THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Bias in AI:

Comparing Algorithm Performance Between US and African Hospitals

Grant Smith, MD and Jordan Perchik, MD

The American Dream (for AI in radiology)

Augment physicians, increase speed and accuracy

Help meet demands of increasing imaging utilization and complexity

Increase efficiency of practices, improve patient care and radiology workflow

Potential Pitfalls in LMICs

Instead of improving workflow in high volume setting, extend reach of physician in a setting with few physicians and limited imaging infrastructure

Algorithm trained on one dataset may perform poorly in other settings

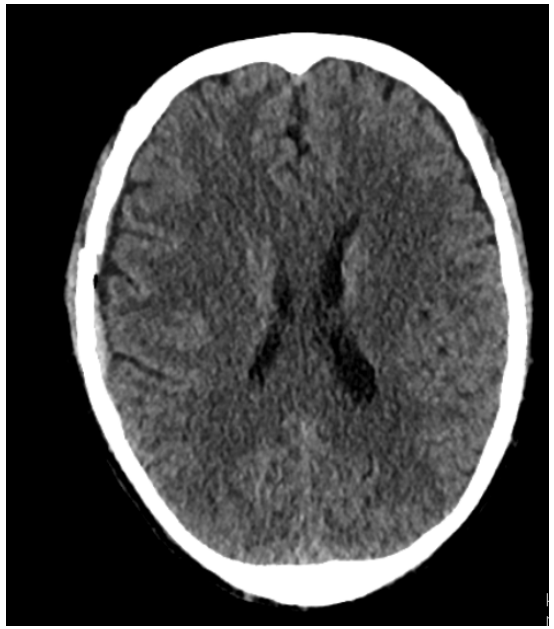
US based AI algorithms performing poorly in limited resource environments perpetuates healthcare disparities

Algorithm Bias

Differential performance of an AI algorithm between groups

Difficult to evaluate

Whose responsibility to manage, vendor or client?



Three examples of False negative studies from African cohort, ICH missed by algorithm.

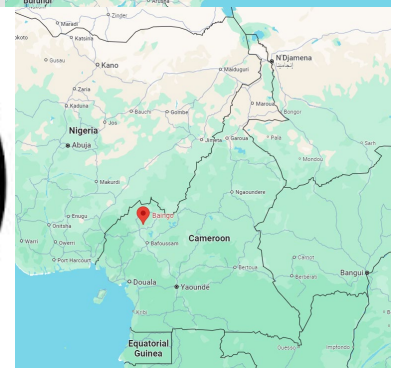
Global Health and Radiology

AI playing a role by increasing efficiency of triage and workflow, expanding access.

Telehealth/teleradiology allows host sites to aid LMIC health systems remotely

SODDO CHRISTIAN HOSPITAL

CAMEROON BAPTIST CONVENTION HEALTH SERVICES 1975



Does a US based AI Application perform well in Cameroon and Ethiopia?

US Cohort

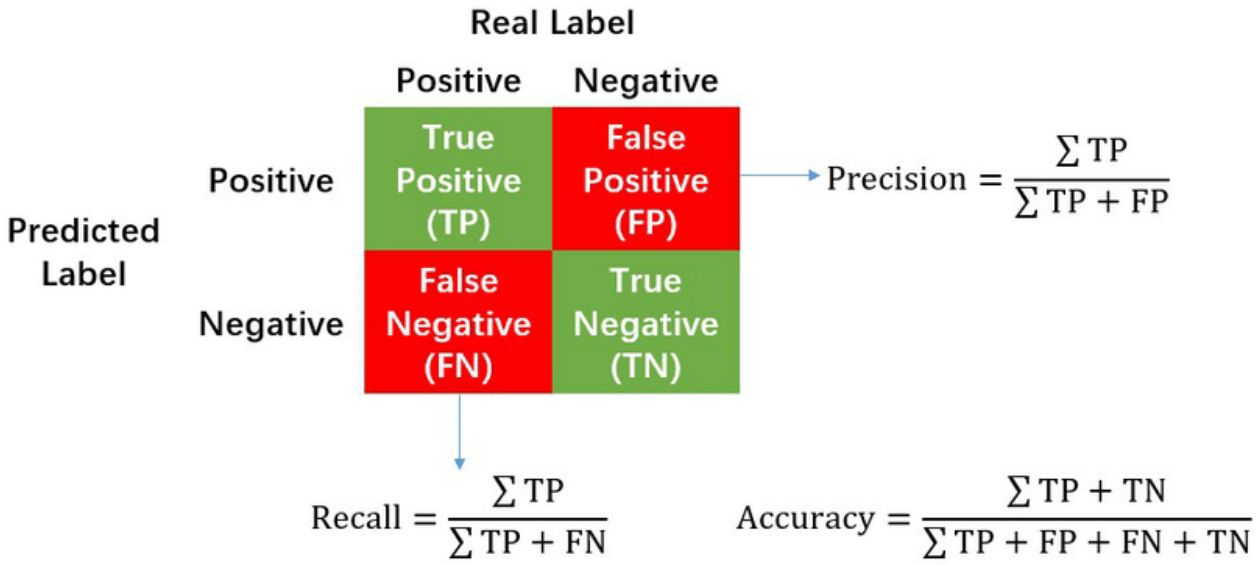
- High volume, multiple CTs per hospital
- Multiple subspecialized radiologists
- High imaging utilization
- Trauma and cancer center

International Cohort

- Single CT at each hospital
- Single general radiologist
- More selective imaging utilization
- Trauma and cancer center

Evaluating AI Algorithms

- FDA-approved ICH triage algorithm (CADt: computer assisted triage and notification)
- Head CTs performed at the Cameroon and Ethiopian sites reviewed by single reader and compared with AI result
- Evaluated recall, precision, and accuracy then compared to vendor calculated values
- January-July 2024



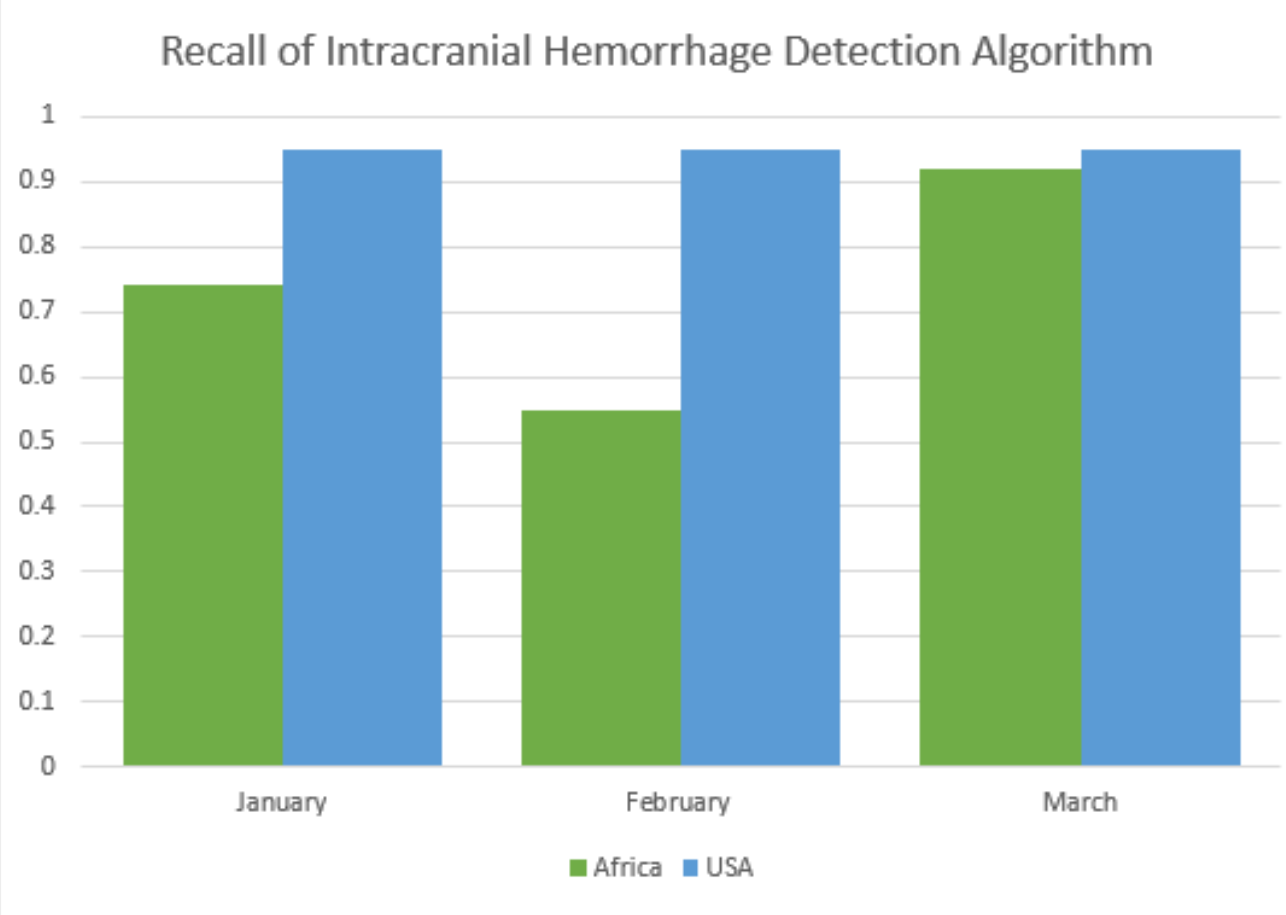
https://www.researchgate.net/figure/Calculation-of-Precision-Recall-and-Accuracy-in-the-confusion-matrix_fig3_336402347

Results

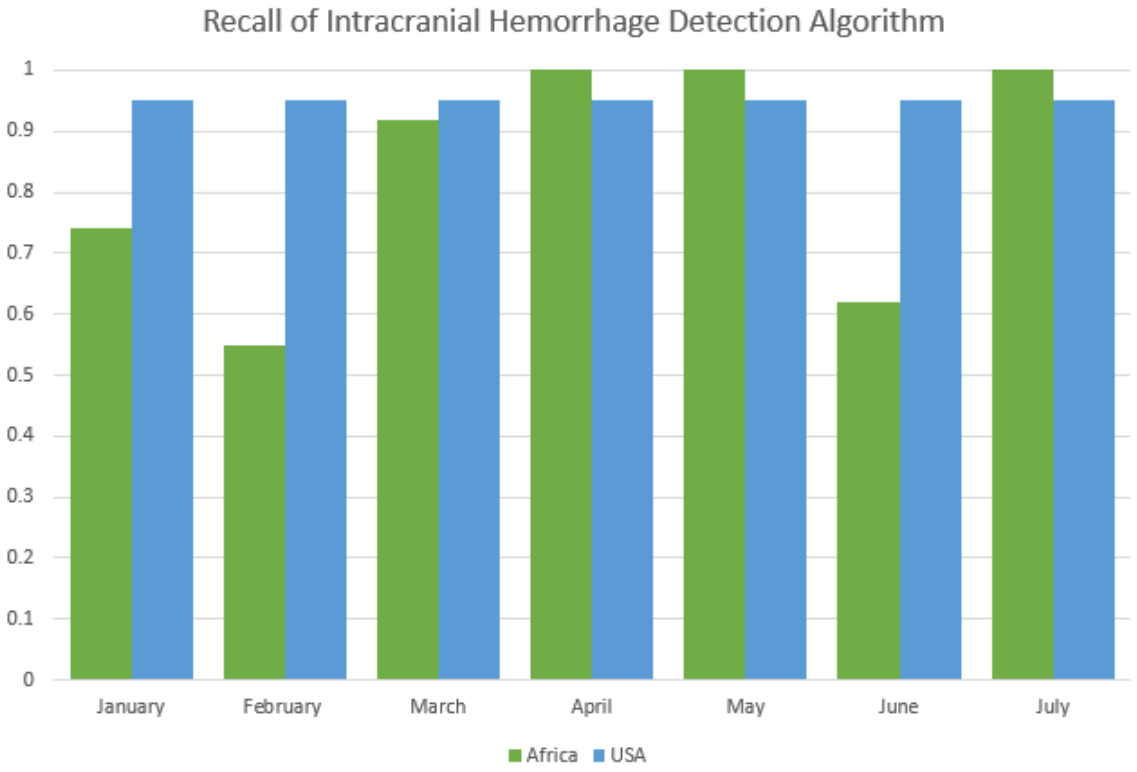
Total	486 Non-contrast head CTs performed at African sites (average 69.4 per month)
Recall	African site average 0.83 US average 0.95
Precision	African site average 0.92 US average 0.78

Results

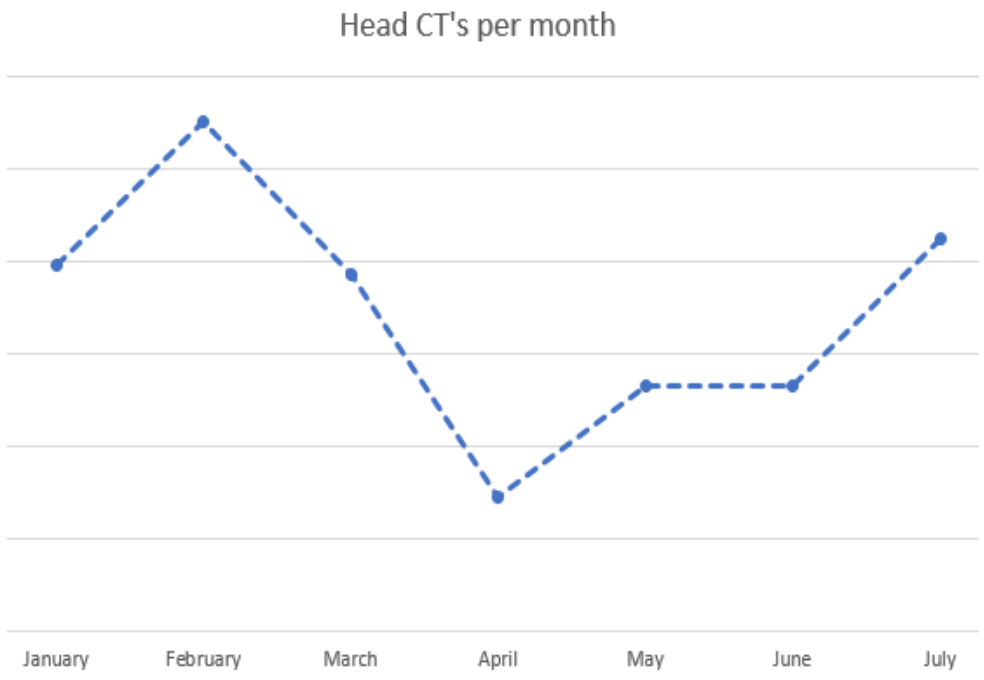
- Evidence of algorithm bias in February, communicated to vendor.
- High rate of false negative studies; informatics error discovered
- Error addressed; study continued



Results following change

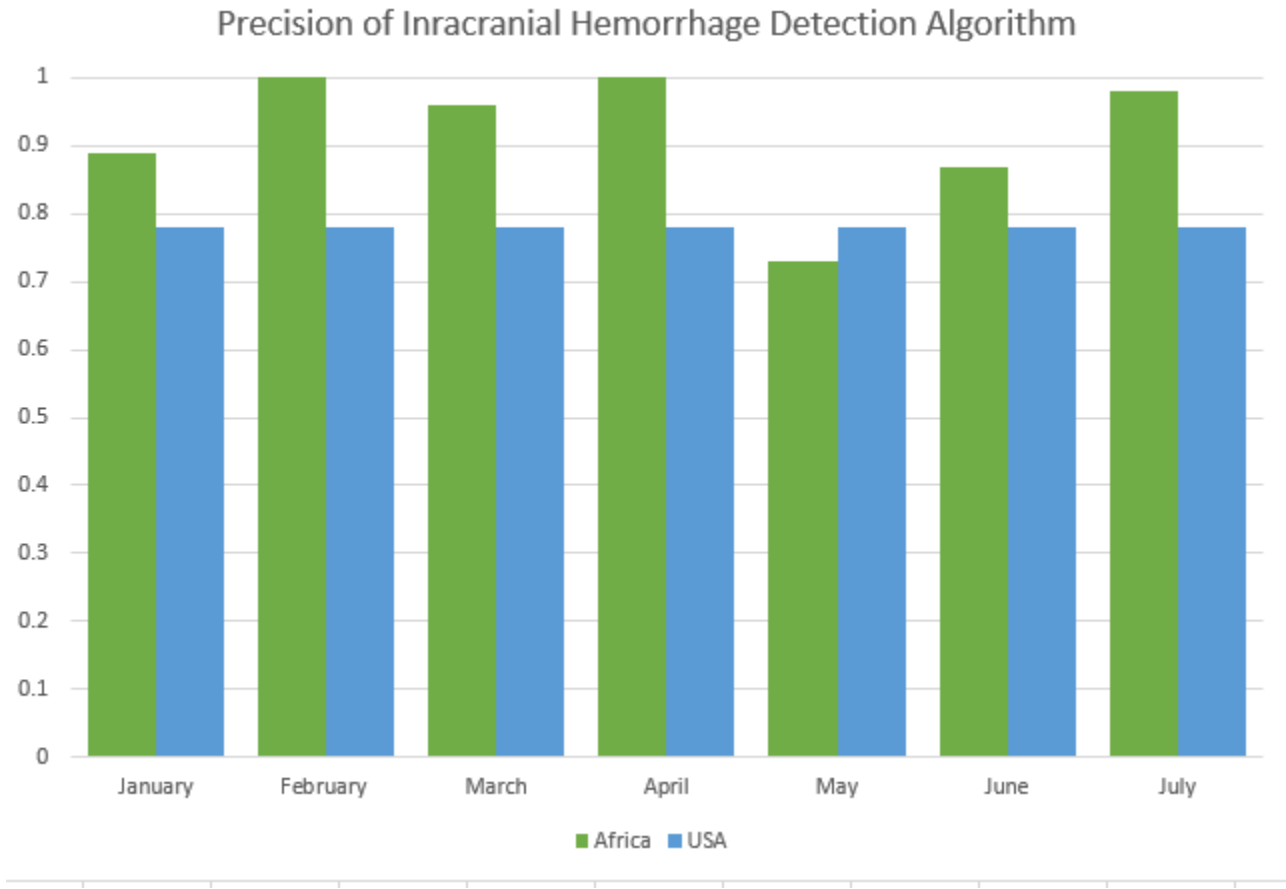


Algorithm processing adjustment performed in March. Low performance in June may be artifact secondary to low CT volume.



Results, cont.

- Precision significantly higher for African sites every month except May (p=0.01)
- Lower false positive rate



Discussion

Differential performance identified early in the study period, adjustment made which resulted in similar recall between US and African sites

Precision significantly higher for African sites

Bias difficult to detect, multitude of components involved

Radiologists ultimately liable, and need to be versed in use of AI and collaborate with vendors to assure high algorithm performance

Conclusion

Radiology AI algorithms can potentially be used to expand healthcare access and reduce disparities in LMICs, with the caveat that they must be carefully evaluated for algorithm bias and performance characteristics.

Vendors and end users can work together to improve the performance of AI algorithms in diverse settings, beyond the original use scenario.

Intracranial hemorrhage detection algorithm performed at a similar level in US and African sites, with better algorithm precision on African data set.

Thank you for your attention

grantsmith@uabmc.edu

References

- •Ciecierski-Holmes T, Singh R, Axt M, et al. Artificial Intelligence for Strengthening Healthcare Systems in Low- and Middle-income Countries: A Systematic Scoping Review. Npj Digital Medicine. 2022. 5(162). Published online October 28, 2022.
- •Mollura DJ, Culp MP, Pollack E, et al. Artificial Intelligence in Low and Middle-Income Countries: Innovating Global Health Radiology. Radiology. 2020;297:513-520. doi: 10.1148/radiol.2020201434.
- •Abuzaid MM, Elshami W, McConnel J, Tekin HO. An Extensive Survey of Radiographers from the Middle East and India on Artificial Intelligence Integration in Radiology Practice. Health Technol. 2021;11:1045-1050. doi: 10.1007/s12553-021-00583-1.
- •Celi LA, Cellini J, Charignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities – A global review. PLOS Digit Health. Published online 31 March 2022. DOI: 10.1371/journal.pdig.0000022
- •Goirand M, Austin E, Clay-Williams R. Implementing ethics in healthcare AI-based applications: A scoping review. Sci Eng Ethics. Published online: 3 September 2021. DOI: 10.1007/s11948-021-00336-3.
- •Li F, Ruijs N, Lu Y. Ethics & AI: A systematic review on ethical concerns and related strategies for designing with AI in healthcare. AI. 2022; 4(1): 28-53. DOI: 10.3390/ai4010003