

VEILLE INTELLIGENCE ARTIFICIELLE

Radiologie et imagerie medicale

Semaine 33 | 9 au 15 aout 2026

Dr Sergey Morozov | drsergeymorozov.com

Compile a partir de sources publiquement disponibles (publications scientifiques indexees, communiqués officiels, bases de données réglementaires).
Ne constitue pas un avis médical, réglementaire ou financier.

8

Articles évalués par les pairs

5

Annonces secteur / FDA

2

Points médiatiques

3

Leaders d'opinion

ENSEIGNEMENT DE GOUVERNANCE DE LA SEMAINE

Si une IA mammographique est en production dans votre service, séparez ce mois-ci le monitoring des rappels par filière. Dans un déploiement d'un an en mois alternés, la détection du cancer a doublé sans ralentir la lecture, mais les interprétations anormales n'ont augmenté qu'en mammographie diagnostique, 18,7 % contre 12,1 %, le dépistage restant stable. Une métrique de rappel unique et poolée aurait caché ce déplacement. Ajoutez le taux d'interprétations anormales par filière à votre dossier mensuel de monitoring ; ce même dossier devient le germe de votre dossier de surveillance post-commercialisation au titre de l'article 72 du règlement européen sur l'IA.

SECTION 0

RESUME EXECUTIF

Tendances clés, semaine du 9 au 15 août 2026

[MONITORING EN CONDITIONS REELLES] L'IA MAMMOGRAPHIQUE AUGMENTE LA DETECTION, ET LES RAPPELS, MAIS SEULEMENT EN DIAGNOSTIC

Lee SE et al. (European Radiology) ont alterné un mois sur deux, pendant un an, l'affichage d'une IA mammographique commerciale sur 4577 examens lus par quatre radiologues. La détection du cancer a environ doublé (22,3 vs 11,5 pour 1000), le temps de lecture n'a pas bougé (65,0 vs 64,3 s), et le taux d'interprétations anormales n'a augmenté qu'en mammographie diagnostique (18,7 % vs 12,1 %), pas en dépistage. Le schéma on-off est en soi un gabarit de monitoring local réutilisable.

[OPERATIONS] LA REDACTION DES COMPTES RENDUS DXA AUTOMATISEE SUR QUATRE SITES, AUDIT A L'APPUI

Lakhani P et al. (JACR) ont déployé une rédaction de comptes rendus DXA par OCR IA indépendante du constructeur dans deux sites académiques et deux sites communautaires. Le temps de création du compte rendu est passé de 3,48 à 0,87 minute, le délai de restitution communautaire a baissé d'environ quatre jours, et un audit de 400 comptes rendus contre les images sources a montré 99,9 à 100 % d'exactitude numérique tout en exposant une complétude de 45 % dans le processus historique.

[BENCHMARK] C'EST LE PROMPT, PAS LE MODELE, QUI A DECIDE SI UN LLM ETAIT UTILISABLE

Zhu J et al. (Journal of Medical Internet Research) ont testé GPT-4o, Claude 3.7 Sonnet, Qwen2.5-VL 72B et Gemma 3 27B sur l'ostéonécrose de la tête fémorale. L'AUC de détection est passée de 0,55 à 0,91 quand un compte rendu radiologique anonymisé accompagnait l'image, la fiabilité du grading d'un ICC de 0,51 à 0,97, et l'ajout d'images supplémentaires a dégradé la performance. Modèles commerciaux et open source ne diffèrent pas globalement ($P = ,83$).

[PHARE] LA PATHOLOGIE 3D DEVIENT LISIBLE : LE TRIAGE LAISSE LE PATHOLOGISTE AUX COMMANDES

Gao G et al. (Nature Biomedical Engineering) présentent TRICARE, qui score chaque niveau 2D d'un volume de pathologie 3D grâce au contexte des profondeurs voisines et route les coupes les plus à risque vers le pathologiste pour le diagnostic final. En stratification du risque sur biopsies prostatiques et en dépistage de l'oesophage de Barrett, l'approche montre un potentiel supérieur à l'histopathologie sur lames, qui n'échantillonne que moins de 1 % du tissu.

[POINT DE VIGILANCE] LE SIGNAL VIT DANS LES DONNEES DESAGREGUES

Trois des resultats de la semaine n'existent que parce que quelqu'un a separe ce que la plupart des services poolent. Le taux d'interpretations anormales en mammographie a bouge en diagnostic mais pas en depistage. Les défauts de completude DXA se logeaient dans les sites communautaires, pas academiques. La fiabilite de grading des LLM est passee d'un ICC de 0,49 a 0,97 sur la seule configuration d'entree. Des metriques agregees auraient masque les trois. Concevez votre monitoring pour desagreger par filiere, par site et par configuration des le premier jour, ou acceptez que votre tableau de bord moyenne le probleme jusqu'a le faire disparaître.

SECTION 1

ARTICLES EVALUES PAR LES PAIRS

Source : PubMed | Indexation verifiee | Aucun preprint | 9 au 15 aout 2026

Impact of AI assistance on reading time, cancer detection rate, and abnormal interpretation rate in screening and diagnostic mammography: a prospective alternating-month study.

Lee SE, et al. | *European radiology* | 2026-08-13

Avant d'etendre l'IA mammographique du depistage vers la liste de travail diagnostique, verifiez ce qu'elle fait a votre taux d'interpretations anormales. Quatre radiologues d'une meme institution ont lu 4577 mammographies consecutives sur 12 mois, avec un systeme d'IA commercial affiche puis masque un mois sur deux, un schema prospectif on-off qui fait aussi office de gabarit de monitoring. La detection du cancer etait plus elevee avec l'IA et le temps de lecture n'a pas bouge, mais les interpretations anormales n'ont augmente que dans la filiere diagnostique.

Indicateurs cles: 4577 examens, age moyen 51,7 ans. Taux global de detection du cancer 22,3 vs 11,5 pour 1000 ($p = 0,005$). Taux d'interpretations anormales en depistage 9,5 % vs 8,4 % ($p = 0,293$), en mammographie diagnostique 18,7 % vs 12,1 % ($p = 0,026$). Temps de lecture moyen 65,0 vs 64,3 s ($p = 0,723$) dans le sous-ensemble de 2917 examens lus en 5 minutes ou moins.

Pertinence clinique: Une metrique de rappel unique et poolee aurait totalement manque le deplacement observe dans la filiere diagnostique ; l'effet n'apparait que si depistage et diagnostic sont surveilles separement. Pour votre comite IA, cela change le monitoring : ajoutez le taux d'interpretations anormales par filiere au dossier mensuel de surveillance post-commercialisation (article 72 du reglement europeen sur l'IA), et retenez le schema on-off en mois alternes comme gabarit local peu couteux pour mesurer l'effet d'un outil sur vos propres lecteurs au titre de l'article 14 (controle humain).

PMID 42593499 DOI 10.1007/s00330-026-12793-0

Prompt Configurations for Multimodal Large Language Models in Diagnosing and Staging Osteonecrosis of the Femoral Head: Multimodel Retrospective Observational Diagnostic Study.

Zhu J, et al. | *Journal of medical Internet research* | 2026-08-10

N'acceptez aucune evaluation d'un LLM multimodal, d'un fournisseur ou de votre propre equipe, qui ne fige pas la configuration exacte du prompt. Quatre modeles (GPT-4o, Claude 3.7 Sonnet, Qwen2.5-VL 72B, Gemma 3 27B) ont diagnostique et stadifie l'osteonecrose de la tete femorale sur radiographies et IRM selon trois configurations d'entree. Ajouter a l'image un compte rendu radiologique anonymise a fait passer la detection d'un niveau proche du hasard a un niveau solide, tandis qu'ajouter davantage d'images a degrade les modeles.

Indicateurs cles: 159 patients radiographies (318 observations par hanche) et 170 patients IRM (340 observations), monocentrique. AUC de detection 0,55 (ET 0,03) avec image seule vs 0,91 (ET 0,01) avec image plus compte rendu. Exactitude de stadification 0,65, 0,78 et environ 0,59 pour image seule, image plus compte rendu et entree multi-images. Fiabilite du grading ARCO : ICC 0,51, 0,97 et 0,49 respectivement, contre un ICC inter-observateur chirurgien de 0,70. Modeles commerciaux vs open source : pas de difference globale significative ($P = ,83$).

Pertinence clinique: La performance est ici une propriete du modele plus sa configuration d'entree, pas du modele seul, et la fiabilite du grading bascule d'inutilisable a quasi parfaite sur les memes cas. Pour votre comite IA, cela change les tests de recette : le rapport de test doit consigner le prompt et la configuration d'entree deployes, et tout changement de cette configuration redeclenche le test (article 15, exactitude et robustesse). La parite entre modeles open source et commerciaux affaiblit aussi tout argument de gouvernance fonde sur la reputation du fournisseur.

PMID 42573734 DOI 10.2196/92919

Vendor-Agnostic Multisite Automated Dual-Energy X-Ray Absorptiometry Reporting Using Artificial Intelligence-Based Optical Character Recognition: Impact on Workflow Efficiency and Accuracy.

Lakhani P, et al. | *Journal of the American College of Radiology* | 2026-08-13

Si vous automatisez la rédaction de comptes rendus, voici le dossier de déploiement à copier. Un pipeline d'OCR par IA, indépendant du constructeur, a extrait les mesures DXA des images DICOM et rédige des comptes rendus complets dans quatre sites d'un même système de santé, deux académiques et deux communautaires, avec une évaluation opérationnelle avant-après et un audit d'exactitude de 400 comptes rendus cotés contre les images sources plutôt que contre le compte rendu antérieur.

Indicateurs clés: Temps médian de création du compte rendu passe de 3,48 à 0,87 minute (académique) et de 1,42 à 0,63 minute (communautaire) ; délai médian de restitution de 2,40 à 0,96 heure (académique) et de 133,82 à 42,10 heures, soit environ quatre jours, dans les sites communautaires (tous $P < ,0001$). Brouillons IA vs comptes rendus originaux : exactitude numérique 99,9 % vs 99,4 % ($P = ,022$) et 100 % vs 99,6 % ($P = ,031$) ; complétude dans les sites communautaires 100 % vs 45 % ($P < ,001$) ; exactitude diagnostique d'au moins 99,5 % dans toutes les cohortes.

Pertinence clinique: La partie transférable est le schéma d'audit : l'exactitude a été définie contre les images sources, ce qui a révélé que la référence humaine était elle-même incomplète dans les sites communautaires. Pour votre comité IA, cela change recette et monitoring : définissez la concordance brouillon contre source comme métrique de recette pour tout outil de rédaction, puis continuez à l'échantillonner chaque mois comme KPI post-commercialisation (article 72 ; clause d'évaluation de la performance d'ISO/IEC 42001).

PMID 42595274 DOI 10.1016/j.jacr.2026.08.007

Adapted foundation models for breast MRI triaging in contrast-enhanced and non-contrast-enhanced protocols.

Nguyen TT, et al. | *European radiology* | 2026-08-13

Le triage des IRM mammaires abrégées par modèle de fondation approche d'un rule-out utilisable : à 97,5 % de sensibilité, environ un examen sur cinq sans anomalie immédiatement actionnable pourrait être repriorisé, avec une performance quasi identique avec ou sans injection de contraste. Un medical slice transformer basé sur DINOv2 a été testé sur quatre protocoles abrégés pour exclure les examens BI-RADS 4 ou plus.

Indicateurs clés: 1847 examens IRM unilatéraux (377 BI-RADS 4 ou plus) d'un jeu de données interne plus 924 examens externes (Duke) ; 1448 patientes, âge moyen 49 ans. AUC 0,77 (ET 0,04) pour T1 soustraction plus T2 et 0,74 (ET 0,04) pour DWI1500 plus T2, sans différence significative entre protocoles. Spécificité à 97,5 % de sensibilité : 19 % avec contraste, 17 % sans contraste. Validation externe du protocole T1sub : AUC 0,77 ; 88 % des cartes d'attention jugées bonnes ou modérées.

Pertinence clinique: Le détail cliniquement important est la concentration des faux négatifs : surtout des lésions de moins de 10 mm et des rehaussements sans masse, précisément les lésions que les protocoles abrégés gèrent le moins bien. La quasi parité des protocoles sans contraste compte pour les programmes de dépistage qui pesent l'exposition au gadolinium, et l'AUC externe de 0,77 s'est maintenue au niveau du résultat interne.

PMID 42593493 DOI 10.1007/s00330-026-12782-3

Deep-learning triage of three-dimensional pathology datasets for comprehensive and efficient pathologist assessments.

Gao G, et al. | *Nature biomedical engineering* | 2026-08-12

La pathologie 3D non destructive image la totalité de la biopsie au lieu du moins de 1 % de tissu échantillonné par une lame standard, mais les volumes produits sont trop lourds pour une revue manuelle. TRICARE, un cadre de triage par apprentissage profond, score chaque niveau 2D du volume en s'appuyant sur le contexte des niveaux de profondeur voisins et remet au pathologiste les coupes les plus à risque, en lui laissant le diagnostic final. Il a été évalué pour la stratification du risque sur biopsies prostatiques et le dépistage de la dysplasie dans l'œsophage de Barrett.

Indicateurs clés: L'abstract rapporte les comparaisons de manière qualitative, sans valeurs numériques : le scoring avec contexte de profondeur surpasse les modèles fondés sur des niveaux 2D isolés, et la pathologie 3D triée par IA montre un potentiel d'amélioration de la détection des maladies à haut risque par rapport à l'histopathologie 2D sur lames, tout en optimisant la charge du pathologiste. Chaque coupe 2D standard représente moins de 1 % du volume de la biopsie.

Pertinence clinique: Retenu pour son mérite scientifique : c'est la voie la plus crédible publiée à ce jour pour faire entrer la pathologie 3D en pratique, précisément parce qu'elle ne retire pas le pathologiste. Trier un format de données illisible de manière exhaustive est un problème d'adoption différent du remplacement d'un lecteur, et les deux cas d'usage, stratification prostatique et surveillance de Barrett, sont des contextes où le sous-échantillonnage 2D a un coût clinique connu.

PMID 42587049 DOI 10.1038/s41551-026-01760-1

Deep Learning-Based Enhancement of Already Diagnostic-Quality MRI for Alzheimer's Disease Classification: Effects on Model Performance and Training Data Requirements.

Zhang Z, et al. | *Journal of magnetic resonance imaging* | 2026-08-13

Le rehaussement d'image s'achete d'habitude pour sauver des images mediocres. Ici, un outil de rehaussement par apprentissage profond autorise par la FDA (SubtleHD) a ete applique a des IRM cerebrales T1 deja de qualite diagnostique avant l'entrainement de classifieurs de la maladie d'Alzheimer, ameliorant a la fois la performance de classification et l'efficience en donnees, ce qui suggere que les definitions conventionnelles de la qualite d'image sous-estiment l'information disponible pour l'apprentissage automatique.

Indicateurs cles: 2293 examens ADNI (1605 entrainement, 229 validation, 459 test interne) et 270 examens externes NACC.

ResNet34 : exactitude de 85,2 % a 88,7 % et macro-AUC de 0,951 a 0,968 avec rehaussement (les deux significatifs) ; differences non significatives pour DenseNet121 ($p = 0,18$ et $p = 0,29$). L'entrainement sur 70 % du jeu rehausse egale le jeu standard complet. Test externe : exactitude 49,2 % et macro-AUC 0,679 pour le modele entraine en standard vs 63,0 % et 0,819 pour le modele entraine sur images rehaussees, qui garde un avantage sur les images externes non rehaussees (macro-AUC 0,772).

Pertinence clinique: Deux lectures honnetes coexistent : le rehaussement a extrait un signal que les definitions conventionnelles de qualite d'image ne voient pas, et la performance externe des deux bras, de 49 % a 63 % d'exactitude, montre la part du resultat interne qui n'a pas voyage. Le resultat d'efficience en donnees, egaler la performance du jeu complet avec 70 % des donnees, est le plus interessant en pratique pour les equipes qui entrainent sur des cohortes locales limitees.

PMID 42593101 DOI 10.1002/jmri.70507

Prediction of Central Lymph Node Metastasis in Papillary Thyroid Microcarcinoma Using a Deep Learning Radiomics Model Based on SAM3 Automatic Segmentation of Ultrasound Images: A Multicenter Cohort Study.

Song J, et al. | *Academic radiology* | 2026-08-12

Dans le microcarcinome papillaire de la thyroide, la decision operatoire depend souvent de metastases ganglionnaires centrales occultes. Un modele de radiomique par apprentissage profond sur echographie, construit sur une segmentation entierement automatique par SAM3, a predit la metastase ganglionnaire centrale a travers quatre centres d'entrainement et un site externe, en supprimant l'etape de contourage manuel qui tenait ces modeles a l'ecart de la pratique courante.

Indicateurs cles: 1859 patients de quatre centres (1674 entrainement, 185 validation) plus 140 d'un etablissement externe, 1999 evalues au total. Dice de segmentation SAM3 : 0,882 (validation) et 0,859 (test externe). AUC du modele combine : 0,901 (entrainement), 0,875 (validation), 0,853 (test externe), contre 0,779 a 0,825 pour la radiomique seule et 0,812 a 0,844 pour l'apprentissage profond seul. L'analyse en courbe de decision soutient l'utilite clinique.

Pertinence clinique: Retenu pour son merite clinique : l'AUC externe de 0,853 avec segmentation entierement automatique est le chiffre qui compte, car les pipelines radiomiques a ROI manuelles ont echoue de maniere repetee au contact de la pratique courante. Savoir si une discrimination de 0,85 suffit a changer les decisions de curage ganglionnaire central prophylactique est la question clinique que l'article laisse ouverte.

PMID 42586895 DOI 10.1016/j.acra.2026.07.064

From voxel discovery to regional interaction: A multi-level interpretable framework for Alzheimer's disease diagnosis.

Shu K, et al. | *Medical image analysis* | 2026-08-12

L'interpretabilite integree a l'architecture plutot qu'ajoutee apres coup : un cadre en trois etages pour le diagnostic de la maladie d'Alzheimer sur IRM structurelle, qui decouvre les voxels pertinents par masquage evidentiel hierarchique, les agrege en regions anatomiquement coherentes, puis modelise les interactions entre regions par un reseau d'attention en graphe, rendant le raisonnement diagnostique visible a chaque niveau plutot que reconstruit post hoc.

Indicateurs cles: L'abstract ne formule que des affirmations comparatives, sans valeurs numeriques : la performance sur les jeux ADNI et AIBL est decrite comme comparable aux modeles boite noire de l'etat de l'art, tandis que le cadre reconstruit de lui-meme des trajectoires neuropathologiques etablies. Le code source est public.

Pertinence clinique: Retenu pour son merite scientifique : les cartes de saillance post hoc ont echoue de maniere repetee a convaincre les cliniciens, a cause des fuites d'arriere-plan et de l'incoherence anatomique, deux defauts que cette architecture traite structurellement plutot que cosmetiquement. Si l'affirmation de performance comparable survit a des tests independants, les architectures nativement interpretables deviennent une alternative realiste aux couches d'explication ajoutees apres l'entrainement.

PMID 42585730 DOI 10.1016/j.media.2026.104243

SECTION 2

SECTEUR ET REGLEMENTATION

*Sources : Registres officiels | Communiqués de presse | Bases FDA***[REGULATION] Medicare approuve un paiement additionnel pour nouvelle technologie pour CARE CT Body Multi-Triage d'Aidoc***Aidoc | 13 août 2026*

Ce que cela couvre : la CMS a finalisé un New Technology Add-on Payment allant jusqu'à 137,53 dollars par cas hospitalisé éligible, environ 65 % du coût moyen de la technologie selon le résumé de l'ACR, à partir du 1er octobre 2026 pour trois ans (code XEZ5XKC). Ce pour quoi l'outil est autorisé : un triage de flux qui signale des anomalies urgentes suspectées sur les scanners thorax, abdomen et pelvis ; une autorisation de triage priorise la liste de travail, elle ne diagnostique pas. Niveau de preuve : le paiement est passé par la voie alternative, qui dispense les dispositifs breakthrough de prouver une amélioration clinique substantielle, et la CMS supprime cette voie à partir de l'exercice 2028. Questions avant de vous y fier : quels cas hospitalisés sont réellement éligibles chez vous, comment le codage sera capturé, et ce que montrent vos propres taux de surclassement et de concordance, car le remboursement change l'économie mais pas la charge de validation locale.

Source: [PR Newswire](#)**[PARTNERSHIP] Douze systèmes de santé américains créent le Diagnostic AI Consortium avec Aidoc***Diagnostic AI Consortium | 11 août 2026*

De quoi il s'agit : douze systèmes, dont Cedars-Sinai, Mount Sinai, Northwell, Northwestern Medicine et Houston Methodist, couvrant environ 20 millions de patients par an, s'engagent sur des standards partagés d'évaluation, de déploiement et de gouvernance de l'IA diagnostique, avec Aidoc comme infrastructure technique et de premiers résultats promis pour 2027. Niveau de preuve : une annonce d'intention, aucune donnée pour l'instant. Pourquoi cela compte pour la gouvernance : des standards d'évaluation et de monitoring côté deployeur sont exactement ce que l'Union européenne fait naître par la réglementation ; aux États-Unis ils émergent par consortium. La question à suivre : des standards construits sur la plateforme d'un seul fournisseur seront-ils portables, et la mesure des résultats sera-t-elle indépendante du fournisseur d'infrastructure évalué.

Source: [Radiology Business](#)**[FDA] Leica Biosystems annonce plusieurs autorisations FDA 510(k), dont le premier logiciel de contrôle qualité assisté par IA en pathologie numérique***Leica Biosystems | 11 août 2026*

Autorise pour : Aperio iQC DX, contrôle qualité autonome assisté par IA pour la pathologie numérique clinique, qui détecte automatiquement six artefacts d'acquisition pendant que les lames sont encore sur le scanner (bulles d'air, traits de feutre, tissu coupé, tissu manquant, zones floues, striures), aux côtés de l'autorisation du scanner Aperio GT 180 DX. Niveau de preuve : annonce du fabricant ; aucune étude de performance publiée n'est citée. Questions au fournisseur : sensibilité et spécificité par type d'artefact, impact mesure sur les taux de rémunération dans un laboratoire en production, robustesse selon les colorations et les parcs de scanners, et journalisation exportable des signalements. Le monitoring qu'il vous faut de toute façon : suivez dès le premier jour le taux de rémunération et les artefacts manquants signalés par les pathologistes ; un outil de contrôle qualité qui sous-signe en silence est pire que pas d'outil.

Source: [Leica Biosystems](#)**[FDA] CliniComp obtient l'autorisation FDA 510(k) d'un visualiseur PACS intégré au dossier patient***CliniComp | 10 août 2026*

Autorise pour : un visualiseur PACS en tant que système de gestion et de traitement d'images médicales (MIMPS), offrant visualisation, manipulation et traitement de qualité diagnostique au sein du dossier patient New Era de l'éditeur. Ce que ce n'est pas : une autorisation de visualiseur ne porte aucune revendication de détection ni de triage, quoi que suggère le discours environnant sur la priorisation native par IA des cas aigus. Niveau de preuve : matériaux du fournisseur uniquement. Questions au fournisseur : quelles fonctions de priorisation par IA sont dans le périmètre autorisé et lesquelles sont des fonctions de flux non réglementées, quelles données de performance soutiennent les revendications de priorisation, et la fidélité d'affichage a-t-elle été validée par modalité, puisqu'un visualiseur change ce que chaque lecteur voit sur chaque cas.

Source: [HIT Consultant](#)

[PARTNERSHIP] Coreline Soft et Optellum annoncent une intention de collaboration sur le parcours du cancer pulmonaire

Optellum et Coreline Soft | 13 aout 2026

De quoi il s'agit : une intention declaree de collaborer, pas un accord de distribution signe, pour rendre le logiciel de detection AVIEW Lung Nodule CAD de Coreline Soft, autorise par la FDA, disponible via la plateforme LungOS d'Optellum, qui porte ses propres autorisations de CADx (evaluation du risque) et de gestion du parcours nodulaire. Le point reglementaire a retenir pour l'acheteur : la detection (CAdE) et le score de risque (CAdx) sont des revendications distinctes, autorisees sur des preuves distinctes ; les cabler dans un meme flux ne cree pas un parcours autorise de bout en bout. Questions aux fournisseurs : quelles revendications s'appliqueront au flux combine dans chaque juridiction, et quelles preuves couvrent le passage de la sortie de detection d'un systeme a l'entree de risque de l'autre.

Source: [PR Newswire](#)

SECTION 3

SIGNAL DE SURVEILLANCE APRES COMMERCIALISATION

Derive | Rappels et MAUDE | Surveillance en conditions reelles | Article 72 du reglement europeen sur l'IA

La SIIM publie des strategies pratiques de monitoring post-production de l'IA en radiologie

SIIM | 13 aout 2026

La SIIM a publie le compte rendu de son webinaire sur le monitoring des IA deployees, qui se lit comme un rapport de terrain sur ce que des obligations de type article 72 donnent en pratique. Penn Medicine fait tourner un comite de gouvernance couvrant tout le cycle, de l'approbation des demonstrations au monitoring post-deploiement. Dasa suit la derive d'un modele interne de debruitage IRM en surveillant dans le temps les distributions de voxels en entree et la similarite entree-sortie. Greensboro Radiology a refuse des outils autorises par la FDA pour cause de performance decevante en conditions reelles et surveille les algorithmes deployes en comparant leurs predictions aux impressions des radiologues extraites des comptes rendus. Message commun : l'autorisation est le debut de l'obligation de preuve, pas sa fin.

Lien: <https://siim.org/artificial-intelligence/effective-strategies-for-post-productio...>

Les radiologues rapportent peu de benefices tangibles de l'IA en imagerie mammaire

Radiology Business | 11 aout 2026

Une enquete de Clinical Imaging aupres de 215 membres de la Society of Breast Imaging, couverte le 11 aout, est un signal d'adoption en conditions reelles a archiver : environ la moitie exercent dans des organisations ayant deploye une IA de detection du cancer du sein, mais seule une minorite rapporte une amelioration mesurable du taux de rappel, du taux de biopsie ou de l'epuisement professionnel. Les taux de rappel rapportes baissent d'environ 35 % chez les utilisateurs actuels contre les attentes des non-utilisateurs, les biopsies inutiles de seulement 9 % environ, et 29,4 % des utilisateurs constatent un effet notable sur l'epuisement. L'adoption court devant le benefice mesure ; les services devraient suivre leurs propres deltas de rappel et de biopsie plutot que de supposer ceux des projections.

Lien: <https://radiologybusiness.com/topics/artificial-intelligence/radiologists-report...>

SECTION 4

FICHE PRATIQUE

*Une mesure concrete pour votre comite IA ce mois-ci***Article 72 du reglement europeen sur l'IA (surveillance post-commercialisation) ; clause d'evaluation de la performance d'ISO/IEC 42001**

Mesure: Echantillonnez 20 sorties d'IA par mois contre le compte rendu final signe et suivez un seul chiffre : le taux de concordance. Tracez-le chaque mois et agissez sur la tendance, pas sur une baisse isolee.

Pourquoi: La derive est invisible dans l'exactitude agreee et visible dans une courbe mensuelle ; une serie commencee maintenant est un historique que vous possederez deja quand l'obligation de monitoring s'appliquera.

Exemple concret: Lakhani P et al. ont audite 400 comptes rendus DXA rediges par IA contre les images sources, et non contre les comptes rendus anterieurs, dans quatre sites : 99,9 a 100 % d'exactitude numerique, et l'audit a revele un taux de completude de 45 % dans la reference humaine des sites communautaires. JACR, PMID 42595274, DOI 10.1016/j.jacr.2026.08.007.

SECTION 5

POINTS MEDIATIKUES

*AuntMinnie | Radiology Business | The Imaging Wire | Diagnostic Imaging | ITN***Les facteurs qui rendent les radiologues moins vulnerables aux erreurs des grands modeles de langage***Radiology Business | 12 aout 2026*

Couverture d'une etude de Radiology (publiee le 11 aout) dans laquelle 10 lecteurs ont interprete 100 cas thoraciques avec et sans assistance LLM, randomises entre un modele a haute exactitude (base GPT-5, 76 %) et un modele a faible exactitude (base GPT-4o, 27 %). La confiance du modele (OR 3,82) et l'expertise du lecteur (OR 2,06) etaient independamment associees a une collaboration reussie, et l'expertise protegeait contre l'acceptation de suggestions incorrectes (OR 0,54). Le resultat inconfortable : une meilleure qualite de justification augmentait l'acceptation des suggestions incorrectes (OR 1,71), les explications persuasives coupant dans les deux sens.

Lien: <https://radiologybusiness.com/topics/artificial-intelligence/factors-make-radiol...>

L'American Board of Radiology adopte une approche prudente de l'IA*Radiology Business | 13 aout 2026*

L'ABR a expose sa position sur l'IA dans une mise a jour du 12 aout : aucune IA n'est actuellement utilisee dans les decisions de certification ni dans la creation du contenu des examens, et les decisions de notation restent le fait d'experts humains qualifies. Le conseil a mis en place un groupe consultatif IA pour la gouvernance et l'ethique et un comite au niveau des trustees sur les cas d'usage en evaluation, avec des garde-fous etablis avant toute adoption. Qu'un organisme certificateur publie sa frontiere de controle humain est en soi un signal de gouvernance pour la specialite.

Lien: <https://radiologybusiness.com/topics/artificial-intelligence/american-board-radi...>

SECTION 6

SOCIETES SAVANTES

SIIM | ACR | RCR | Publications et couverture officielles

L'ACR publie son resume detaille de la regle finale IPPS FY2027

13 aout 2026

L'ACR a publie le 13 aout son resume detaille pour ses membres de la regle finale CMS FY2027 sur les paiements hospitaliers (IPPS) : revalorisation de 2,3 % des paiements hospitaliers, environ 779 millions de dollars de paiements pour les cas impliquant de nouvelles technologies, un paiement additionnel pour un dispositif de triage d'imagerie par IA, et la fin de la voie alternative des NTAP a partir de l'exercice 2028, apres quoi toute technologie demandant un paiement additionnel devra demontrer une amelioration clinique substantielle. Pour les fournisseurs comme pour les acheteurs d'IA, la barre de preuve du remboursement hospitalier americain vient de monter.

Source: <https://www.acr.org/News-and-Publications/2026/fy2027-ipp-final-rule-detailed-summary>

SECTION 7

LEADERS D'OPINION

LinkedIn | Declarations publiques | Semaine du 9 au 15 aout 2026

Curt Langlotz

Professeur de radiologie, de medecine et de science des donnees biomedicales, Stanford ; directeur du Stanford AIMI

Semaine de curation active (publications des 12 et 15 aout) : il a amplifie les travaux d'Eric Horvitz sur le controle par confiance de la generation de comptes rendus cliniques, une approche a seuil ou un point de fonctionnement sur la courbe des faux positifs conditionne ce que le LLM ecrit dans le brouillon ; repartage le nouveau travail du Stanford AIMI sur un modele d'IA en echocardiographie pediatrique pour les cardiopathies congenitales publie dans Circulation ; et relaye l'analyse de Nature Health portant sur 617 827 conversations de sante avec Microsoft Copilot, en notant que Copilot est la petite plateforme a cote des 40 millions de requetes de sante rapportees pour ChatGPT.

Source: [LinkedIn](#)

Amine Korchi

Radiologue et investisseur HealthTech, Geneve

Publication du 13 aout : il signale Vocca comme une startup a suivre au moment ou elle s'etend de San Francisco a New York, dans la continuite de son fil sur les agents vocaux d'IA qui absorbent la charge administrative des cabinets medicaux plutot que les taches diagnostiques, le segment de l'IA medicale ou la friction de deploiement et la charge reglementaire sont les plus faibles.

Source: [LinkedIn](#)

Daniel Pinto dos Santos

Radiologue, hopitaux universitaires de Cologne et Francfort ; conseil d'administration d'EuSoMII

Repartage du 12 aout : il a amplifie l'annonce par Marius George Linguraru de la session conjointe EuSoMII et MICCAI sur l'avenir de l'IA en imagerie medicale au congres annuel d'EuSoMII a Heraklion en octobre, avec pour theme affiche l'integration sure des outils d'IA en pratique clinique, signe de la convergence des societes cliniques et techniques vers la securite du deploiement plutot que le developpement de modeles.

Source: [LinkedIn](#)

NOTE D'ASSURANCE QUALITE

Compile par le Dr Sergey Morozov a partir de sources publiquement disponibles : articles evalues par les pairs indexes dans PubMed (PMID et DOI verifiees, aucun preprint ni arXiv), communiquees officielles, bases reglementaires et medias specialises. Ne constitue pas un avis medical, reglementaire ou financier. Tous les articles ont des dates verifiees dans la fenetre du 9 au 15 aout 2026. La selection est deterministe (requete PubMed fixe, liste de revues Q1/phases, exclusion des revues pures [meta-analyses conservees], de la radiomique seule et de la radiologie interventionnelle, puis un score de classement transparent). Les annonces secteur proviennent de communiquees officiels ; les items KOL de l'activite publique LinkedIn dans la fenetre.