June 2024

# The value of Google Cloud Medical Imaging Suite (MIS) in radiology diagnostics

**Nupur Gautam**

# Abstract

**Main body:** All medical specialties have an inherent, intrinsic risk of missed diagnosis that has an impact on health care, and radio diagnosis is not immune to this risk. This paper discusses the challenges faced in the field of radiology, a global radiologist shortage, and an ever-aging human population that requires more health care. All the factors combined put more pressure on the already overworked medical professionals, including radiologists. This paper also captures the steady losses incurred by health care. Google Medical Imaging Suite provides solutions that can mitigate these challenges to a great extent and provide overall efficiency to the health care systems.

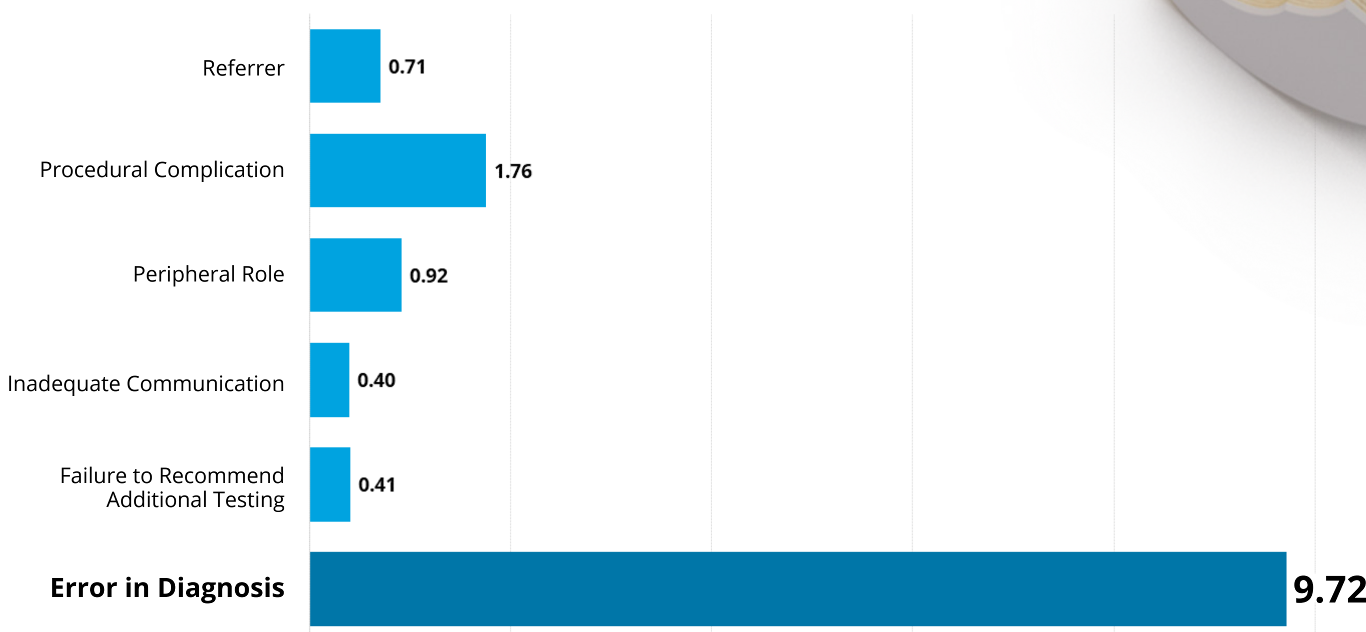
**Conclusion:** Google MIS can accelerate imaging diagnostics and reduce burden on radiologists by imaging analysis, thereby improving access to better patient care and outcomes.

# Challenges in the field of radiology

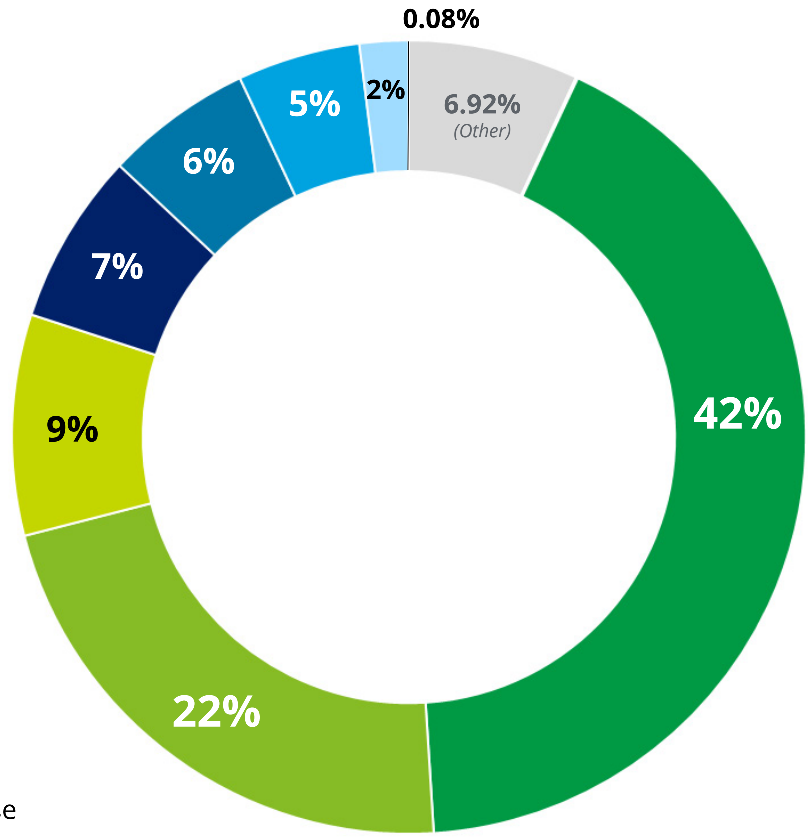
## Diagnostic errors

Per the Radiology Society of North America, the leading cause of errors by radiologists arise from errors in diagnosis, not procedural complications or failures to communicate or recommend additional studies.[1]

**SUM OF CLAIMS PER 1,000 PERSONS-YEAR**



As per Kim & Mansfield below is the radiologic error categorization, 2014.[2]




**42%**

**Under-reading**

Abnormality visible but not reported

**22%**

**Satisfaction of search**

After having identified a first abnormality, radiologist fails to continue to look for additional abnormalities

**9%**

**Faulty reasoning**

Abnormality identified but attributed to wrong cause

**7%**

**Abnormalities outside area of interest (but visible)**

Many on first or last image of CT or MR series, suggesting radiologist's attention not fully engaged at beginning or end of reviewing series

**6%**

**Satisfaction of report (alternative reasoning)**

Uncritical reliance on previous report in reaching diagnosis, leading to perpetuation of error through consecutive studies

**5%**

**Failure to consult prior imaging studies**

**2%**

**Inaccurate or incomplete clinical history**

**0.08%**

**Correct report failing to reach referring clinician**

**6.92%**

**Others**

## Factors contributing to radiology error

### Global radiologist shortage:

The annual growth rate of radiologists in the United States between 2010 and 2020 entering the workforce is only 2.5%[3] with 29,530 radiologists as of May 2021. This equates to about one radiologist serving 11,200 people in the United States. The US House of Representatives introduced the Resident Physician Shortage Act of 2021.[4] Europe has 13 radiologists per 100,000 population, but in the United Kingdom, the rate is only 8.5 per 100,000. In some other countries like Malaysia this number can be around 30 radiologists per million.[5]

### Mental fatigue:

Excessive continuous-duty shifts and work hours for many health care professionals combined with sleep deprivation has been shown experimentally to produce effects on certain mental tasks equivalent to alcohol intoxication.[7] Continuous prolonged decision-making results in decision fatigue and leads to unconscious taking of shortcuts in cognitive processes, resulting in poor judgment and diagnostic errors. Radiology trainees providing preliminary interpretations during off-hours are especially prone to this effect.[8]

### Visual fatigue:

Krupinski and the co-authors established that long radiology workdays reduce detection and accommodation accuracy.[6]

### Inattentional blindness:

Research at Harvard University's Visual Attention Lab established that inattentional blindness describes the phenomenon wherein observers miss an unexpected but salient event when engaged in a different task.[9]

# Other technical challenges

### Distribution drift:

Medical imaging lacks standardized acquisition protocols, and there is a large variation in terms of equipment and scanning settings. This leads to the "distribution drift" phenomenon.[10]

### Isolated data:

Due to requirements around patient privacy, imaging data is distributed across different hospitals and imaging centers.[11]

### Noisy labels:

Labeling or annotation of medical images can be nonstandard and time consuming. Because of variable experience and different conditions, both inter-user and intra-user labeling inconsistency is high, resulting in noisy labels.[12]

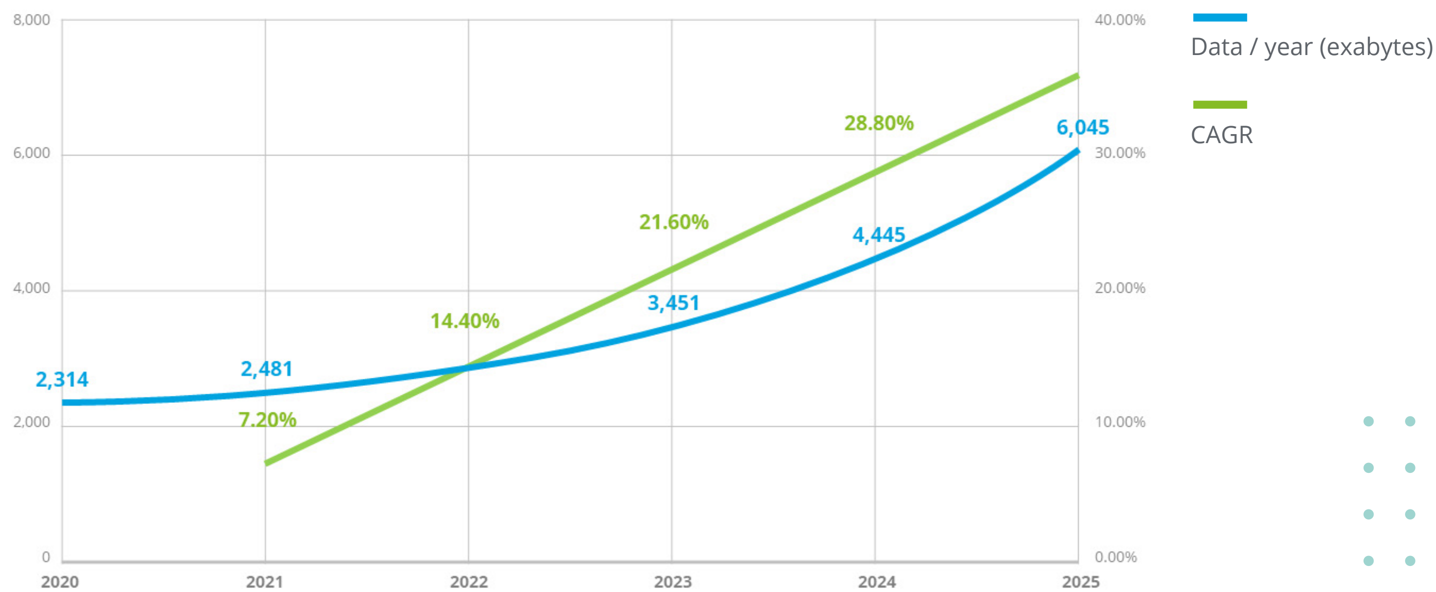
### Heterogeneous and imbalanced samples:

The ratio between positive and negative samples is extremely uneven. For example, the number of pixels belonging to a tumor is usually one to many orders of magnitude less than that of normal tissue.[13]

# Future of radiology without AI/MIS

**Ever growing imaging data:** The amount of medical data produced every year is huge, and the rate of growth is exponential. The 2020 global annual pre-COVID estimate for health care data was around 2,314 exabytes and an expected compound annual growth rate (CAGR) by 2025 is 36%.[14] A linear extrapolation of growth shows that there will be 6,035 exabytes of health care data produced in 2025, with 90% of that consisting of medical imaging.[15]

**FORECASTED HEALTHCARE DATA PER YEAR**




Each medical image frame typically ranges from 8 to 50 megabytes[16] and taking the high end of that would mean 119.6 trillion images.

**Aging population increases imaging needs:** According to the World Health Organization, the proportion of the world's population over 60 years of age will be 22% by 2050, nearly double that of 2015.[17]

**Ever growing workload on medical staff:** To put things in perspective, at the low end it takes about 15 minutes[18] for a radiologist to process simpler images (such as for pneumonia) and longer for more complicated cases. **It would take 3.4 billion radiologists working 24 hours a day, 7 days a week, 365 days a year, to process all the data in 2025 manually.**

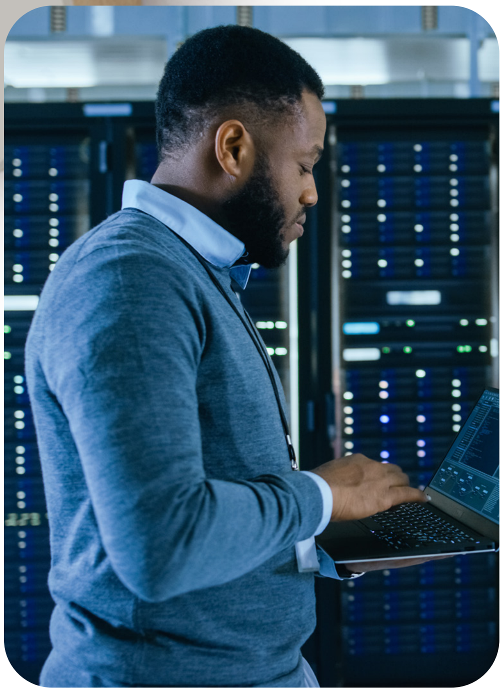
**119.6 Trillion** Images X **15 Minutes** per image

**29.9 Trillion Hours**

↓

**3.4 Billion Radiologists**

Working 24 / 7 / 365



# Google Cloud

**Google Cloud:** Google Cloud Platform is a suite of cloud-computing services that runs on the same infrastructure that Google uses internally for its end-user products.

# Cloud Healthcare API

**Cloud Healthcare API** allows easy and standardized data exchange between health care applications and solutions built on Google Cloud. With support for popular health care data standards, such as HL7® FHIR®, HL7® v2, and DICOM®, Cloud Healthcare API provides a fully managed, highly scalable, enterprise-grade development environment for building clinical and analytics solutions securely on Google Cloud. The Cloud Healthcare API also includes additional value-added capabilities, such as automated DICOM and FHIR de-identification (de-ID) to better prepare data for these solutions.

Cloud healthcare API provides a pathway to intelligent analytics and machine learning capabilities in Google Cloud with prebuilt connectors for streaming data processing in Dataflow, scalable analytics with BigQuery, and machine learning with Vertex AI.

The Cloud healthcare API is backed by Google Cloud's privacy and security features, supports HIPAA compliance, and is in scope for Google Cloud's ISO/IEC 27001, ISO/IEC 27017, and ISO/IEC 27018 certifications. In addition, Google Cloud is HITRUST CSF certified.[19]

# Medical Imaging Suite

**Medical imaging suite(MIS):** MIS Platform, offered by Google, is a suite of medical imaging components that leverages Google Cloud's world-class infrastructure. MIS offers the following components:

**Imaging storage:** Secure, scalable, standardized and managed cloud storage environment with integrated de-ID

**Imaging lab:** AI-assisted labeling and annotation tools to automate highly repetitive tasks

**Imaging datasets and dashboards:** Easily view and search petabytes of data for advanced analytics and cohort building

**Imaging AI pipelines:** Easily transform images and annotations into Vertex AI datasets for a faster model training process

**Imaging deployment:** Flexible options for cloud, on-prem or edge deployment, and real-time insights

MIS accelerates development of AI for medical imaging by making imaging data accessible, interoperable, and useful.

- Support convergence of diverse storage formats to the DICOM standard
- Seamless integration with on-premises storage via NetApp or Change Healthcare cloud-native PACS
- AI-assisted annotation environment powered by NVIDIA and the MONAI open-source framework
- Support healthcare-specific security and compliance[20]

# Benefits of Medical Imaging Suite

## Accelerating imaging diagnostics with interoperability:

Sharing imaging data using Google's scalable ingestion frameworks reduces time and resources to deliver scalable AI/ML.

## Reducing the burden on radiologists:

AI-assisted annotation and labeling automate highly repetitive tasks, streamlining radiologist workflow.

## Imaging analysis and datasets:

Imaging dashboards and datasets allow users to easily search petabytes of data for advanced analytics and cohort building.

## Privacy and security:

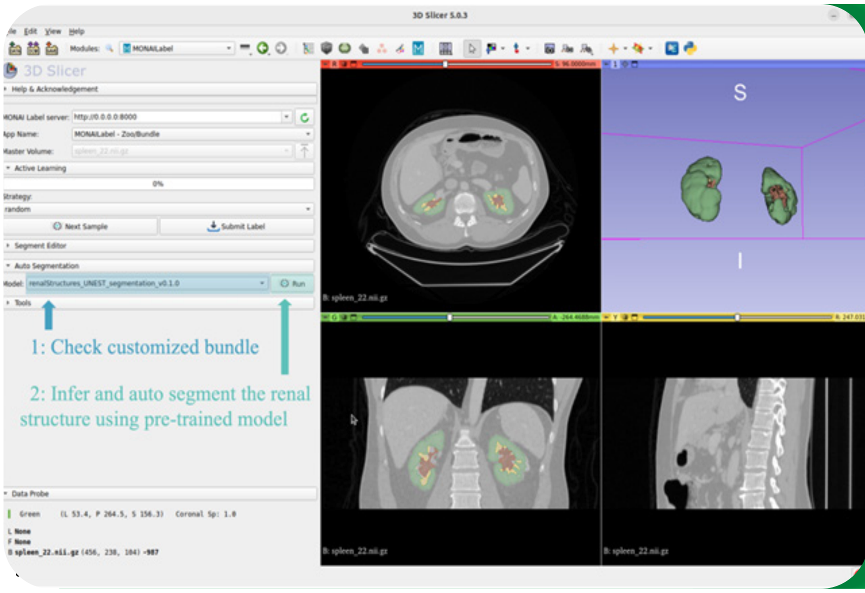
Ensure protected health information (PHI) and personally identifiable information (PII) is protected using Google's best-in-class security frameworks, identity and access management, and VPC Service controls. MIS identifies and remediates PII and PHI in images and metadata, creating datasets where researchers can analyze data and train models without needing internal review board approvals.

## Imaging AI pipelines and deployment:

MIS easily transforms images and annotations into Vertex AI datasets for a faster model training process. Flexible options are available for cloud, on-prem, or edge deployment. Google Distributed Cloud, enabled by Anthos, extends Google Cloud's infrastructure and services to the edge.

## Improve access to better patient care and outcomes:

Transform disease detection and diagnosis by prioritizing critical cases, augmenting treatment decisions, or expanding screenings in areas where there are shortages of doctors.




Deloitte makes imaging data **interoperable and useful,** driving accessibility of medical images and enabling labeling and detection and diagnosis of disease

**MONAI 3D SLICER**[21]

# How MIS meets these objectives

**Ingest data:** MIS is seamlessly integrated with Google Cloud and can ingest DICOM images into Cloud Healthcare DICOM Stores through Cloud Storage Import or through DICOMWeb GKE Adapter.

**Annotate image data:** Allow professionals to annotate image data within Vertex AI by drawing segments around areas of interest (e.g., cancers) using open-source MONAI label on Google Compute Engine (see Figure 5, MONAI 3D Slicer).[21]

**Enrich image data:** Drawing the segment around an area of interest is challenging. Various ML models adjust the outline to fit perfectly.

**Train model:** Create an ML model to find areas of interest (e.g., cancer) by training it on the annotated images in Vertex AI.

**Evaluate model:** Apply the model to a sample of annotated data it has never seen before and see how it performs compared to the annotator.

**Validate model:** Determine which confidence threshold the model needs in order to make various business decisions (e.g., notify a radiologist of possible cancer).

**Deploy model:** Model is deployed to Vertex AI API endpoint to respond to applications with predictions. Enable MLOps to measure model drift and auto retrain if needed.

**Deidentify data:** Medical Imaging Suite de-identification pipelines use configurable algorithms to identify and redact PHI and PII in image data, as well as in extracted DICOM metadata tags.

# Case studies by Deloitte

**Study 1:**

# Medical imaging ingestion and de-identification

## Engagement Summary

- Medical Imaging Suite was deployed at a major US health care provider.
- The solution ingested nearly 1 billion image files, more than 500 terabytes, and subsequently de-identified images and metadata.
- DICOM tags were exported to BigQuery for both original and de-identified datasets.
- A segmentation model trained on CT scans was built to detect physical anomalies.
- The model was applied to the provider's data on Google Cloud production environment.

## Our Approach

- About 1 petabyte of medical images were stored in Cloud Storage bucket.
- Deloitte worked with client cloud engineers to create secure environment for ingestion of medical images using Google Cloud Platform.
- Deloitte was able to ingest and subsequently de-identify almost 1 billion image files.
- Deloitte built a pipeline for ingestion and de-identification of new images placed in Cloud Storage bucket using Cloud Functions and Python.
- For easier user access and analytics enablement, the imaging metadata was streamed into a BigQuery data warehouse, allowing users to build dashboards on top of the BigQuery tables.
- Analytics dashboards were built using Looker; visualizations were reviewed and customized per the client team inputs.
- Deloitte trained a machine learning model on a Google-owned public lung images dataset that can segment areas within DICOM instances (i.e., lung nodules to detect cancer) and saved the model. PyTorch framework and GCP Jupyter Notebooks with 8vCPUs, 30 GB RAM, and NVIDIA Tesla V100x1 were used for this.

## Key Outcomes

- De-identification algorithms that met client requirements for data privacy (excluded images with barcodes and maintained key data elements [e.g., accession number]) were successfully implemented.
- The pipeline for ingestion and de-identification was moved to production, processing about 100K images/hour.
- Dashboard visualizations provided insights into image cohorts.
- Required DICOM attributes were successfully analyzed and collected.
- The image instance segmentation model to detect lung nodules was built.
- The trained model was implemented on the provider's images to plot results.
- The trained model was successfully saved and stored; Vertex AI Jupyter notebooks with results were implemented to the client-owned repository for future use and enhancements.

**Study 2:**

# Brain tumor identification and location

## Engagement Summary

- A federal medical research agency wanted an automated method to identify brain tumors in a variety of MRI volume views (i.e., T1, T2, etc.). They also wanted to be able to scan their historical images in bulk to determine whether any hard-to-detect tumors were present and went undetected during their initial review by radiologists.
- Many of the tumors had indistinct boundaries between malignant tissue and healthy tissue and were present in both white matter and grey matter, making them difficult to identify.

## Our Approach

- An ML model was developed using Google Vertex AI to review all MRI images and determine the presence and location of brain tumors of various sizes, tissue types, locations, and shapes.
- 3D reconstruction of the 2D DICOM images was used to assist the provider in surgery planning.

## Key Outcomes

- The model was able to identify tumors with high performance for a sample set of tumors with different sizes, shapes, tissue types, and locations in the brain.
- The model was highly scalable. While not deployed into production, a model built on Google Vertex AI allows it to be deployed to multiple machines to provide inferences at scale and meet high demand. It also allows the model to be deployed to an API endpoint with ease, making MLOps more manageable.
- The model has not yet been deployed for production use, but it would be considered for use as the first step of tumor identification, with all true positives and false positives needing to be reviewed by medical personnel.

# Endnotes


1. T. Jeremy S. Whang et al., "The causes of medical malpractice suits against radiologists in the United States," Radiology 266, no. 2 (2013).

2. Young W. Kim and Liem T. Mansfield, "Fool me twice: Delayed diagnoses in radiology with emphasis on perpetuated errors," American Journal of Roentgenology 202, no. 3 (2014).

3. Mary Henderson, "Radiology facing a global shortage," Radiological Society of North America (RSNA), May 10, 2022.

4. Senate Bill 834 – Resident Physician Shortage Reduction Act of 2021, 117th Cong. (2021–2022).

5. Henderson, "Radiology facing a global shortage."

6. Adrian P. Brady, "Error and discrepancy in radiology: Inevitable or avoidable?," Insights into Imaging 8, no.1 (2017): pp. 171–82.

7. Elizabeth A. Krupinski et al., "Long radiology workdays reduce detection and accommodation accuracy," Journal of the American College of Radiology 7, no. 9 (2010): pp. 698–704.

8. David M. Gaba and Steven K. Howard, "Fatigue among clinicians and the safety of patients," New England Journal of Medicine 347 (2002): pp. 1249–55.

9. Trafton Drew, Melissa L.-H. Võ, and Jeremy M. Wolfe, "The invisible gorilla strikes again: Sustained inattentional blindness in expert observers," Psychological Science 24, no. 9 (2013): pp. 1848–53.

10. S. Kevin Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," Proceedings of the IEEE 109, no. 5 (2021): pp. 820–38.

11. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises."

12. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises."

13. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises."

14. Nick Culbertson, "The skyrocketing volume of health care data makes privacy imperative," Forbes, August 6, 2021.

15. Zhou et al., "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises."

16. Dandu Ravi Varma, "Managing DICOM images: Tips and tricks for the radiologist," Indian Journal of Radiology and Imaging 22, no. 1 (2012): pp. 4–13.

17. Henderson, "Radiology facing a global shortage."

18. Whitney L. Jackson, "In radiology, turnaround time is king," Diagnostic Imaging, November 15, 2015.

19. Google Cloud, Cloud Health care API: Unlock your healthcare data to power insights and AI, accessed February 21, 2024.

20. Henderson, "Radiology facing a global shortage."

21. Google Cloud, Medical Imaging Suite, accessed February 21, 2024.

## Contacts

**Nupur Gautam**
Manager
Deloitte Consulting LLP
nugautam@deloitte.com