procedures

Philips / Imricor | 27 August 2026

Philips and Imricor announced a commercially available interventional MR lab combining Philips' 1.5T cardiac MRI platform, including SmartHeart AI-powered cardiac planning, with Imricor's MR-compatible mapping systems and electrophysiology catheters, enabling MR-guided cardiac ablation workflows. Cleared status: available immediately in CE-marked markets and in some configurations in the US, with further Imricor configurations pending regulatory clearances; check the exact clearance status of each component for your jurisdiction before procurement. Evidence level: vendor announcement of a product configuration; the announcement cites no comparative clinical outcomes for the integrated lab. For imaging leaders, the relevance is infrastructural: MR moving from diagnosis into the procedure room changes siting, staffing and MR-safety governance, and the wheelchair-detection paper in this issue is a reminder of how much of that governance is about the anteroom, not the magnet.

Source: Royal Philips / GlobeNewswire

SECTION 3

POST-MARKET SIGNAL

Drift | Recalls and MAUDE | Real-world monitoring | EU AI Act Article 72

ECRI opens a national reporting channel for AI errors in patient care

ECRI / PR Newswire | 25 August 2026

Patient safety organization ECRI expanded its Problem Reporting Network, the confidential device-problem channel it has run since 1972, to explicitly capture errors, malfunctions and near misses involving AI tools and AI-enabled devices, and called on providers nationwide to submit incidents. The gap it fills is the one every AI committee knows: there is no centralized, healthcare-specific mechanism tracking how often AI outputs are wrong and how often those outputs reach a patient. ECRI's own survey of 124 quality and safety leaders shows why self-report alone fails: 31% encountered an AI output they believed incorrect in the past year, 9% said an error reached a patient or care decision, and 35% simply did not know. Each report is investigated by ECRI's clinical and engineering staff, with hazard advisories issued when a flaw poses sector-wide risk. For deployers, this is a concrete addition to the incident-reporting path: alongside MAUDE and vendor complaint files, there is now a neutral place to send the near miss your radiologist caught.

Link: https://www.prnewswire.com/news-releases/ecri-expands-problem-reporting-network-...

FDA posts Class II recall of Fujifilm Synapse PACS over inaccurate ruler measurements

FDA recall record Z-3030-2026, reported by Medical Daily | 26 August 2026

The FDA posted a Class II recall record on 26 August for FUJIFILM Synapse PACS version 7.4.310: a software design flaw can make ruler measurements inaccurate on the reading workstation, with the disclosed workaround pointing at M-mode ultrasound measurements, which are central to echocardiography and obstetric monitoring. Fujifilm initiated the correction on 22 July; six units are affected across five US states and Australia, and no patient harm is reported. The governance lesson does not scale with the unit count: measurement is now a software function, a measurement defect does not announce itself the way a broken gantry does, and the affected sites found out because the manufacturer's own surveillance worked. Worth checking locally: whether your QA program would catch a workstation-side measurement error at all, and whether your PACS version is in any open recall record this quarter.

Link: https://www.medicaldaily.com/fujifilm-synapse-pacs-recall-ruler-measurement-erro...

SECTION 4

PLAYBOOK SNIPPET

One concrete step your AI committee can take this month

EU AI Act Article 10 (data governance)

Step: Add one clause to your next imaging AI RFP: the supplier discloses the composition of training and test data, including treatment phase, scanner mix and demographics, with performance reported by subgroup. Record refusal or silence as a risk item under Article 9 rather than treating it as acceptable.

Why: Training-set composition, not architecture, is what decides whether a model works on your use case, and it is the one thing a demo will never show you.

Worked example: Zielke J et al. trained a pediatric diffuse midline glioma segmentation model with post-treatment scans included and tested it externally on a prospective trial: whole-tumor Dice 0.90 vs 0.81 for the BraTS-PEDs 2024 challenge winner trained on pre-treatment data, with the largest gap on post-treatment scans, exactly where response assessment happens. AJNR Am J Neuroradiol, PMID 42648875, DOI 10.3174/ajnr.A9609.

SECTION 5

MEDIA HIGHLIGHTS

AuntMinnie | Radiology Business | The Imaging Wire | Diagnostic Imaging | ITN

Ten years after Hinton: how AI is actually changing the radiologist's job

Ars Technica / Knowable Magazine | *25 August 2026*

A long feature marking a decade since Geoffrey Hinton's replacement prediction, built around Curt Langlotz and Nina Kottler. The useful material for deployers is concrete: Kottler describes monitoring radiologist-AI agreement rates as a governance tool (if readers accept an AI known to be 95% accurate 99 times out of 100, someone talks to that reader), argues tools should report per-case confidence rather than binary outputs, and notes that a 2026 AMA survey found over a quarter of physicians received no AI training at all. Langlotz's framing stands: machine and human intelligence fail differently, which is why the radiologist's job becomes evaluating AI decisions, and why automation bias and automation complacency are now operational risks to manage, not abstractions.

Link: https://arstechnica.com/health/2026/08/ai-wont-replace-radiologists-but-it-will-...

CE-first radiology AI waits a median 17.5 months for FDA clearance; FDA-first devices get CE in 3.5

Medspark, covering npj Digital Medicine | *23 August 2026*

An npj Digital Medicine analysis of 239 AI-enabled radiology devices found a persistent regulatory asymmetry: devices that obtained a CE mark first waited a median 17.5 months for subsequent FDA clearance, while FDA-first devices obtained European authorization in a median 3.5 months. Of the 239 devices, 128 carried only a CE mark. For European deployers the practical reading is that a large share of the tools on your market will never see the FDA's independent review pathway at all, and for vendors the study is a planning roadmap: launching in Europe first buys speed now at the cost of a long second act in the US.

Link: https://medspark.ai/2026/08/23/radiology-ai-that-clears-europe-first-faces-a-17-...

FDA's generative AI discussion paper: comment window open to 19 October

Noah Intelligence, covering the FDA discussion paper of 18 August | *27 August 2026*

Coverage continued this week of the FDA discussion paper on generative AI-enabled medical devices, published 18 August with comments due 19 October 2026 under docket FDA-2026-N-7874. The elements most relevant to imaging: a risk framework graded by how independently a system acts, a competency-based premarket model with standardized benchmarks for safety behavior, generalisability and resistance to prompt injection, heavier reliance on post-market monitoring precisely because generative outputs resist complete premarket testing, and a floated voluntary Foundation Model Device Master File for model developers. ACR is preparing a radiology response and collecting member input. If your institution deploys or plans report-drafting AI, this docket is where the rules that will govern it are being drafted; the paper is not guidance and changes no rule today.

Link: https://noah-news.com/fda-seeks-input-on-new-risk-framework-for-generative-ai-me...

SECTION 6

PROFESSIONAL SOCIETIES

SIIM | ACR | RCR | Official publications and event coverage

RSNA highlights new evidence that cognitively biased prompts degrade radiology LLM accuracy

26 August 2026

RSNA published a news feature on 26 August on a Radiology: Artificial Intelligence study from the University of Toronto testing 10 LLMs on 400 radiology board-style questions. Baseline accuracy was 84.8% on text-only and 59.5% on multimodal questions; biased prompts written to sound clinically plausible (authority, complexity, anchoring) cut accuracy by up to 21.1% and 44.9% respectively, with authority bias the most damaging. Mitigation prompts recovered only part of the loss. The accompanying commentary makes the deployment point: these are not adversarial tricks but recognizably human framings, exactly what a model embedded in clinical conversation will encounter.

Source: https://www.rsna.org/news/2026/august/cognitive-biases-mislead-radiology-llms

ACR reports advocacy for NIH imaging AI programs, including the PRIMED-AI initiative

27 August 2026

ACR reported on 27 August that it and other members of the Friends of NIBIB Coalition met with NIH officials on 21 August to support the National Institute of Biomedical Imaging and Bioengineering. The discussions covered upcoming programs relevant to radiology, notably PRIMED-AI (Precision Medicine with AI: Integrating Imaging with Multimodal Data), which will fund development of AI clinical decision support combining medical imaging with other health data. ACR continues to advocate for increased NIH and NIBIB funding.

Source: https://www.acr.org/News-and-Publications/2026/advocating-for-imaging-programs-nih

SECTION 7

KEY OPINION LEADERS

LinkedIn | Public statements | Week of 23 to 29 August 2026

Amine Korchi

Radiologist, healthtech investor; Sifted Top 25 Expert, Imaging Wire Top 10 Radiology AI KOL

Authored post (28 August) announcing two milestones for ThinkSono, the AI-guided deep vein thrombosis ultrasound startup in his investment portfolio, with images showing phone-guided leg scanning in use at Chelsea and Westminster Hospital. His framing is the point: the biggest impact of AI in healthcare will not be helping specialists do the same things better, it will be redesigning how care is delivered, in this case moving DVT ultrasound acquisition away from the sonographer's room. In a separate post (27 August) he pushed back on the Swiss healthcare cost debate: a tariff hour is not an hour worked, and high radiologist productivity reflects technology and organization, not overpayment; systems should measure and reward value rather than penalize efficiency.

Source: LinkedIn

Curt Langlotz

Professor of Radiology, Medicine and Biomedical Data Science, Stanford

Shared (26 August) the Modern Healthcare feature 'What radiologists want from the next generation of AI', continuing his season of interviews on where imaging AI actually stands a decade after the replacement predictions: current tools assist rather than replace, and the radiologist's role shifts toward evaluating AI output, with automation bias as the operational risk to manage.

Source: LinkedIn

Daniel Pinto dos Santos

Radiologist and imaging informatics researcher, University Hospital Cologne / Frankfurt

Amplified (28 August) the ACR Data Science Institute's ACR-EuSoMII 2026 announcement built around four governance questions: is our AI performing as expected, who is responsible for monitoring it, how do we choose the next model, and is your governance governing. The checklist doubles as a self-test for any department's AI committee ahead of the October meeting in Crete.

Source: LinkedIn

QUALITY ASSURANCE NOTE

Compiled by Dr. Sergey Morozov from publicly available sources: peer-reviewed papers indexed in PubMed (PMIDs and DOIs verified, no preprints or arXiv), official press releases, regulatory databases and specialist media. It does not constitute medical, regulatory or financial advice. All papers have publication dates verified within 23 to 29 August 2026. Paper selection is deterministic (pinned PubMed query, Q1/ flagship journal allow-list, exclusion of pure reviews [meta-analyses kept], radiomics-only and interventional-radiology work, then a transparent ranking score). Industry items come from official press releases / regulatory notices; KOL items from public LinkedIn activity within the window.