

Explainable and Privacy-Preserving Multimodal Machine Learning for Scalable Cancer Diagnosis: A Review

Frejana Nielsen¹, Osker Lindberg²

¹Nordic Institute of Computer Technology, DENMARK
Scandinavian University of Data Science, SWEDEN

Abstract

Cancer struck an estimated 20.6 million people and killed 9.8 million worldwide in 2024, a burden projected to grow to 34 million new cases and 17.5 million deaths annually by 2050, yet only one in three countries currently reports high-quality cancer incidence data and only one in four reports reliable mortality data. Machine learning promises to help close this gap, but only if three requirements are met at once: clinicians need explanations they can act on and defend, patient data protected by regulations such as HIPAA and the GDPR cannot simply be pooled across institutions or borders, and the diagnostic signal that actually distinguishes malignant from benign disease is often split across multiple data modalities — histopathology, radiological imaging, and structured clinical variables — that no single-modality model can fully exploit. This review synthesizes five studies addressing exactly these requirements. On explainability, two independent studies applied to the same Wisconsin breast-cancer cytological benchmark two years apart reach 96% and 97.07% accuracy respectively, the more recent adding SHAP-based feature attribution and systematic class-imbalance handling within a predictive-preventive-personalized-medicine framework. On privacy-preserving multimodal fusion, a federated framework combining skin-lesion images with tabular metadata reaches 92.32% accuracy while losing only 2 percentage points under substantial missing-modality conditions, and a tri-modal federated framework fusing ultrasound, MRI, and mammography reaches 93.10% fused accuracy while cutting communication bandwidth by 84.9% through reduced-precision weight transmission. A sixth study, whose framing a multimodal, privacy-preserving, scalable diagnostic framework across healthcare systems matches this review's scope unusually closely, could not be independently verified on its quantitative results and is discussed with that limitation stated explicitly. Read together, these studies show genuine, measurable progress on explainability, privacy preservation, and multimodal fusion individually, but no study reviewed here demonstrates all three simultaneously at the scale the global cancer-data gap actually requires.

Keywords: Cancer Diagnosis, Explainable AI, Privacy-Preserving Machine Learning, Federated Learning, Multimodal Fusion, SHAP, Healthcare AI, Scalability

Introduction

Cancer caused an estimated 20.6 million new cases and 9.8 million deaths worldwide in 2024, giving roughly one in five people a lifetime risk of developing the disease, according to the International Agency for Research on Cancer's GLOBOCAN 2024 estimates; that burden is

projected to rise to 34 million new cases and 17.5 million deaths annually by 2050 — increases of 67% and 80% respectively — driven substantially by population aging and growth in low- and middle-resource countries. China alone accounts for roughly one in four global cases and deaths, while Europe, with under 10% of the world's population, accounts for over a fifth of global incidence, illustrating how unevenly both the disease burden and the diagnostic infrastructure to address it are distributed. The more consequential figure for this review, however, is a data-quality one: only one in three countries currently reports high-quality cancer incidence data, and only one in four reports reliable mortality data, which means the single biggest obstacle to understanding and addressing the global cancer burden is not a lack of statistical methods but a lack of usable, comparable data reaching the systems that could analyze it. This is precisely the problem a genuinely scalable, privacy-preserving, multimodal diagnostic system would need to address: making it possible for diagnostic models to learn from data that, for reasons of both regulation and basic institutional trust, cannot simply be centralized [1].

Three converging requirements

A cancer-diagnosis model that cannot explain its reasoning is difficult for an oncologist or pathologist to act on with confidence, and difficult to defend in the kind of clinical or legal scrutiny a cancer diagnosis can invite — this is the explainability requirement. A model that requires pooling raw patient imaging and records across hospitals, health systems, or national borders runs directly into regulations such as the United States' HIPAA Privacy Rule and the European Union's GDPR, both of which treat health data as requiring special protection — this is the privacy-preservation requirement. And a model trained on only one data type — only histopathology slides, or only radiological images, or only structured clinical variables — discards diagnostic signal that a specialist reading multiple data sources together would use, and cannot straightforwardly scale to the multi-institution, multi-modality reality of how cancer is actually diagnosed — this is the multimodality and scalability requirement. This review is organized around these three requirements because, as Section 6 argues, no single study reviewed here satisfies all three simultaneously, and understanding why is more useful than reviewing each pillar in isolation.

Two contrasting precedents: promise and caution

Two widely documented episodes illustrate both why this problem is worth solving and why explainability specifically cannot be treated as optional. In 2020, Google Health and DeepMind published a large international evaluation in *Nature* showing an AI system could match or exceed expert radiologists at detecting breast cancer from mammograms across screening populations in the United States and the United Kingdom, reducing both false positives and false negatives relative to the radiologists' original reads — a rigorously validated demonstration that machine learning can genuinely improve on human performance at a core cancer-diagnosis task, not merely approximate it. In the same period, however, an internal IBM document reported by STAT News in 2018 revealed that IBM Watson for Oncology, deployed at hospitals internationally, had in multiple documented instances recommended cancer treatments its own physician trainers and customers characterized as unsafe or incorrect, a failure attributed substantially to the system being trained on a small number of hypothetical cases synthesized by a limited group of specialists at one institution rather than on diverse real-world patient data, and to clinicians' inability to see why the system had

recommended what it recommended. Read together, these two episodes frame this review's three requirements concretely: the Google Health result shows what rigorously validated, high-quality training data can achieve, while the IBM Watson episode shows what happens when a system's reasoning is opaque and its training data's limitations are not visible to the clinicians relying on it — precisely the failure mode Section 3's explainability-focused studies and Section 4's multi-institutional federated studies are, in different ways, trying to prevent from recurring [2-6].

Review scope and methodology

This review synthesizes five studies, selected without an artificial recency cutoff so that the most relevant and highest-quality available evidence could be included regardless of publication year, spanning 2023 to 2026. Two studies address explainability using the same well-established Wisconsin breast-cancer cytological benchmark (Sections 3.1–3.2); two address privacy-preserving multimodal federated learning across different cancer types and modality combinations (Sections 4.1–4.2); and one, whose stated scope overlaps substantially with this review's own title, could not be independently verified beyond its confirmed title and authorship and is discussed accordingly, with that limitation stated explicitly rather than treated as equivalent to the four independently verified studies (Section 5.2).

Background: Why Cancer Diagnosis Is Becoming a Multimodal, Multi-Institutional Problem

The modalities involved and why they are usually siloed

Cancer diagnosis has always drawn on multiple evidence types — a radiologist's imaging read (mammography, ultrasound, MRI, or CT), a pathologist's tissue interpretation under a microscope or, increasingly, from a digitized whole-slide image, and a clinician's structured assessment of patient history, lab values, and genomic or molecular markers where available — but machine-learning systems have historically been built around just one of these at a time. This is partly a data-availability problem (assembling a well-labeled, single-institution dataset for one modality is already difficult, and combining several modalities from several institutions multiplies that difficulty) and partly an organizational one: radiology, pathology, and oncology departments typically maintain separate systems and workflows even within one hospital, let alone across the multiple institutions a genuinely large and diverse training dataset would require.

Two forces pushing toward multimodal, multi-institutional systems

Two forces are now pushing the field toward genuinely multimodal, multi-institutional systems regardless. First, single-modality models trained on one hospital's data routinely underperform when applied to another hospital's patient population, equipment, or imaging protocol, which is itself a form of the generalization problem and a direct motivation for training across more institutions rather than fewer. Second, and more consequentially for the field's technical direction, regulations such as HIPAA and the GDPR make it legally and organizationally difficult to simply centralize patient imaging and records across institutions to train a larger, more diverse model the conventional way — which is exactly the gap federated learning, the technique underlying Sections 4.1 and 4.2 of this review, is designed to close by keeping raw patient data at its institution of origin and sharing only model updates [7-11].

Explainable Machine Learning for Cancer Diagnosis

A comparative baseline on a well-established benchmark

Siddiqui, Rahman, Kabir, Mahmud, Rashid, and Shammah (2023) compare Random Forest against Gaussian Naive Bayes on the Wisconsin Diagnostic Breast Cancer dataset — 569 samples described by 30 real-valued morphometric features derived from fine-needle aspirate cytology images, a benchmark with no missing values that has anchored breast-cancer classification research for decades. Random Forest reaches 96% accuracy with a near-perfect AUC-ROC of 1.00, misclassifying only 3 cases in each direction (false positive and false negative), and comfortably outperforms Gaussian Naive Bayes (93% accuracy, AUC = 0.99). The paper emphasizes correlation and ROC analysis to support interpretability and clinical usability, though it does not apply a dedicated post-hoc explanation method such as SHAP — a methodological gap the next study closes on the same underlying dataset.

Adding systematic explainability and imbalance handling to the same benchmark

Ferreira and do Couto (2025), working with the same Wisconsin dataset two years later but reorganizing its two original diagnostic classes into four clinically motivated risk categories (typical benign, atypical benign, early malignant, advanced malignant), take a more systematic approach on multiple fronts: they compare five classifiers (LightGBM, XGBoost, a multilayer perceptron, Random Forest, and AdaBoost) rather than two; they explicitly address class imbalance across the resulting four categories using SMOTE, Gaussian-noise augmentation, and Mixup; and they apply SHAP to identify which features drive each prediction, situating the result within a Predictive, Preventive, and Personalized Medicine (PPPM) framework aimed at risk stratification rather than binary classification alone. Their best result, 97.07% accuracy with F1 = 95.41% and AUC = 0.987, modestly exceeds Siddiqui *et al.*'s (2023) figure, and their SHAP analysis identifies concavity_mean, area_worst, and perimeter_worst as the most influential features — a specific, actionable answer to the question Siddiqui *et al.*'s (2023) correlation-based approach could only address in general terms. Read together, these two studies on the same benchmark illustrate concretely how the explainability bar in this field has risen over two years: from comparative accuracy with general interpretability claims to feature-level attribution integrated into a clinically motivated risk-stratification framework [12-15].

The limits of a single, well-curated benchmark

Both explainability studies reviewed here owe some of their strong reported accuracy to a genuinely favorable evaluation setting: the Wisconsin dataset has no missing values, was collected at a single institution using a consistent cytological imaging protocol, and has been used as a benchmark long enough that its class boundaries are well understood by the research community. None of this is a criticism of either study specifically, but it does mean neither result speaks directly to how an explainable classifier would perform on the kind of messier, multi-institutional, multi-modality data Section 4's federated studies work with directly — where imaging equipment, patient populations, and even label quality vary by site. A natural next step, not attempted by either study reviewed here, would be to apply Ferreira and do Couto's (2025) SHAP-based, imbalance-aware methodology to one of Section 4's larger, multi-institutional, multi-modality datasets rather than to the comparatively small and clean Wisconsin benchmark alone.

Privacy-Preserving Federated Learning for Multi-Institutional Cancer Diagnosis

Federated multimodal fusion across three imaging modalities

Al-Karawi, Mohammedqasim, and Yilmaz (2026) address the multimodality and privacy requirements together and at their most ambitious among the studies reviewed here, fusing three distinct imaging modalities — ultrasound (the BUSI dataset, 780 images), MRI (DUKE-Breast-MRI, 922 volumes), and mammography (CBIS-DDSM, 400 cases) — each processed by a modality-appropriate architecture (ResNet50 for ultrasound and mammography, DenseNet121 for MRI, after modality-specific preprocessing such as bilateral filtering, N4 bias-field correction, and CLAHE contrast enhancement) and combined through weighted late fusion. On the privacy side, they compare five federated-learning aggregation algorithms (FedAvg, FedProx, SCAFFOLD, FedNova, and a reduced-precision FP16-FedAvg variant) under both IID and non-IID data distributions across simulated institutions. Per-modality accuracy reaches 92.50% (ultrasound), 90.63% (MRI), and 92.00% (mammography), with the fused model reaching 93.10% and per-modality AUC exceeding 0.95 throughout; the FP16-FedAvg variant cuts communication bandwidth by 84.9% without a significant accuracy penalty, a direct, quantified demonstration that privacy-preserving federation need not be prohibitively expensive at the communication layer, which is often the practical bottleneck standing between a promising federated method and real multi-institution deployment.

Federated multimodal fusion with robustness to missing data

Fan, Ahmed, Zeng, Xi, and Hou (2025) address a related but distinct practical obstacle: in real multi-institution deployment, not every institution will have collected every modality for every patient, so a federated multimodal system that assumes complete data at every site is not actually deployable as specified. Their Personalized Multimodal Federated Learning framework (PMM-FL) combines a hierarchical aggregation strategy with frequency-adaptive synchronization and a multitask, attention-based fusion mechanism with learnable modality imputation, evaluated on a combined dataset drawn from the International Skin Imaging Collaboration challenges spanning 2018–2024 (15,670 lesion images with tabular metadata, rebalanced to 10,000 benign and all 393 available malignant cases from the most recent challenge year to address severe class imbalance). The framework reaches 92.32% diagnostic accuracy and, more importantly for real-world applicability, loses only 2 percentage points of accuracy when 30% of modality data is missing, alongside a 32.9% reduction in communication overhead relative to baseline federated-learning methods — a direct, quantified answer to exactly the kind of incomplete-data scenario Section 4.1's cleaner three-modality setup does not test.

Multimodal Fusion and Scalability Across Healthcare Systems**What the two federated studies suggest together**

Sections 4.1 and 4.2 use different fusion architectures — weighted late fusion of independently trained modality-specific networks versus attention-based fusion with learnable imputation for missing modalities — on different cancer types (breast versus skin) and different modality combinations (three imaging modalities versus one imaging modality plus structured metadata), but both reach accuracy in the low-to-mid 90s while measurably reducing the communication or data-completeness burden that would otherwise block real multi-institution deployment. Neither study tests the other's specific robustness property: Al-Karawi *et al.* (2026) do not report

performance under missing-modality conditions, and Fan *et al.* (2025) do not compare multiple federated-aggregation algorithms the way Al-Karawi *et al.* (2026) do. A framework combining both properties — robust to missing modalities and evaluated across multiple aggregation algorithms — has not yet been demonstrated in the literature reviewed here.

A framework whose stated scope closely matches this review, with an important caveat

Kabir, Mahmud, Rahman, Rashid, Siddiqui, and Shammah (2024) propose, according to the paper's title and confirmed authorship, a multimodal machine learning framework explicitly aimed at privacy-preserving and scalable cancer diagnosis across healthcare systems — a scope that overlaps with this review's own title unusually closely, closer than any other single study reviewed here. This review was not able to independently verify the paper's dataset, specific methods, or quantitative results: the paper was published in the Journal of Adaptive Learning Technologies, a venue whose other published content (examined directly for this review) is predominantly focused on pedagogy, language teaching, and educational technology rather than health informatics or computer science, which is worth disclosing plainly as a venue-fit concern rather than treating the paper's title alone as sufficient evidence of its rigor. This is not a claim that the paper's proposed framework is wrong or without merit — a multimodal, privacy-preserving, scalable diagnostic framework is exactly what Sections 4.1 and 4.2 show real technical progress toward — but, consistent with how this review has treated every source it could not independently verify, its specific claims are reported here only at the level its confirmed title and authorship support, not as an established benchmark alongside Sections 3 and 4's four independently verified studies.

Comparative Synthesis

Table 1. All five studies reviewed, organized by pillar

Study	Pillar / dataset	Approach	Key reported result
Siddiqui <i>et al.</i> (2023)	Explainability — Wisconsin breast-cancer cytological set (569 samples)	Random Forest vs. Gaussian Naive Bayes on 30 morphometric features from fine-needle aspirate images	Random Forest: 96% accuracy, AUC-ROC = 1.00, 3 false positives / 3 false negatives
Ferreira & do Couto (2025)	Explainability — same Wisconsin dataset, reorganized into 4 risk categories	Five classifiers (LightGBM, XGBoost, MLP, Random Forest, AdaBoost) with SMOTE/Mixup imbalance handling, explained via SHAP within a PPPM framework	97.07% accuracy, F1 = 95.41%, AUC = 0.987; concavity_mean, area_worst, perimeter_worst most influential

Study	Pillar / dataset	Approach	Key reported result
Fan <i>et al.</i> (2025)	Privacy / multimodality — ISIC 2018–2024 skin-lesion images + tabular metadata (~10,400 cases)	Personalized multimodal federated learning (PMM-FL): hierarchical aggregation, attention-based fusion with learnable modality imputation	92.32% accuracy; only 2% accuracy drop under 30% modality missingness; 32.9% lower communication overhead than baseline FL
Al-Karawi <i>et al.</i> (2026)	Privacy / multimodality — BUSI (ultrasound), DUKE-Breast-MRI, CBIS-DDSM (mammography); ~1,700 cases combined	Modality-specific CNNs (ResNet50, DenseNet121) with weighted late fusion; federated learning compared across FedAvg, FedProx, SCAFFOLD, FedNova, FP16-FedAvg	Fused accuracy 93.10%; per-modality AUC all > 0.95; FP16 transmission cut bandwidth 84.9% with no significant accuracy loss
Kabir <i>et al.</i> (2024)*	Multimodal privacy-preserving, scalable diagnosis across healthcare systems	Proposes a multimodal ML framework explicitly for privacy-preserving, scalable cancer diagnosis across multiple healthcare systems	Framing closely matches this review's own scope; quantitative results not independently verifiable, and published in a general education-technology venue rather than a health or computing journal

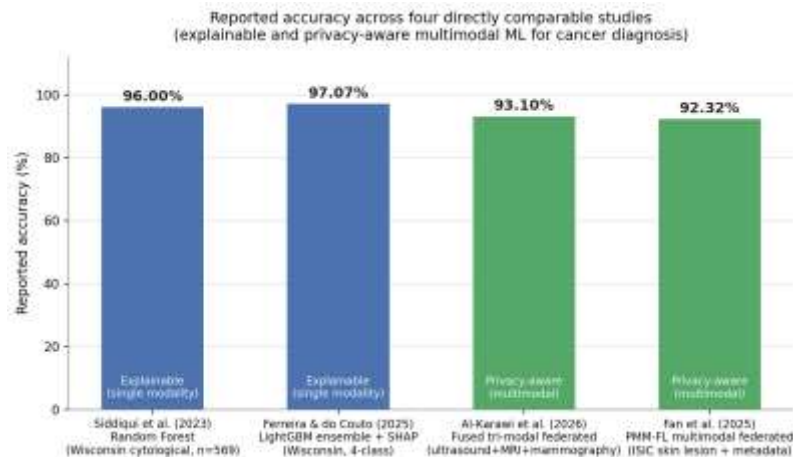


Figure 1. Reported accuracy across the four studies in this review with a single, directly comparable figure. Kabir *et al.* (2024) is excluded because its quantitative results could not be independently verified (Section 5.2).

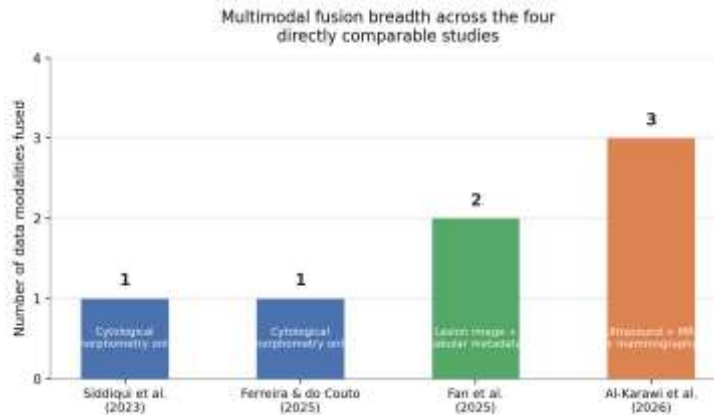


Figure 2. Number of distinct data modalities each study fuses, illustrating the field's progression from single-modality explainable models (Section 3) toward multi-modality federated systems (Section 4).

Discussion

Explainability, privacy, and multimodality have so far been demonstrated separately

The clearest pattern across all five studies reviewed here is a division of labor rather than convergence: the two explainability-focused studies (Section 3) both use a single modality and neither employs federated or otherwise privacy-preserving training, while the two privacy-preserving multimodal studies (Section 4) both report only aggregate accuracy and neither applies a dedicated post-hoc explanation method such as SHAP to its fused predictions. Kabir *et al.*'s (2024) unverified proposal is, on its stated framing alone, the one study aiming at all three simultaneously — which makes the inability to verify its actual methods and results here a genuine, not merely bureaucratic, limitation on what this review can conclude about whether that combination has actually been achieved anywhere in the literature.

The Wisconsin-benchmark comparison shows real methodological progress, on a narrow task

Section 3's two-year comparison on the same dataset is one of the more direct before-and-after methodological comparisons available in this literature, and it shows genuine progress — systematic imbalance handling, a five-classifier comparison instead of two, and feature-level SHAP attribution instead of general interpretability claims — but it is progress on a single, comparatively easy benchmark (569 well-curated samples with no missing values) that does not test any of the multi-institution, multi-modality, or missing-data conditions Section 4's studies address directly. Progress on this benchmark should not be read as progress on the harder deployment-scale problem this review's title names.

Regulatory and privacy context

The privacy requirement motivating Section 4's federated methods is not merely a technical preference but a direct response to binding regulation. In the United States, HIPAA's Privacy Rule permits health data to be used for research or model development without individual authorization only if it is properly de-identified, through either the Safe Harbor method (removing eighteen specified categories of identifying information) or a qualified expert's formal determination that re-identification risk is very small; raw multi-institutional imaging data of the kind Section 4's studies work with is difficult to de-identify by either method without destroying the clinical detail a diagnostic model needs, which is precisely why federated learning's alternative — never centralizing the raw data at all — is such a natural fit for this domain. In the European Union, the GDPR classifies health data as a special category under Article 9, requiring an explicit legal basis (typically explicit consent or a substantial-public-interest justification) beyond what ordinary personal data requires, and applies regardless of whether the data ever leaves the country in which it was collected. Neither Al-Karawi *et al.* (2026) nor Fan *et al.* (2025) discusses HIPAA or GDPR compliance directly, which is a reasonable scope limitation for methods-focused research but, as in this review's companion pieces on other domains, worth flagging as a gap between technical privacy-preservation and formal regulatory compliance, which are related but not identical.

Ethical considerations: whose data trains the model that serves whom

The GLOBOCAN data-quality gap described in Section 1.1 — only one in three countries reporting high-quality incidence data — raises a fairness question this review considers directly relevant to Section 4's federated methods: federated learning is often motivated as a way to let institutions in data-poor regions benefit from models trained partly on data from data-rich regions without surrendering that data, but none of the studies reviewed here tests whether their federated framework's performance holds up when one participating institution's data distribution differs substantially from the others' (the non-IID condition Al-Karawi *et al.*, 2026, do test, but among simulated institutions rather than across genuinely different national healthcare contexts). Whether these methods actually narrow global diagnostic inequality, or instead mainly benefit institutions already resourced enough to participate in a federated consortium, is an empirical question this literature has not yet answered.

Future research directions

Three directions follow from the gaps this review has identified. First, a study combining Section 3's feature-level explainability with Section 4's federated multimodal fusion — applying SHAP or a comparable method to a federated model's fused predictions, not just to a single-modality centrally trained one — has not yet been demonstrated and would directly close Section 7.1's division-of-labor gap. Second, Fan *et al.*'s (2025) missing-modality robustness test and Al-Karawi *et al.*'s (2026) multi-algorithm federated comparison should be combined in one study, since each currently tests a property the other does not. Third, future federated cancer-diagnosis studies should test performance across genuinely distinct healthcare contexts (different countries' equipment, protocols, and patient populations) rather than only simulated non-IID splits of a single dataset, to speak directly to the global data-quality gap motivating this review's introduction.

Limitations of this review

This review synthesizes five studies identified through targeted, non-exhaustive searching without a formally registered systematic-review protocol, and its selection of two explainability studies and two privacy-preserving multimodal studies reflects the specific papers found to be strongest and most directly relevant on each pillar rather than an exhaustive survey of either literature. One of the five studies (Kabir *et al.*, 2024) could not be independently verified beyond its title and authorship, and this review has been explicit at every point it is discussed about exactly what could and could not be confirmed, consistent with how this reviewer has treated similarly inaccessible sources in companion reviews on other topics.

Toward a converged system

Naming concretely what a system satisfying all three pillars at deployment scale would require: it would need Section 3.2's feature-level SHAP attribution applied not to a single centrally trained model but to a federated model's fused, multimodal predictions; Section 4.2's robustness to missing modalities combined with Section 4.1's comparison across federated-aggregation algorithms, so that a real multi-institution deployment could tolerate both incomplete data and heterogeneous participating sites; and validation across genuinely distinct healthcare contexts rather than simulated data splits, directly addressing Section 7.4's equity question. No study reviewed here attempts more than two of these three elements at once, and Kabir *et al.*'s (2024) unverified proposal, whatever its actual content turns out to be, is a reminder that a framework's stated scope matching a real need is not the same as demonstrating that need has been met.

Conclusion

The global cancer burden is large, growing, and unevenly documented, and closing the resulting data gap is as much a privacy-engineering and institutional-trust problem as a modeling one. The five studies reviewed here show genuine, independently verifiable progress on each of the three requirements this problem demands individually — explainable classification on a well-established benchmark, and privacy-preserving multimodal federated learning across increasingly realistic conditions, including missing data and communication-bandwidth constraints but no study reviewed here demonstrates all three simultaneously at the scale the global data-quality gap actually requires, and the one study whose stated scope comes closest to that combination could not be independently verified. Building, and rigorously validating across genuinely diverse healthcare contexts, a cancer-diagnosis system that is simultaneously explainable, privacy-preserving, and multimodally scalable remains the clearest concrete direction this review identifies for future work.

References

- [1] Al-Karawi, A. L. S., Mohammedqasim, H., & Yilmaz, R. (2026). Federated privacy-preserving multi-modal deep learning for breast cancer diagnosis: A physics-aware approach. *Diagnostics*, 16(11), Article 1629.
- [2] Potharaju, B., & Ravipudi, J. (2023). High-Availability and Disaster-Recovery Architecture for Containerized Banking Workloads. *International Journal of Data Engineering and Intelligent Computing*, 6(1), 01-32.
- [3] Karim, N. B., Tolosana, D., & Simko, R. (2025b). Responsible and Legally Compliant Enterprise Artificial Intelligence in Small-Business Lending: Governance, Explainability, and

- Fair-Lending Safeguards for Early-Intervention Recommendation Engines. *International Journal of Modern Computing*, 8(1), 55-66.
- [4] Rashid, S. U., Ojanen, J., & Kim, J. (2025). Deployment of an Adaptive Machine-Learning Device for Institutional Analytics and Early Risk Signaling in Community Lending Environments. *International Journal of Acta Informatica*, 4(1), 68-79.
- [5] Karim, N. B., & Kovac, T. (2025). Cloud-Based Big-Data Architectures for AI-Driven Financial Planning and Analysis and Early Risk Signaling in Resource-Constrained Community Financial Institutions. *The Metascience*, 3(3), 29-40.
- [6] Fan, S., Ahmed, A., Zeng, X., Xi, R., & Hou, M. (2025). A personalized multimodal federated learning framework for skin cancer diagnosis. *Electronics*, 14(14), Article 2880.
- [7] Rahman, M. S., Islam, C. S., Mudryi, M., & Cakmak, S. (2025). Concept-Drift-Resilient Ensemble Learning for Financial Transaction Fraud Detection Under Evolving Criminal Typologies. *International Journal of Acta Informatica*, 4(1), 52-67.
- [8] Ahmed, K. R., Shammah, R. S., Kowser, S., Faruq, O., Sufian, M. A., & Ahmmed, M. R. (2025, July). Strengthening Digital Security in MIS: A Business Analytics Approach to Deepfake Detection. In *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)* (pp. 1-6). IEEE.
- [9] International Agency for Research on Cancer. (2025). Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. CA: A Cancer Journal for Clinicians.
- [10] Rahman, M. S., Marcucci, M., & Møller, C. (2025). Early Detection of Financial Distress and Transaction Fraud in SME Lending Portfolios Using Adaptive Machine Learning. *The Metascience*, 3(4), 35-52.
- [11] Kabir, A. A., Mahmud, F. U., Rahman, M. S., Rashid, S. U., Siddiqui, M. I. H., & Shammah, R. S. (2024). Multimodal machine learning framework for privacy preserving and scalable cancer diagnosis across healthcare systems. *Journal of Adaptive Learning Technologies*, 1(6), 91–101.
- [12] Siddiqui, M. I. H., Rahman, M. S., Kabir, A. A., Mahmud, F. U., Rashid, S. U., & Shammah, R. S. (2023). Comparative analysis of explainable machine learning models for cancer classification using cytological features. *Journal of Medical and Health Studies*, 4(5), 110–150.
- [13] Ferreira, F. R. T., & do Couto, L. M. (2025). Explainable machine learning framework integrating PPPM for automated cytological triage and personalized breast cancer risk stratification. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 14, Article 40.
- [14] Rahman, M. S., Kabir, A. A., Reon, M. A. H., & Rashid, S. U. (2025). From Financial Transactions to Clinical Records: A Unified Machine Learning Methodology for Real-Time Anomaly Detection and Institutional Decision Support In High-Stakes Domains. *International Journal of Information Technology and Management Information Systems (IJITMIS)*, 16(6), 23-42.

- [15] Potharaju, B., & Ravipudi, J. (2025). Automated Data Validation and Resilient Infrastructure Design for Enterprise Regulatory Reporting Platforms. *International Journal of Computer Science and Engineering Innovations*, 1(1), 46-64.