P35

CIBM Annual Symposium 2024

Forum Rolex Learning Center, EPFL, Lausanne Switzerland | 7th November 2024

20th Anniversary

CIBM Center for Biomedical Imaging

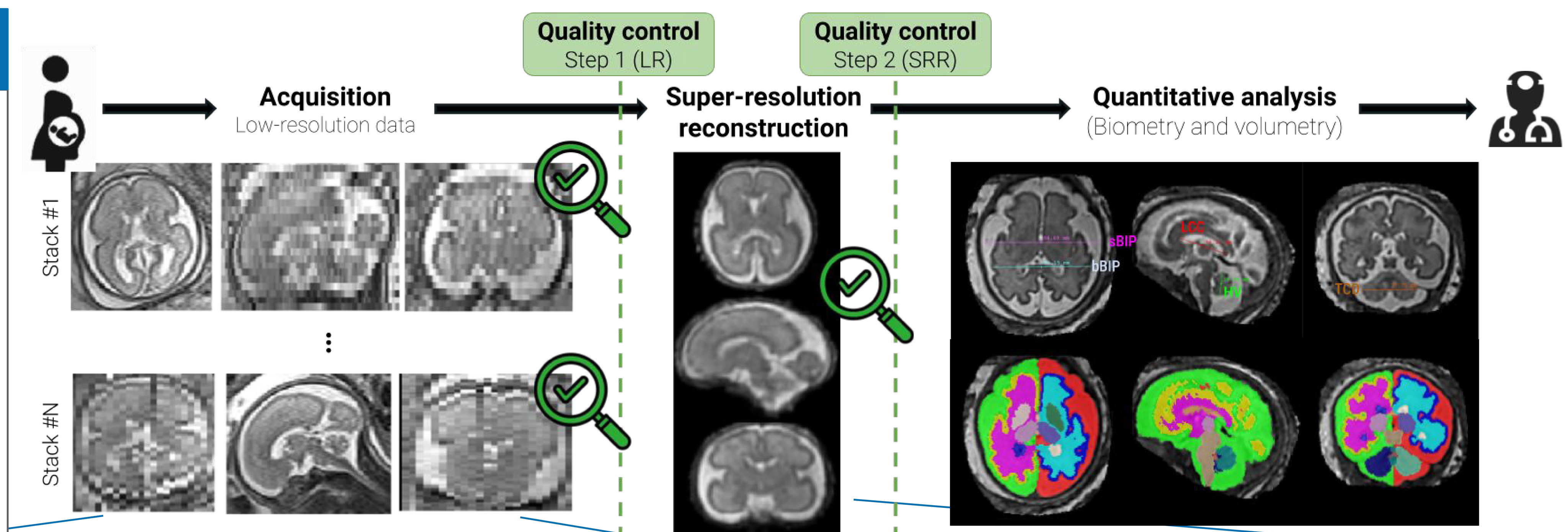
Improving reliability in fetal brain MRI analysis

Thomas Sanchez[1,2], Oscar Esteban[2], Angeline Mihailov[3], Yvan Gomez[4,5], Alexandre Pron[3], Gerard Martí Juan[6], Mériam Koob[2], Vincent Dunet[2], Nadine Girard[3,7], Andras Jakab[8,9,10], Elisenda Eixarch[5,11], Guillaume Auzias[3], Meritxell Bach Cuadra[1,2]

[1]CIBM – Center for Biomedical Imaging
[2]Department of Diagnostic and Interventional Radiology, CHUV-UNIL
[3]Aix-Marseille Université, CNRS, Institut de Neurosciences de La Timone, Marseilles
[4]Department Woman-Mother-Child, CHUV
[5]BCNatal Fetal Medicine Research Center (Hospital Clínic and Hospital Sant Joan de Déu), Universitat de Barcelona
[6]Universitat Pompeu Fabra, Barcelona
[7]Service de Neuroradiologie Diagnostique et Interventionnelle, Hôpital Timone, AP-HM, Marseilles
[8]Center for MR Research, University Children's Hospital Zurich, University of Zurich
[9]Neuroscience Center Zurich, University of Zurich
[10]Research Priority Project Adaptive Brain Circuits in Development and Learning (AdaBD), University of Zürich
[11]IDIBAPS and CIBERER, Barcelona

Context

- Complex acquisition and processing pipeline in fetal brain MRI
- Heterogeneous data and domain shifts



Question

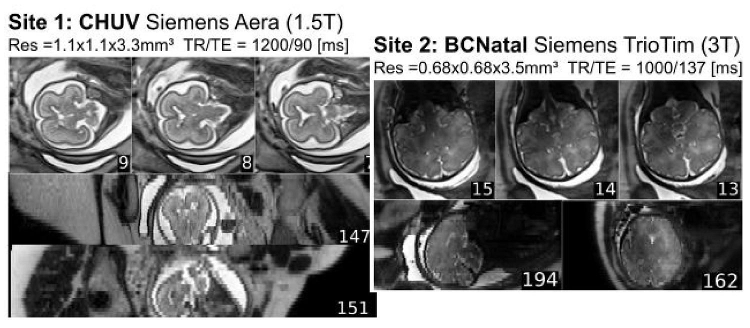
How can we ensure reliability at every step of the pipeline?

1. Quality control
2. Analysis of biases and reproducibility

Tackling domain shifts with FetMRQC [2]

The problem. Heterogeneity across scanners and sites → Machine learning models fail to generalize [1].

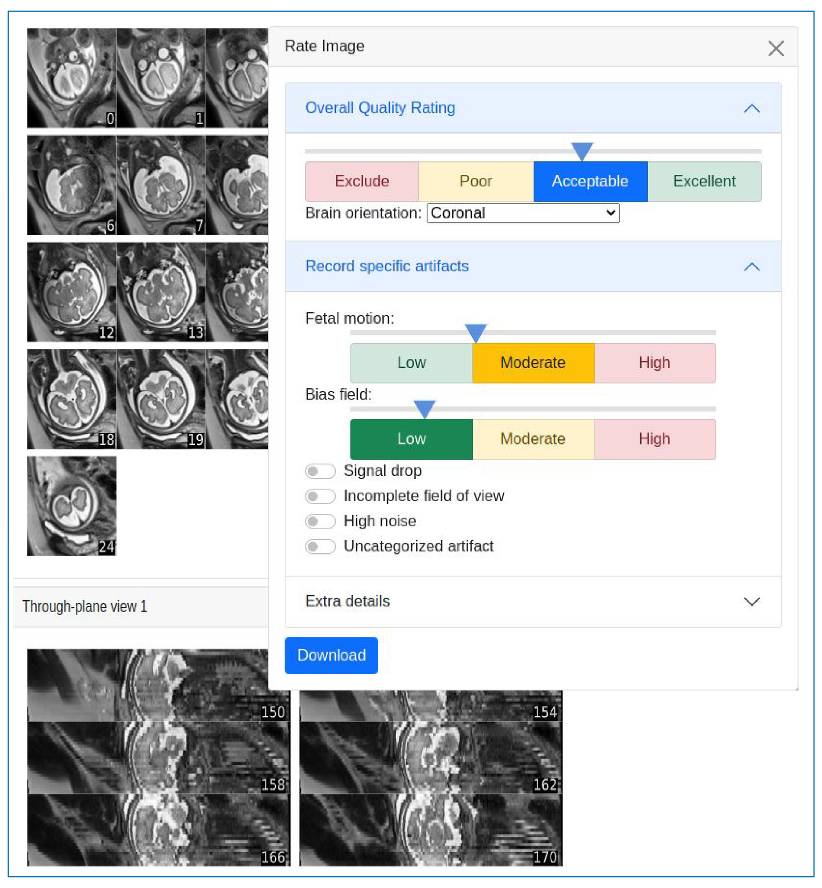
Our solution



Step 1. Standardized ratings.

Annotation interface and multiple raters [3]

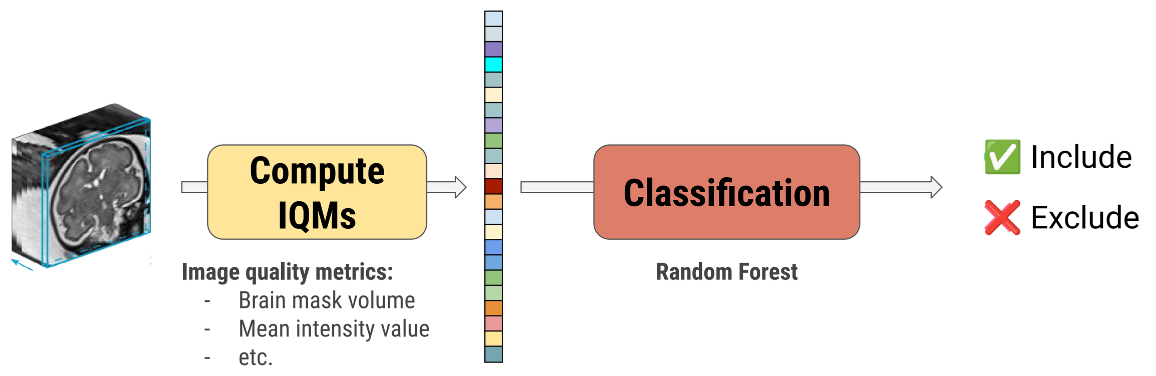
Result. Two experts annotated more than 1600 LR T2w scans from 13 scanners across 4 hospitals.



Annotation interface based on MRIQC [3]

Step 2. Automated prediction of quality.

Insight. Use a simple model. More complex models using nested cross validation and more sophisticated predictors failed to generalize out-of-domain.

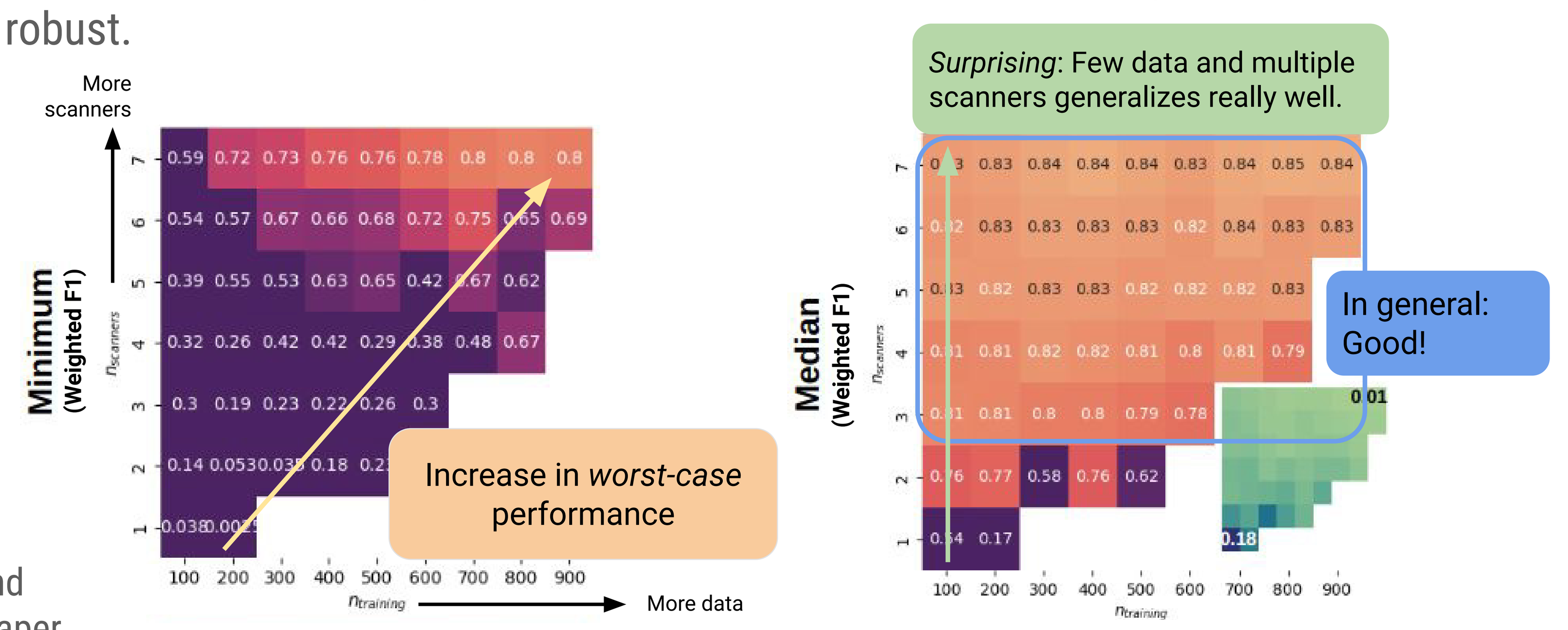


Step 3. Robust evaluation.

Use leave-one-site-out cross-validation and report worst-split performance with multiple metrics.

Experiment. How well can we expect to generalize given limited # of scanners and # of training data ?

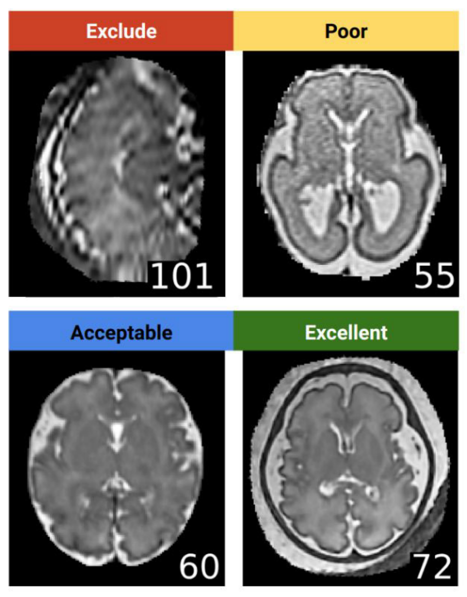
Insight. Heterogeneous data in training are key to a strong *median* generalization ability. More data helps make the model more robust.



More details and results in the paper

Towards more reproducible quality control [4]

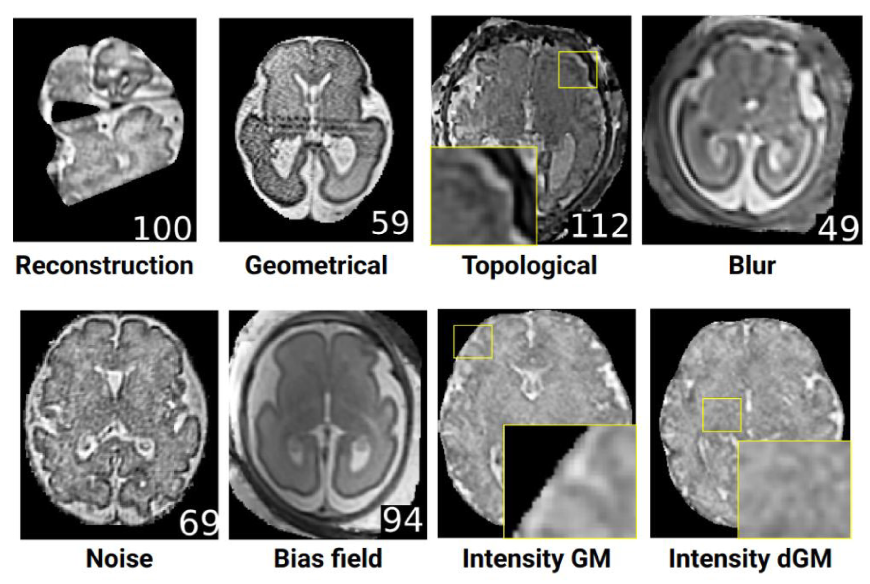
The problem. How reproducible are quality annotations from different raters? Can we make the quality rating depend on *specific* criteria rather than subjective assessment?



Step 1. A taxonomy of quality.

Super-resolution reconstruction can lead to various data quality.

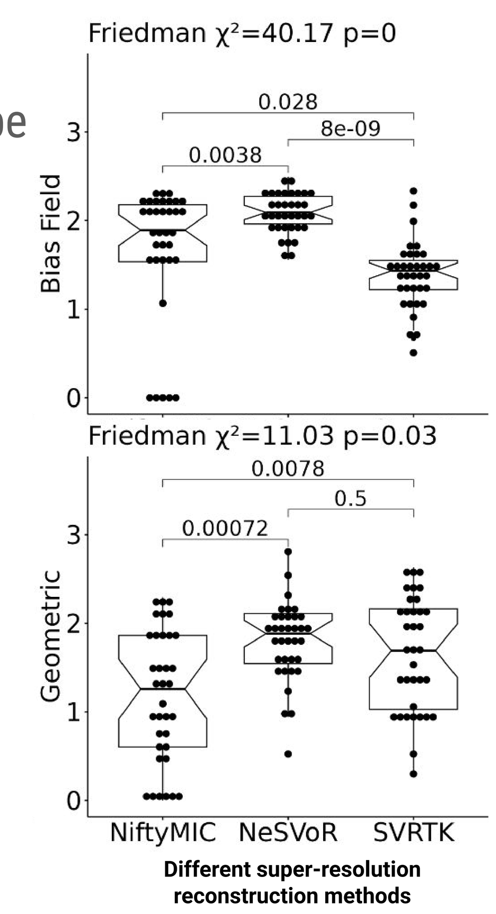
Data can feature very different *artifacts.*



Step 2. Multi-annotator rating.

105 reconstructions annotated twice by four raters.

Result.
1. The global quality score is highly reliable. Most specific criteria can be reliably rated.
2. The scores allow us to understand where different SRR methods underperform



More details and results in the paper

mial Medical Image Analysis Lab · UNIL | Université de Lausanne · neuron · Swiss National Science Foundation · CHUV Centre hospitalier universitaire vaudois · upf. Universitat Pompeu Fabra Barcelona · BCNATAL · Fetal Medicine Research Center · int · Universitäts-Kinderspital Zürich

Financial support: SNSF 31NE30_203977 - Era-Net Neuron MULTIFACT

References.
[1] Dockès J. et al. "Preventing dataset shift from breaking machine-learning biomarkers." GigaScience (2021)
[2] Sanchez T. et al. "FetMRQC: A robust quality control system for multi-centric fetal brain MRI." *MedIA* (2024)
[3] Esteban O. et al. "MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites." *PloS one* (2017)
[4] Sanchez T. et al. "Assessing data quality on fetal brain MRI reconstruction: a multi-site and multi-rater study." MICCAI PIPPI Workshop (2024)