*Digital Health and Informatics Innovations for Sustainable Health Care Systems*
*J. Mantas et al. (Eds.)*
*© 2024 The Authors.*
*This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).*
*doi:10.3233/SHTI240517*

# How Data Infrastructure Deals with Bias Problems in Medical Imaging

Feifei LI[a,1], Ekaterina KUTAFINA [a], Mirjam SCHONECK [b], Liliana Lourenco CALDEIRA [b] and Oya BEYAN [a,c]
[a] *Institute for Biomedical Informatics, University of Cologne, Faculty of Medicine and University Hospital Cologne, Germany*
[b] *Institute for Diagnostic and Interventional Radiology, University of Cologne, Faculty of Medicine and University Hospital Cologne, Germany*
[c] *Fraunhofer Institute for Applied Information Technology, FIT, Germany*
ORCiD ID: Feifei Li https://orcid.org/0000-0003-4815-8547, Ekaterina Kutafina https://orcid.org/0000-0002-3430-5123, Mirjam Schoneck https://orcid.org/0000-0002-1630-4012, Liliana Lourenco Caldeira https://orcid.org/0000-0002-9530-5899, Oya Beyan https://orcid.org/0000-0001-7611-3501

**Abstract**. The paper discusses biases in medical imaging analysis, particularly focusing on the challenges posed by the development of machine learning algorithms and generative models. It introduces a taxonomy of bias problems and addresses them through a data infrastructure initiative: the PADME (Platform for Analytics and Distributed Machine-Learning for Enterprises), which is a part of the National Research Data Infrastructure for Personal Health Data (NFDI4Health) project. The PADME facilitates the structuring and sharing of health data while ensuring privacy and adherence to FAIR principles. The paper presents experimental results that show that generative methods can be effective in data augmentation. Complying with PADME infrastructure, this work proposes a solution framework to deal with bias in the different data stations and preserve privacy when transferring images. It highlights the importance of standardized data infrastructure in mitigating biases and promoting FAIR, reusable, and privacy-preserving research environments in healthcare.

**Keywords.** Medical Imaging, Machine Learning, Bias, Federated Learning, Differential Privacy, Data Infrastructure

## 1. Introduction

Bias in the medical imaging domain can manifest in various forms, influencing diagnostic accuracy and patient care. Before the development of machine learning applications in medical imaging analysis, the bias only came from data collection and image acquisition, and this bias, together with human oversights, could result in diagnosis errors. Afterthe rapid development of computational algorithms, the medical imaging

[1] Corresponding Author: Feifei Li; E-mail: feifei.li@uk-koeln.de.

field will face additional sources of bias. For instance, biased algorithms demonstrate inconsistent performance when evaluated among sub-groups classified by attributes like age, ethnicity, gender, socioeconomic status, and other pertinent factors [1].

In this paper, we propose an illustrative taxonomy for analyzing bias problems in medical imaging after the emergence of artificial intelligence methods. Without claiming completeness, we aim to illustrate the need for alertness and preventive actions, such as the employment of distributed infrastructure to enable the use of generative methods for data augmentation and a differential privacy-preserving strategy.

## 2. Illustrative taxonomy of biases in medical imaging

The order of the proposed taxonomy is based on the timeline of the medical image analysis process, as shown in Figure 1:

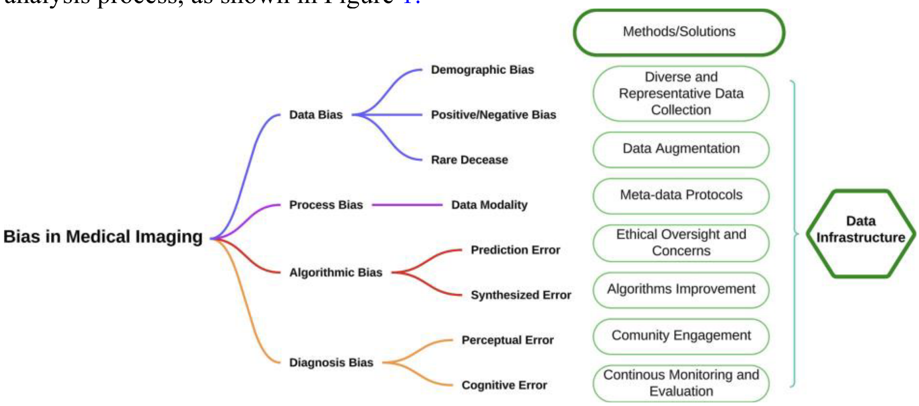



**Figure 1.** Illustrative taxonomy for bias types and available solutions in medical imaging

**Data bias:** This category describes biases from data source subgroup imbalances, including demographic diversities, positive/negative imbalances, and rare diseases. Deep learning methods heavily rely on training data, exacerbating data biases [2]. Demographic biases, such as age, race, and gender biases, contribute to disparities in data collection access. Positive/negative imbalances and rare diseases pose unique challenges in understanding and developing treatments due to limited data availability.

**Process Bias:** Medical imaging often suffers from differences in data modality and selection bias in data collection. For each imaging modality, the imaging process is based on the signal reconstruction process, such as signal misrepresentation and distortion, which can lead to some bias/error in the images [3].

**Algorithmic Bias:** Algorithmic bias can lead to performance disparities in machine learning models across demographic groups, often due to imbalanced training data or inherent biases in dataset development. It is crucial to consider both in-distribution and out-of-distribution errors when evaluating prediction results affected by algorithmic bias. Additionally, synthesized bias can occur when generative machine learning algorithms are used in medical image analysis, potentially augmenting datasets but also introducing data pollution if misused.

**Diagnosis Bias:** Since diagnosis is the last step in the whole process, its bias will be the accumulation of all the above sources of bias. Diagnosis bias is introduced by healthcare professionals during image interpretation and is influenced by factors such as prior experience, unconscious stereotypes, or institutionalized practices. Especially in the Radiology department, some works [4,5] analyse and summarize the causes of diagnostic imaging errors and biases based on practical cases and divide them into two groups: perceptual error and cognitive error.

Addressing bias in medical imaging requires a multifaceted approach, as illustrated in Figure 1, many methods can be used for each type of bias. All the approaches require a reliable data infrastructure to acquire, transfer, communicate and analyse the data.

## 3. Methods

PADME [6] is a Distributed Analytics (DA) infrastructure that supports i.a. the analysis of Computed Tomography (CT) scans. In this work, we propose a PADME infrastructure as a solution for cross-training of generative methods to deal with the data bias problem (see Figure 2). PADME is developed in line with the Personal Health Train (PHT) which complies with present privacy guidelines and considers the heterogeneous problems in each local data source. PHT was previously tested in multi-center research setup [7].

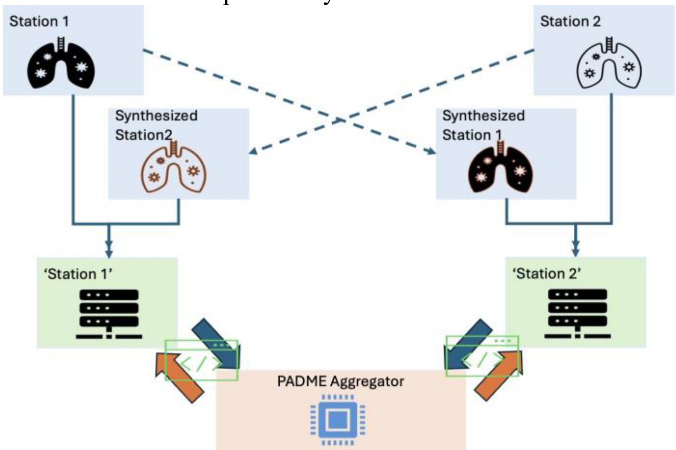



**Figure 2.** The framework of PADME. Generative models can be trained locally and then sent to other stationsfor data enhancement.

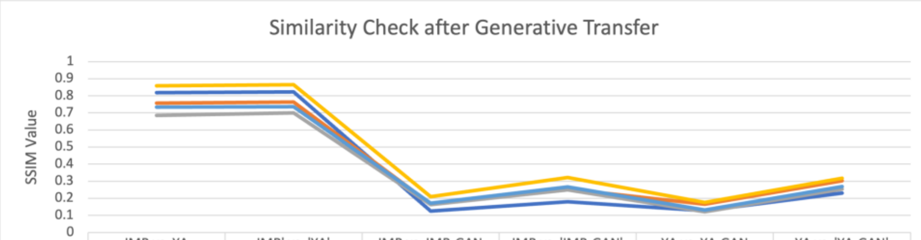
This framework shows the proof-of-concept using two "stations" or data hubs with potentially different data distributions. We apply generative methods to synthesize data from the complementary station and combine it with the original data to eliminate the bias between these two data distributions. This is made possible through distributed computing, preventing the leakage of sensitive patient data. Our infrastructure ensures that the algorithms are distributed to the local data station and sent back to the aggregator after the separate training. We take one example of process bias sourced from two different CT reconstruction methods, IMR (Iterative Model Reconstruction, Station 1) and YA iDose (Intelligent Dose reconstruction with Ya kernel, Station 2) [8]. Reconstruction is a processing step to convert raw data to an image, typically provided by the vendor.

In dealing with the modality transfer, we applied CycleGAN [9] and score-based diffusion model [10], which are known to perform well in mitigating both process and synthesis bias. For CycleGAN, we use 137 images from Station 1 and 86 images from Station 2 to compose the training dataset. For the diffusion model, we use the pretrained model from the MRI denoising work [10].

In the evaluation process, we use 5 paired patient data from the two stations. For the metric results, we use the structural similarity index measure (SSIM, 1=identical, 0=completely different) to evaluate the similar characteristics between medical images from different sources.

## 4. Results

In Figure 3, we can observe that a process bias existed in the first column since the two reconstruction methods resulted in differences in the visualization of the final images. When we process the data via diffusion models, the SSIM values are increased. From the [IMR vs. IMR-GAN] and [YA vs. YA-GAN] columns, large differences have been observed after the GANs-based synthesized data. In the [IMR vs. 'IMR-GAN'] and [YA vs. 'YA-GAN'] columns, the increase in the metrics illustrated the diffusion model's performance in reducing the algorithmic bias.




| | IMR vs. YA | IMR' vs. 'YA' | IMR vs. IMR-GAN | IMR vs. 'IMR-GAN' | YA vs. YA-GAN | YA vs. 'YA-GAN' |
|---|---|---|---|---|---|---|
| P1 | 0.8186 | 0.8226 | 0.1251 | 0.1788 | 0.1259 | 0.2308 |
| P2 | 0.7565 | 0.7638 | 0.1661 | 0.2535 | 0.1657 | 0.3019 |
| P3 | 0.6841 | 0.6996 | 0.1622 | 0.2501 | 0.1205 | 0.2545 |
| P4 | 0.8574 | 0.865 | 0.2088 | 0.3214 | 0.1766 | 0.3172 |
| P5 | 0.7338 | 0.7357 | 0.1709 | 0.2655 | 0.1311 | 0.2676 |

**Figure 3.** Similarity check between two different CT reconstruction methods IMR and YA. Items with a quote″ denote the image after the denoising diffusion transfer, and end with GAN denotes the image after the CycleGAN transfer, respectively, for both modules.

## 5. Discussion and Conclusions

In our work, we provided a taxonomy of bias problems in the medical imaging domain, introduced the platform for distributing computing, PADME, and experimented with results on generative methods performance via PADME to show that it can help build a FAIR, standardized, and privacy-preserving research environment for decentralized data analysis necessary for decreasing the bias problem in medical image analysis when dealing with different data sources. We were able to successfully test PADME to train generative algorithms on two data stations which may represent data bias linked to the differences in reconstruction methods. The infrastructure supports the enrichment of the local data modalities by generating bias-decreasing data by deep learning algorithms trained on the corresponding complementary stations. The two tested generative algorithms showed the superiority of the diffusion-based model, however the results still need improvement before tackling medically relevant problems. Importantly, the current infrastructure allows researchers to explore and test different generative methods and use diverse data hubs without compromising sensitive medical information. Ongoing work is focused on the preparation of segmented data sets to test the results of distributive data enhancement on clinically relevant tasks. Another possible approach to decreasing the data bias, supported by our infrastructure, is incremental learning. It could provide an alternative to the generative models by sending the central algorithm for training to different stations, hence taking advantage of the diverse data directly and not via synthetic supplements.

Our work presents the capabilities of decentralized data analysis infrastructure in the healthcare sector, which enables the maintenance of high-security standards while supporting data-driven research. Well-built data infrastructure can help identify and reduce data bias and improve the safety and quality of AI applications in the medical domain.

## Acknowledgement

This work was done as part of the NFDI4Health Consortium and is published on its behalf (www.nfdi4health.de). We acknowledge the financial support of the DFG, German Research Foundation, project number 442326535.

## References

[1] Ricci Lara MA, Mosquera C, Ferrante E, Echeveste R. Towards Unraveling Calibration Biases in Medical Image Analysis. In: Clinical Image-Based Procedures, Fairness of AI in Medical Imaging, and Ethical and Philosophical Issues in Medical Imaging. Springer Nature Switzerland; 2023. p. 132-41.

[2] Zhang Z, Zhang X, Ichiji K, Bukovsky I, Homma N. How intra-source imbalanced datasets impact the performance of deep learning for COVID-19 diagnosis using chest X-ray images. Scientific Reports. 2023 11;13.

[3] Banerjee I, Bhattacharjee K, Burns JL, Trivedi H, Purkayastha S, Seyyed-Kalantari L, et al. “Shortcuts” Causing Bias in Radiology Artificial Intelligence: Causes, Evaluation, and Mitigation. Journal of the American College of Radiology. 2023;20(9):842-51.

[4] Busby L, Courtier J, Glastonbury C. Bias in Radiology: The How and Why of Misses and Misinterpretations. RadioGraphics. 2017 12;38:170107.

[5] Zhang L, Wen X, Li JW, Jiang X, Yang XF, Li M. Diagnostic error and bias in the department of radiology: a pictorial essay. Insights into Imaging. 2023 10;14.

[6] Welten S, Mou Y, Neumann L, Jaberansary M, Ucer Y, Kirsten T, et al. A Privacy-Preserving Distributed Analytics Platform for Health Care Data. Methods of Information in Medicine. 2022 01;61.

[7] Tahar K, Martin T, Mou Y, Verbuecheln R, Graessner H, Krefting D. Rare Diseases in Hospital Information Systems—An Interoperable Methodology for Distributed Data Quality Assessments. Methods of Information in Medicine. 2023 Sep;62(03/04):071-89. Publisher: Georg Thieme Verlag KG.

[8] Li F, Wang Y, Beyan O, Scho¨neck M, Caldeira LL. Voxel-wise Medical Images Generalization for Eliminating Distribution Shift. ACM Trans Knowl Discov Data. 2024 Jan.

[9] Zhu JY, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In: Computer Vision (ICCV), 2017 IEEE International Conference on; 2017. .

[10] Chung H, Ye JC. Score-based diffusion models for accelerated MRI. Medical Image Analysis. 2022:102479.