



DICOM Image ANalysis and Archive (DIANA): an Open-Source System for Clinical AI Applications

Thomas Yi^{1,2} · Ian Pan^{1,2} · Scott Collins¹ · Fiona Chen^{1,2} · Robert Cueto⁶ · Ben Hsieh¹ · Celina Hsieh^{1,2} · Jessica L. Smith³ · Li Yang⁴ · Wei-hua Liao⁵ · Lisa H. Merck^{2,3,6} · Harrison Bai^{1,2} · Derek Merck^{1,2,6}

Received: 23 June 2020 / Revised: 22 February 2021 / Accepted: 19 April 2021 / Published online: 2 November 2021
© Society for Imaging Informatics in Medicine 2021

Abstract

In the era of data-driven medicine, rapid access and accurate interpretation of medical images are becoming increasingly important. The DICOM Image ANalysis and Archive (DIANA) system is an open-source, lightweight, and scalable Python interface that enables users to interact with hospital Picture Archiving and Communications Systems (PACS) to access such data. In this work, DIANA functionality was detailed and evaluated in the context of retrospective PACS data retrieval and two prospective clinical artificial intelligence (AI) pipelines: bone age (BA) estimation and intra-cranial hemorrhage (ICH) detection. DIANA orchestrates activity beginning with post-acquisition study discovery and ending with online notifications of findings. For AI applications, system latency (exam completion to system report time) was quantified and compared to that of clinicians (exam completion to initial report creation time). Mean DIANA latency was 9.04 ± 3.83 and 20.17 ± 10.16 min compared to clinician latency of 51.52 ± 58.9 and 65.62 ± 110.39 min for BA and ICH, respectively, with DIANA latencies being significantly lower ($p < 0.001$). DIANA's capabilities were also explored and found effective in retrieving and anonymizing protected health information for “big-data” medical imaging research and analysis. Mean per-image retrieval times were 1.12 ± 0.50 and 0.08 ± 0.01 s across x-ray and computed tomography studies, respectively. The data herein demonstrate that DIANA can flexibly integrate into existing hospital infrastructure and improve the process by which researchers/clinicians access imaging repository data. This results in a simplified workflow for large data retrieval and clinical integration of AI models.

Keywords DICOM · Machine learning · Artificial intelligence · Intracranial hemorrhage · Bone age · PACS

Background

Hospital medical imaging services rely on Picture Archiving and Communications Systems (PACS) as central medical imaging repositories. However, most PACS setups are

optimized to support a single departmental workflow and do not readily accommodate new retrospective or prospective image analysis workflows for research or clinical purposes. Accessing a large collection of images from a PACS, as required for artificial intelligence (AI) modeling, often requires repetitive, manual labor to retrieve and post-process each of thousands of studies [1]. This challenge hinders multiple independent tasks. For example, the review of large datasets is contingent on the ability to assemble said large amount of data in the first place—it follows that manual assembly could serve as the rate-limiting factor. Other hindrance presents itself in the real-time application of clinical analytics as data can be difficult to obtain with manual methods in both a prospective and time-conscious manner; at best, imaging devices can be reconfigured to directly send desired data to specific hosts, but this requires coordination from different parties. Additionally, the administrative and operational complexity of managing the embedded protected health information (PHI) in medical imaging is

✉ Harrison Bai
Harrison_bai@brown.edu

¹ Department of Diagnostic Imaging, Rhode Island Hospital, Providence, RI, USA

² Warren Alpert Medical School of Brown University, Providence, RI, USA

³ Department of Emergency Medicine, Rhode Island Hospital, Providence, RI, USA

⁴ Department of Neurology, Second Xiangya Hospital, Changsha, China

⁵ Department of Radiology, Xiangya Hospital, Changsha, China

⁶ University of Florida, Gainesville, FL, USA

another commonly cited difficulty in the use of electronic health records and medical images in research studies often before manual retrieval tasks can even begin [2].

In an effort to alleviate some of the aforementioned issues, academic institutions have moved towards alternative solutions including vendor neutral archives (VNAs) with the hope of providing an improved, universal image viewing experience by decoupling the archival part of PACS from proprietary dependencies [3]. However, the switch from a pure PACS to a VNA system is associated with significant data migration challenges [4]. For research purposes specifically, some institutions have attempted to leverage open-source public medical image informatics platforms such as XNAT, but such systems have challenges in setup and implementation, which may limit scalability [5, 6]. Integrating image data into clinical data warehouses has also been limited by the same issue of scalability, and also by the massive investment of resources required to establish a clinical data warehouse [7–9]. Due to the obstacles in handling medical data, researchers have begun to implement bespoke methods for image data handling focused on training and delivery of AI models. Other research has been exploring methods to optimize delivery of AI models and their results to clinical settings [10].

The DICOM Image Analysis and Archive (DIANA) system was designed as a Python command line interface that interfaces with hospital PACS to automate common medical image data handling processes.¹ The development of modern computing frameworks such as container systems (e.g., Docker) and text databases (e.g., Splunk) all motivated the possibility of DIANA's existence, and refinement of AI containers (e.g., Docker with Keras/Tensorflow) remains a constantly improving field of research [11, 12]. This work aims to describe and evaluate the core functionality of DIANA and its potential in improving both research and clinical practice in the context of big-data retrieval, analysis, and AI delivery for medical imaging research.

Methods

The DIANA system leverages several open-source tools including Docker, a container system that facilitates replicability, and Orthanc, an open-source DICOM service that serves as a mini-PACS capable of communicating with other PACS systems [13, 14].

At a high-level, DIANA was evaluated in the context of two core tasks:

- i. Automated, retrospective batch retrieval of specific studies from PACS with optional DICOM metadata PHI anonymization
- ii. Automated, prospective monitoring of PACS for incoming studies of interest for extension to pre-trained, live AI pipelines and notification system

The bulk of the system's functionality will be reviewed in the context of these tasks. Miscellaneous functionality, such as radiation dose monitoring, is presented in the context of the overall system. DIANA was run on a headless server with an Intel® Xeon® CPU E5-2637 v3 @ 3.50 GHz. All AI pipelines were run without graphics acceleration.

Retrospective Batch Retrieval

The process of retrospective batch retrieval offers an introductory overview of the DIANA system as outlined in Fig. 1. DIANA is setup up as a container housing all the necessary analysis and control scripts on an institutional machine on the same network as the PACS. As part of data retrieval, an Orthanc service is also created in the form of a container. The Orthanc node is added as an allowed DICOM node on the network, with privilege to query and retrieve from the PACS.

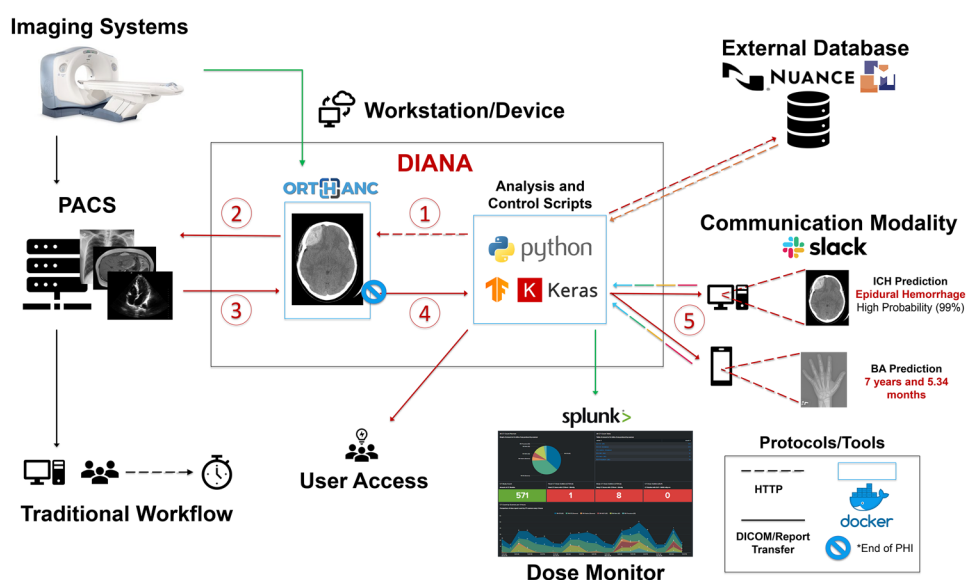
From there, the steps in image data retrieval are enumerated below:

1. Image identifiers (accession numbers) are input into DIANA, and HTTP requests are initiated to Orthanc to identify and retrieve each imaging study from the PACS. Arbitrarily large collections of image identifiers can be input in the form of a text file pre-populated with the desired institutional accession numbers.
2. Orthanc converts each HTTP request into DICOM C-Find, and C-Move requests and submits them to the hospital PACS.
3. Imaging data and structured reports (radiation dose reports) are retrieved from the PACS into Orthanc.

At this point following the retrieval, the data may be automatically stripped of all of PHI DICOM metadata tags as a follow-up to the Python script that initiated the request (step 1 in Fig. 1), using a built-in Orthanc anonymization protocol. This anonymization consists of erasing all the tags that are specified in Table E.1–1 from PS 3.15 of the DICOM standard 2008 or 2017c (default), but may be customized [15]. DIANA will automatically generate sham values for certain fields (e.g., randomized fake names or DICOM Unique Identifiers with embedded validators) if desired. Of note, DIANA does not strip raw images of any PHI (e.g., a coincidental dog tag on chest x-ray), requiring users to be cognizant of such markers.

¹ <https://github.com/derekmerck/diana2>

Fig. 1 Overview of DIANA system. For lines, solid indicates DICOM command or dose report transfer and dashed indicates HTTP request. As in the legend, red reflects DIANA-specific communications, and green is transfer of dose reports for radiation monitoring. Multi-colored and orange lines reflect Slack® and Nuance mPower communications, respectively. Blue outlines represent containers. The traditional workflow is shown on the left



- PHI-free data are either saved locally to the machine where DIANA runs for post-processing or may be immediately forwarded along to other services such as the AI pipeline (step 5 in Fig. 1), as detailed in the next section.

Data that are stored locally are hereafter available for users to access from the local disk as flat files. As part of the anonymization process, a comma-separated values key that links original PHI to the anonymized studies may be generated. In case of network interruption during the data collection process, DIANA pre-screens for accession numbers that have already been downloaded upon restarting a job, and only proceeds with pulling missing data. Additional clinical data (e.g., radiologist reports, ordered labs) related to an accession number may be queried against mPower (Nuance, Burlington, MA, USA—formerly Montage) via DIANA and contextualized with the metadata.

The DIANA workflow described above represents a more flexible, ad hoc approach to institutional PACS integration and use. Long-term archival of data pulled from the PACS can be maintained by taking advantage of Orthanc’s scalability when using a Postgres backend database [13]. The Orthanc helper system can then act as a local cache for previously accessed image data. This reduces the query burden on the PACS when image collections are rebuilt or updated.

The automated DICOM data collection pipeline is most commonly used for AI research, but can also be useful in other contexts. An additional workflow is shown in green in Fig. 1 that highlights a separate DIANA tool—the dose monitor. The dose monitor allows 24/7 monitoring of radiation-dependent imaging (e.g., XR, CT) across several hospitals and was developed with the Splunk text database as the backend and dashboarding system [16]. Radiation-

dependent imaging devices of interest are configured to submit a duplicate dose report to a dedicated DIANA Orthanc container. DIANA monitors this repository and automatically forwards JSON representations of the DICOM structured dose reports to a Splunk instance. Splunk dashboards and reports then provide customizable, detailed safety, and operational summaries to the imaging technologists and service managers. For example, data may be filtered down to all the imaging studies performed in adults within the past 48 h at a single hospital and reviewed for unexpected radiation exposure levels.

To evaluate DIANA retrieval performance, both total and per-image retrieval times (including anonymization time) for 5, 10, 20, 40, and 80 imaging studies were calculated where the images were either computed tomography (CT) scans or x-rays (XRs). Per-image retrieval time was calculated as the total time required to download and anonymize a cohort of images divided by the total number of images in that cohort. In the context of CT scans, per-image is defined as per-slice.

Prospective AI Pipelines

DIANA was configured to monitor the institutional PACS for incoming pediatric bone age (BA) x-rays and emergency department (ED) head trauma CT scans over the course of 2 months. BA XRs were processed with an AI classifier that received the second-place award in the 2017 Radiological Society of North America (RSNA) machine learning challenge and is now being applied clinically [17, 18]. Head trauma CTs were processed with an internally developed classifier derived from the 2019 RSNA machine learning challenge [19]. The output for the BA-XR AI pipeline was predicted bone age from radiographs, and ground-truth values were clinically determined BA using the standard

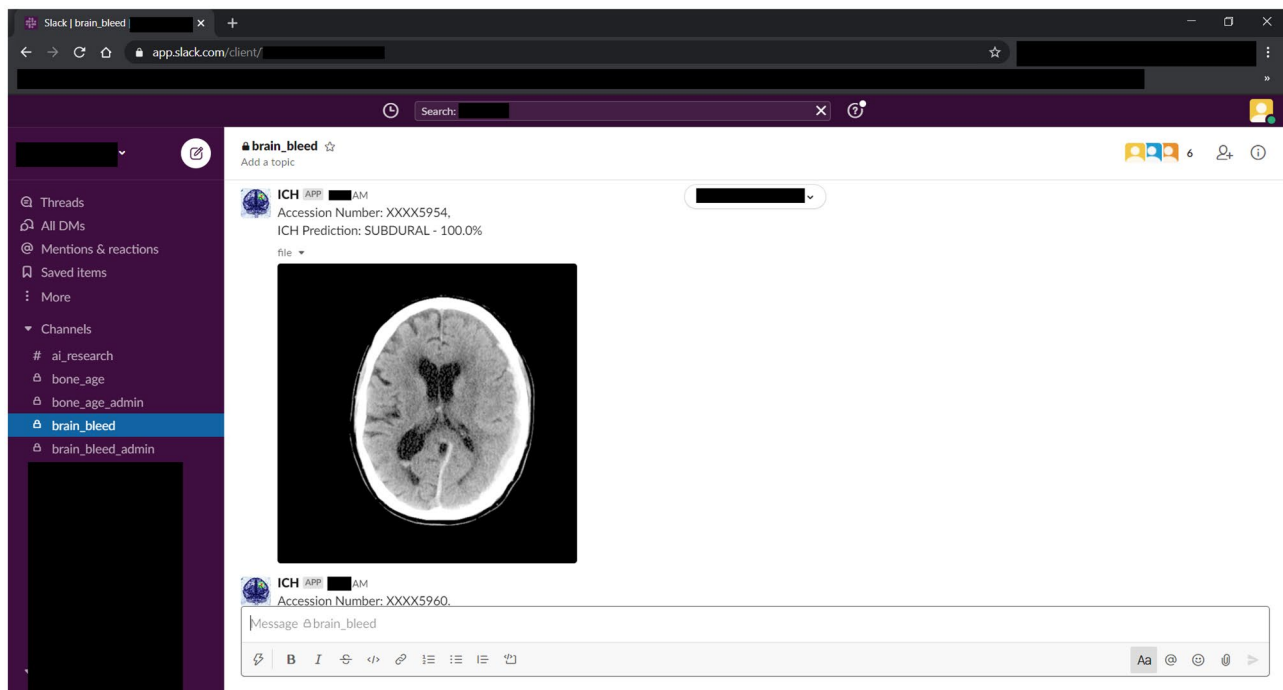


Fig. 2 Sample view of a Slack® channel demonstrating the end user view for the ICH clinical AI pipeline. Users may message study information here for the DIANA backend to selectively process studies.

Greulich-Pyle method [20]. The outputs for the intracranial hemorrhage (ICH) AI pipeline were probability of and type of bleed, including intraparenchymal, intraventricular, and subarachnoid hemorrhage, as well as subdural and epidural hematomas. Ground-truth values were compiled by semi-automated review of relevant clinical reports.

The results of the prospectively conducted AI assays were forwarded to an internally monitored Slack® (cloud-based messaging platform) channel for display, as shown in Fig. 2. The system allows selective processing by user-specified information via direct Slack® messages in the respective channel (e.g., requesting a retrospective review for a study via accession number) and may be applied to any classifier managed by DIANA.

Statistical Analysis

The Python seaborn package was used to fit regression curves to the batch retrieval latency analysis. Notable computations were otherwise performed using Python's SciPy Statistics module. The institutional mPower database was queried for ground truth, defined as the findings of radiologist report readings, and latency for reading imaging studies [21]. A comprehensive search through mPower was performed to count all completed studies and their timeframes of interest to evaluate the number of cases DIANA caught

and processed. Primary system latency was calculated for radiologists and DIANA as time between exam completion until report creation or Slack® posting, respectively. In the acute ICH setting, secondary latency was calculated for radiologists as time from exam completion until report finalization; that is, after a preliminary report has been reviewed and accepted by an attending physician. Paired t-tests were performed for case-by-case latency to assess differences between DIANA and radiologist in both pipelines.

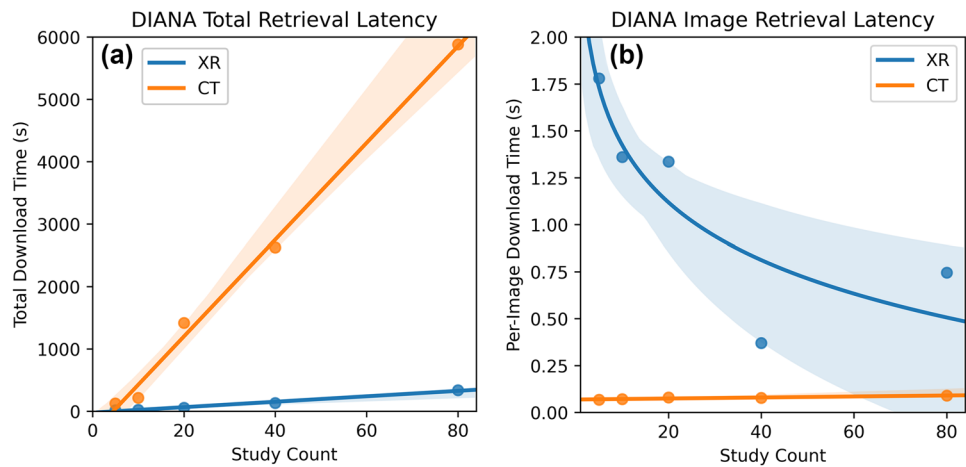
AI model-specific performance metrics were assessed for each pipeline. For BA, the discrepancies between AI and radiologist readings were pruned from radiologist reports. For ICH, AI model performance was quantified as an Area Under Curve (AUC) in the binary classification problem of ICH presence on head trauma CT scan. Detection threshold for this cohort was set to 57.6% based on analysis of a prior cohort of DIANA ICH AI data [22].

AI model-specific performance metrics were assessed for each pipeline. For BA, the discrepancies between AI and radiologist readings were pruned from radiologist reports. For ICH, AI model performance was quantified as an Area Under Curve (AUC) in the binary classification problem of ICH presence on head trauma CT scan. Detection threshold for this cohort was set to 57.6% based on analysis of a prior cohort of DIANA ICH AI data [22].

Results

Evaluation of time required to retrieve images through the batch retrieval pipeline yielded the results shown in Fig. 3. Total retrieval times ranged from 17.80 to 339.71 s and 128.18 to 5882.26 s for XR and CT studies, respectively. Mean per-image retrieval times were 1.12 ± 0.50 and 0.08 ± 0.01 s across XR and CT studies, respectively.

Fig. 3 Latency analysis for retrospective batch retrieval where predominant patient study was either XRs or CTs demonstrating **a** total download times as a function of number of studies retrieved and **b** per-image download times across studies retrieved. Shaded regions represent the 95% confidence interval about the regression curves



Preliminary assessment of BA measurement across 44 cases revealed a mean discrepancy of 9.25 ± 10.86 months for AI prediction compared to ground truth (radiologist reading). Mean DIANA latency of 9.04 ± 3.83 min was significantly lower than radiologist latency of 51.52 ± 58.9 min ($p = 1.84 \times 10^{-5}$). BA results are summarized in Fig. 4.

For ICH cases, the DIANA AI captured 100% of ED head CTs completed during the study period. In the 626 studies included in analyses, DIANA demonstrated 92% sensitiv-

ity and 94% specificity in detecting presence or absence of ICH. The AUC for DIANA was 97% with optimal detection threshold at 50.7%. Mean latency for DIANA was 20.17 (median 17.6) ± 10.16 min. Primary and secondary latencies for radiologists were 65.62 (median 26.0) ± 110.39 and 91.67 (median 36.0) ± 157.82 min, respectively. Speed of interpretation for DIANA was significantly greater than radiologists' ($p = 1.28 \times 10^{-22}$, 2.70×10^{-27}). ICH results are summarized in Table 1 and Fig. 5.

Fig. 4 DIANA BA AI results: **a** distribution of differences between radiologist and AI readings per case; **b** case-by-case assessment of DIANA and radiologist reading times (note the semi-log axes). Size and color correspond to the difference in BA reading between AI and radiologist

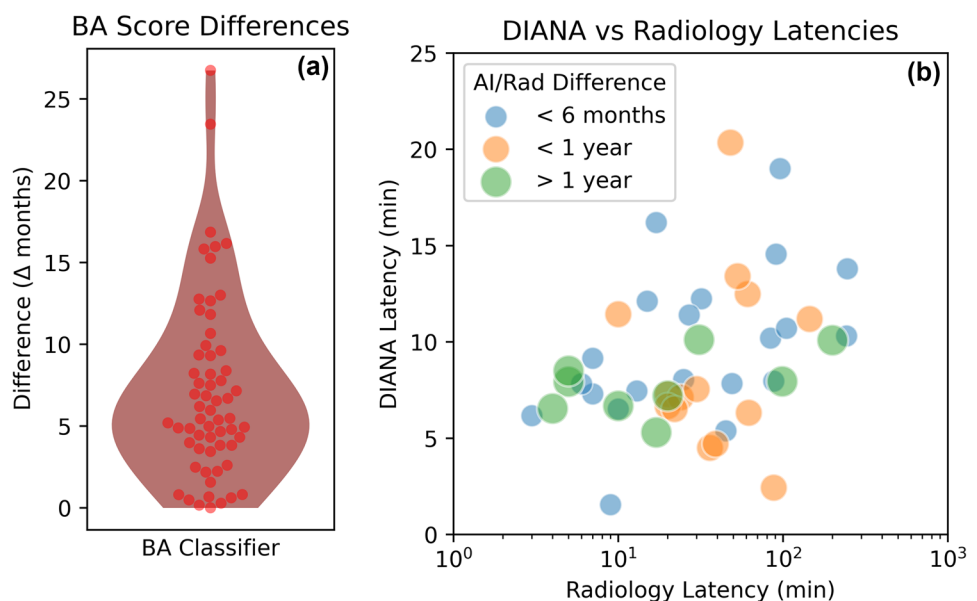


Table 1 Summary of DIANA ICH statistics

DIANA ICH overview			
True positives	108	False positives	28
False negatives	10	True negatives	480
Sensitivity	0.92	Specificity	0.94
Subtyping performance			
Bleed type	Correctly identified	Total cases	Proportion correct
Intraventricular	7	7	1.00
Intraparenchymal	34	39	0.87
Subdural	41	45	0.91
Subarachnoid	22	24	0.92
Epidural	3	3	1.00

Discussion

In this work, DIANA demonstrated retrospective batch retrieval latencies that suggest efficiency and consistency in handling multi-image jobs. In general, the results demonstrate that pull times scale linearly with query burden.

The true value of such speeds is best appreciated in context. As an example of automation in action, a 90 x-ray cohort was pulled and anonymized via DIANA then transferred to and displayed from a research PACS in under 2 h. By comparison, manual labor is intensive, and from experience, it may take days to weeks to pull even a mere 100 studies.

The AI pipeline results demonstrate that DIANA can potentially integrate into existing workflow as an adjunct tool without hindering clinical work times. The AI pipelines ultimately aim to increase clinician confidence for the benefit of patients similar to other automated engineering

solutions for medical image analysis problems [23–25]. In the context of clinical workflow, DIANA hopes to avoid obstacles in configuring modalities to send to specific hosts for post-processing through its comprehensive, institutional network presence. Likewise, the system aims to improve the accessibility of AI model predictions on live data through integration with ubiquitously accessible platforms.

Each pipeline's results have their own nuances worth examining. A potential improvement for the BA estimation pipeline is to display the BA standard deviations for a patient on the Slack® channel based on gold standard literature tables. However, these standard deviations are read from the patient's age by birthdate, not bone age. Since the institution limits DIANA's access to patient chart data other than imaging, these metrics are not currently being automatically reported by the system described. Still, many of the patients assessed in Fig. 4 were teenagers, which put their bone age standard deviation at 10 months [20]. It is encouraging to see that even when the DIANA AI model and radiologist readings differed, the difference was almost always within two standard deviations of expectation.

For ICH, it is interesting to note that the average ICH read time by radiologists is about 1 h—much more than one might expect in an emergent setting. One potential explanation lies in the lumping of follow-up ED cases with acute patient scans. Acute scans may be more readily identifiable and would be expected to be read much more quickly, likely within minutes. Beyond this, the AI model used in this work demonstrated remarkable performance with respect to radiologist ground truth classification of bleed types. The AUC of 0.97 suggests that models such as this one have strong, timely potential in increasing clinician confidence in reads. This is further celebrated by the notion that there are always increasingly excellent models on the horizon.

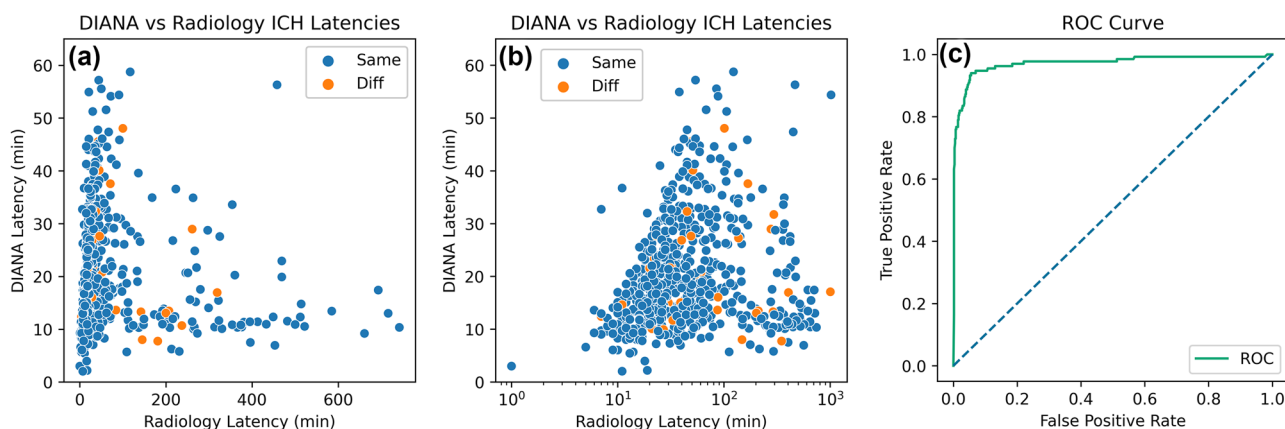


Fig. 5 DIANA ICH AI results: case-by-case visualization of latencies for DIANA and radiologist defined as **a** time until initial report creation and **b** time until report finalization. Orange represents cases

where binary detection of bleed presence differed between DIANA and radiologists. **c** Receiver operating characteristic (ROC) curve for ICH AI model

It is important to note that the accuracy of latency calculations for radiologists in this work is limited by workflow realities. In the context of emergent cases such as ICH, it is entirely likely that the reading radiologist would have called the attending ED physician to report any acute findings. The practical latency in a case like this would be the time from imaging completion to the time of such phone call. However, given the challenge of tracking such reports, the closest institutional surrogate available was the time of initial radiologist report creation. When looking at latency with the endpoint defined as time of report completion, the longer times suggest another key component of workflow: radiologists are responsible for assessing innumerable other features in imaging studies beyond singular findings.

In this work, freely available models were integrated with DIANA and evaluated. However, there is no restriction on model integration beyond the abstraction of models to accept an arbitrary image-related input and provide an accessible output result. This opens the door for DIANA to be used in conjunction with commercial models, if desired, with subsequent ability to refine those models as prospective results are obtained and analyzed. With the flexibility of model choice, DIANA could improve delivery of AI models from diverse trades to clinical settings.

Creating PHI-sensitive interinstitutional datasets is a high priority in prospective clinical translational imaging research [26]. Overall, many healthcare environments could benefit from AI augmentation including automated screening for stroke on CT or magnetic resonance imaging, chest XR evaluation, and detection/grading of neoplastic pathology [27, 28]. Although AI training and deployment are obvious targets, there are also many other research and quality assurance applications for readily available imaging data and metadata.

For example, the radiation dose monitor has offered a comprehensive quality-assurance system by which annual national dose guidelines may be readily maintained. Prospective considerations could include meticulous monitoring of pediatric imaging radiation dose on account of the increased radiation morbidity and mortality associated with pediatric patients [29, 30].

DIANA strives to achieve maximum usability for clinicians and researchers while minimizing workflow disruptions that could influence acceptance. From the perspective of a healthcare system, there is room for continuous improvement in DIANA's components. As an example, the present AI pipelines leverage Slack® as a prototype endpoint, substitutable for another endpoint. That said, other clinical environments may benefit more by leveraging an existing institutional, real-time notification system. Ideally, a system such as DIANA would integrate with electronic medical record (EMR) services as well—mPower serves as a partial EMR surrogate at the present institution. How-

ever, tighter integration with established clinical systems also shifts DIANA away from its strengths in rapid ad hoc use and into the domain of established commercial medical informatics products.

Anecdotally, once appropriate approvals were secured, a new DIANA system was integrated with a local clinic's radiology PACS and downloading image data in a matter of hours. However, if anything, this demonstrated how disproportionately challenging it is to navigate bureaucratic approval to interface directly with PACS due to requirements for maintaining PHI confidentiality and minimizing potential impact on clinical IT workloads. Medical imaging is notoriously difficult to completely anonymize given the vast variability in DICOM header field utilization. From an operations standpoint, the risks associated with batch access to the hospital PACS are non-trivial. Extremely aggressive batch queries during high clinical volume times can result in significant system latency. And large batch queries can result in explosion of local or network file storage. Even dedicated resources can be overwhelmed by a single DIANA request that results in thousands of CT scans and hundreds of gigabytes of fulminant storage [31].

The final discussion this work will touch upon is in navigating some of the real-world challenges of deploying a system such as DIANA into enterprises that are understandably resistant to such tools. In making imaging data more accessible via DIANA, it was important to address the degree to which image PHI could be accessed by various users in the data retrieval process from PACS as well as the network location of the data at all points. In addressing these concerns, the DIANA research team worked closely with the institutional information security team to thoroughly explore these points. This collaborative discussion was streamlined through several means:

- i. Outlining system functionality and associated image PHI flow through system diagrams such as Fig. 1. Such depictions provide additional assurance to all parties that there are no invasive modifications required within PACS itself.
- ii. Review of storage locations of images (e.g., behind institutional firewall on the network, external machine) and possible degrees of PHI associated with images at all access points.
- iii. Manual comparison of random imaging studies and modalities post-anonymization via DIANA and prior institutional gold standard (e.g., manually “hand scrubbing” away PHI by technical expert) to, at minimum, establish anonymization parity between DIANA and original anonymization means. At the local institution, DIANA anonymization was even shown to be slightly more meticulous compared to traditional anonymization, with kudos to Orthanc's capabilities.

By working through these valid concerns, the research team was able to bring a live DIANA setup to fruition. Establishment of this work aims to promote more wide-spread acceptance of such systems in the future.

Conclusion

DIANA empowers big data projects with image access and new routes of hospital communication that can extend AI applications within the academic medical center environment. Moreover, the facilitation of AI delivery to clinical settings could reduce workflow burdens on clinicians.

Acknowledgements The authors would like to acknowledge Drs. Jonathan Movson and Guarav Jindal for their input on the clinical utilities of DIANA. In addition, the authors would like to thank Wendy Smith, Alex Todorovich, Anthony DeLuca, Alison Chambers, and Lifespan Information Services (Frank Kucienski, John Haley, and others) for input on the PHI-security of the system.

Funding This research was supported by the department of diagnostic imaging at the local institution.

Availability of Data and Material Any experimental data mentioned above is available upon request to authors.

Code Availability The repository of interest is publicly available at <https://github.com/derekmerck/diana2>.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. M. Alhajeri and S. G. S. Shah, Limitations in and solutions for improving the functionality of picture archiving and communication system: an exploratory study of PACS professionals' perspectives, *Journal of Digital Imaging*, 2019, <https://doi.org/10.1007/s10278-018-0127-2>.
2. C. A. Kushida, D. A. Nichols, R. Jadrnicek, R. Miller, J. K. Walsh, and K. Griffin, Strategies for de-identification and anonymization of electronic health record data for use in multi-center research studies, *Medical Care*, 2012, <https://doi.org/10.1097/MLR.0b013e3182585355>.
3. C. Sirota-Cohen, B. Rosipko, D. Forsberg, and J. L. Sunshine, Implementation and benefits of a vendor-neutral archive and enterprise-imaging management system in an integrated delivery network, *Journal of Digital Imaging*, 2019, <https://doi.org/10.1007/s10278-018-0142-3>.
4. T. K. Agarwal and Sanjeev, Vendor neutral archive in PACS, *Indian Journal of Radiology and Imaging*, 2012, <https://doi.org/10.4103/0971-3026.111468>.
5. D. S. Marcus, T. R. Olsen, M. Ramaratnam, and R. L. Buckner, The extensible neuroimaging archive toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data, *Neuroinformatics*, 2007, <https://doi.org/10.1385/NI.5:1:11>.
6. R. Herrick, W. Horton, T. Olsen, M. McKay, K. A. Archie, and D. S. Marcus, XNAT Central: open sourcing imaging research data, *NeuroImage*, 2016, <https://doi.org/10.1016/j.neuroimage.2015.06.076>.
7. S. L. MacKenzie, M. C. Wyatt, R. Schuff, J. D. Tenenbaum, N. Anderson, Practices and perspectives on building integrated data repositories: results from a, 2010 CTSA survey, *Journal of the American Medical Informatics Association*, 2012 <https://doi.org/10.1136/amiajnl-2011-000508>.
8. D. J. Foran *et al.*, Roadmap to a comprehensive clinical data warehouse for precision medicine applications in oncology, *Cancer Informatics*, 2017, <https://doi.org/10.1177/1176935117694349>.
9. F. A. Meineke, S. Staübert, M. Löbe, and A. Winter, A comprehensive clinical research database based on CDISC ODM and i2b2, in *Studies in Health Technology and Informatics*, 2014, <https://doi.org/10.3233/978-1-61499-432-9-1115>.
10. A. Sedghi *et al.*, Tesseract-medical imaging: open-source browser-based platform for artificial intelligence deployment in medical imaging, 2019, <https://doi.org/10.1117/12.2513004>.
11. A. Grupp, V. Kozlov, I. Campos, M. David, J. Gomes, and Á. López García, Benchmarking deep learning infrastructures by means of TensorFlow and containers, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, https://doi.org/10.1007/978-3-030-34356-9_36.
12. P. Xu, S. Shi, and X. Chu, Performance evaluation of deep learning tools in Docker containers, in *Proceedings - 2017 3rd International Conference on Big Data Computing and Communications, BigCom 2017*, 2017, <https://doi.org/10.1109/BIGCOM.2017.32>.
13. S. Jodogne, The Orthanc ecosystem for medical imaging, *Journal of Digital Imaging*, 2018, <https://doi.org/10.1007/s10278-018-0082-y>.
14. S. Jodogne, C. Bernard, M. Devillers, E. Lenaerts, and P. Coucke, Orthanc—a lightweight, restful DICOM server for healthcare and medical research, in *Proceedings - International Symposium on Biomedical Imaging*, 2013, <https://doi.org/10.1109/ISBI.2013.6556444>.
15. “Anonymization and modification—Orthanc Book documentation.” [Online]. Available: <https://book.orthanc-server.com/users/anonymization.html>. [Accessed: 21-Apr-2020].
16. D. Merck, S. Collins, and K. Laurie, Monitoring radiation exposure With DICOM and Splunk, in *Splunk .conf*, 2017.
17. S. S. Halabi *et al.*, The rSNA pediatric bone age machine learning challenge, *Radiology*, 2019, <https://doi.org/10.1148/radiol.2018180736>.
18. I. Pan, H. H. Thodberg, S. S. Halabi, J. Kalpathy-Cramer, and D. B. Larson, Improving automated pediatric bone age estimation using ensembles of models from the 2017 RSNA machine learning challenge, *Radiology: Artificial Intelligence*, 2019, <https://doi.org/10.1148/ryai.2019190053>.
19. A. E. Flanders *et al.*, Construction of a machine learning dataset through collaboration: the RSNA 2019 brain CT hemorrhage challenge, *Radiology: Artificial Intelligence*, 2020, <https://doi.org/10.1148/ryai.2020190211>.
20. W. W. Greulich and S. I. Pyle, *Radiographic atlas of skeletal development of the hand and wrist*. Stanford University Press, 1959.
21. “mPower Clinical Analytics for medical imaging | Nuance. [Online]. Available: <https://www.nuance.com/healthcare/diagnostics-solutions/radiology-performance-analytics/mpower-clinical-analytics.html>. [Accessed: 28-Apr-2020].
22. T. Yi *et al.*, Identification of intracranial hemorrhage using an original artificial intelligence system, in *Society of Academic Emergency Medicine 2020*.
23. A. Manbachi *et al.*, Clinical translation of the LevelCheck algorithm for automatic localization of target vertebrae in spine sur-

- gery,” *The Spine Journal*, 2017, <https://doi.org/10.1016/j.spinee.2017.07.290>.
24. T. de Silva *et al.*, Utility of the level check algorithm for decision support in vertebral localization, *Spine*, 2016, <https://doi.org/10.1097/BRS.0000000000001589>.
 25. S. Oh, J. H. Kim, S. W. Choi, H. J. Lee, J. Hong, and S. H. Kwon, Physician confidence in artificial intelligence: an online mobile survey, *Journal of Medical Internet Research*, 2019, <https://doi.org/10.2196/12422>.
 26. H. A. Piwowar and W. W. Chapman, Public sharing of research datasets: a pilot study of associations, *Journal of Informetrics*, 2010, <https://doi.org/10.1016/j.joi.2009.11.010>.
 27. H. X. Bai *et al.*, AI augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other etiology on chest CT, *Radiology*, 2020, <https://doi.org/10.1148/radiol.2020201491>.
 28. K. Chang *et al.*, Automatic assessment of glioma burden: a deep learning algorithm for fully automated volumetric and bidimensional measurement, *Neuro-Oncology*, 2019, <https://doi.org/10.1093/neuonc/noz106>.
 29. P. L. Khong *et al.*, ICRP Publication 121: radiological protection in paediatric diagnostic and interventional radiology, *Annals of the ICRP*, 2013, <https://doi.org/10.1016/j.icrp.2012.10.001>.
 30. D. J. Brenner, C. D. Elliston, E. J. Hall, and W. E. Berdon, Estimated risks of radiation-induced fatal cancer from pediatric CT, *American Journal of Roentgenology*, 2001, <https://doi.org/10.2214/ajr.176.2.1760289>.
 31. DICOM Library—about DICOM most common features of study. [Online]. Available: <https://www.dicomlibrary.com/dicom/study-structure/>. [Accessed: 28-Apr-2020].

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