

Deep Learning-based Pseudo-Mass Spectrometry Imaging Analysis for Precision Medicine

Xiaotao Shen PhD

Stanford University School of Medicine
Department of Genetics
Dr. Snyder lab

October 9th, 2022 @ CASMS



shenxt@stanford.edu



[@xiaotaoshen1990](https://twitter.com/xiaotaoshen1990)



shenxt.info



github.com/jaspershen

About Me

 **Name:** Xiaotao Shen/申小涛

 **Affiliation:** Stanford University School of Medicine

 **Title:** Postdoctoral Research Fellow

 **Email:** shenxt@stanford.edu

 **Personal website:** shenxt.info

 **GitHub:** github.com/jaspershen

 **Twitter:** @xiaotaoshen1990



Research:

Metabolomics & Bioinformatics

Microbiome

Multi-omics

Systems biology

Aging

Precision medicine

</> **Background**

</> **DeepPseudoMSI project: Deep Learning-based Pseudo-Mass Spectrometry Imaging**

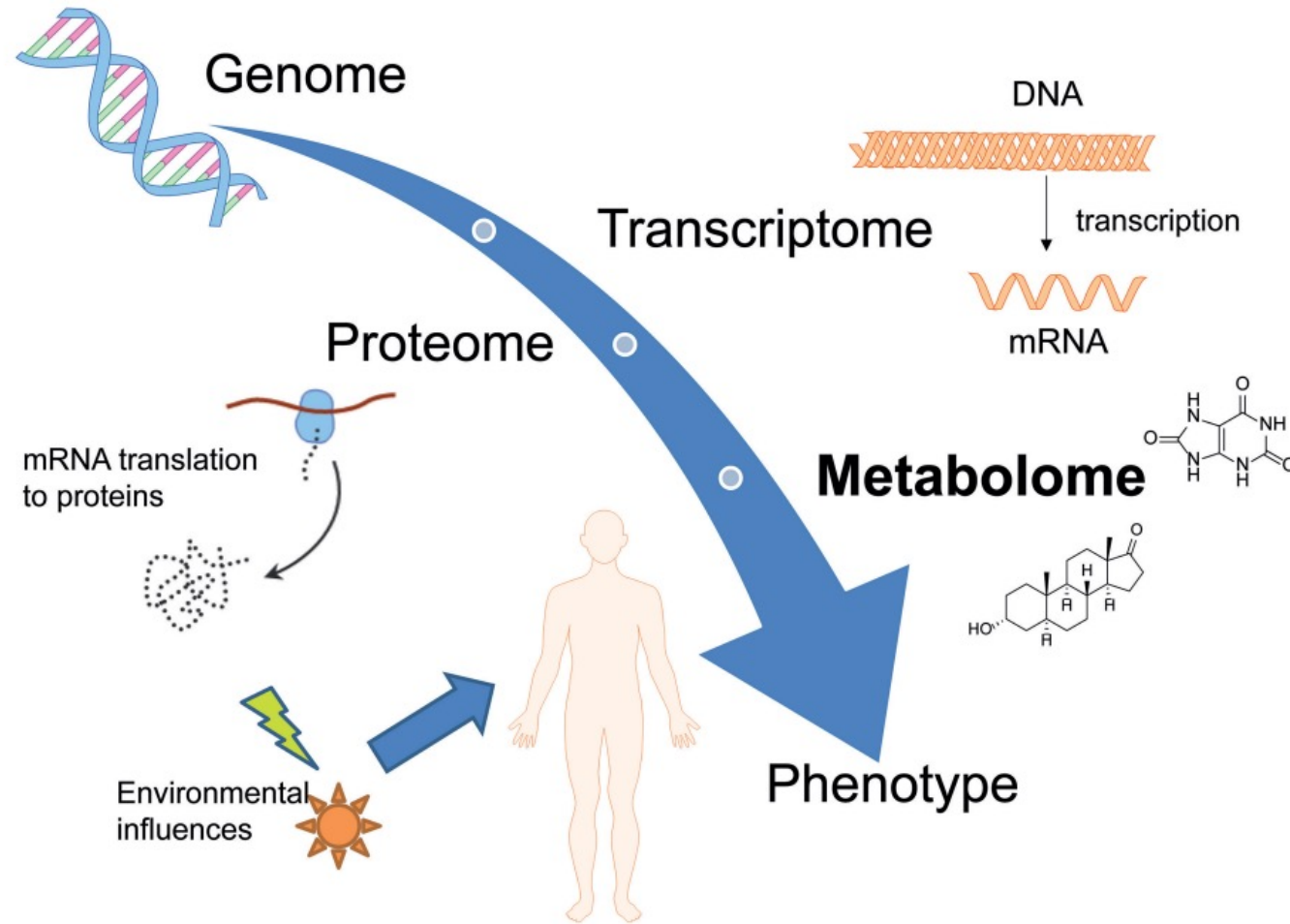
</> **Applications of deepPseudoMSI**

</> **Summary**

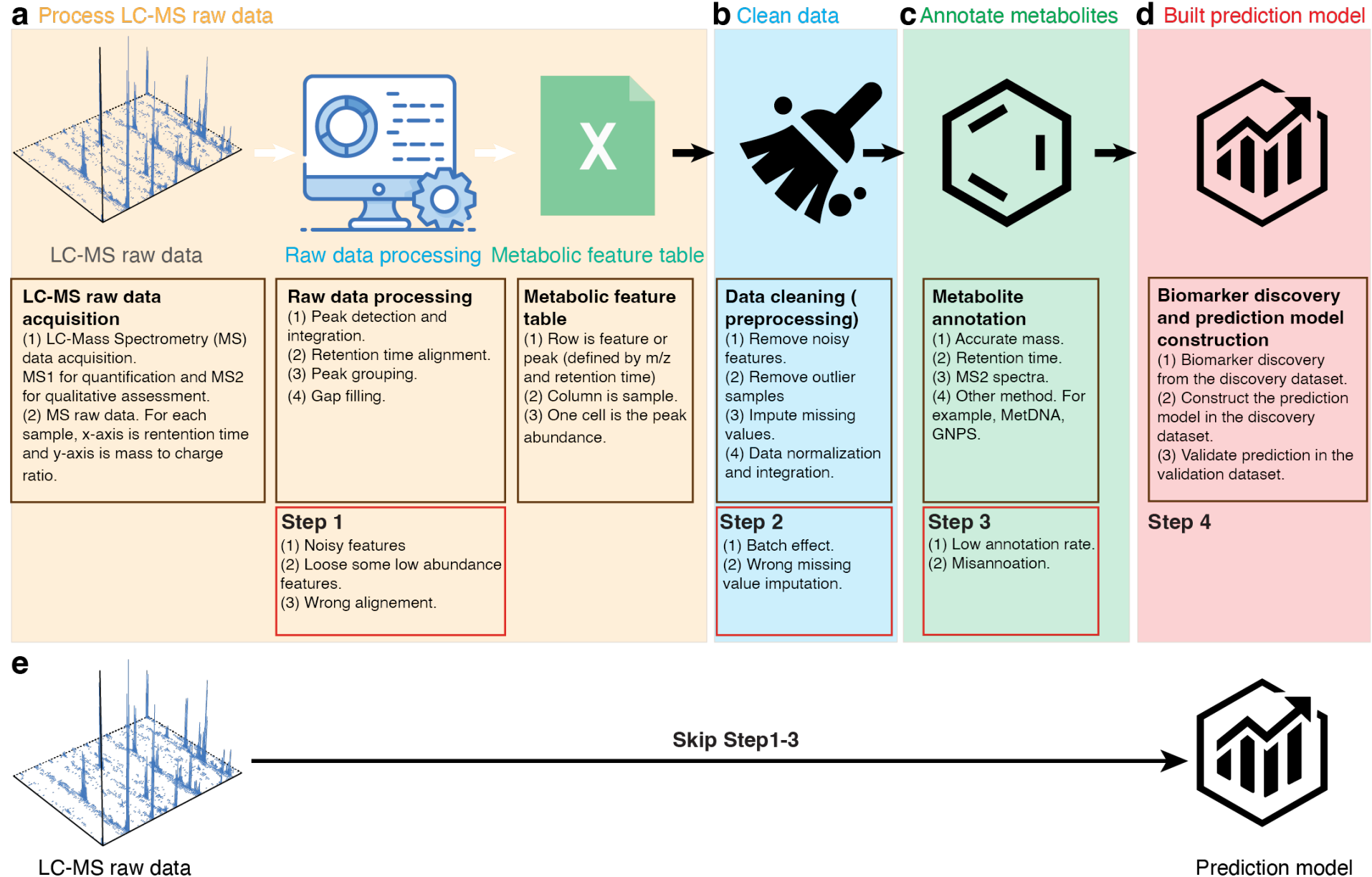


Background

Metabolomics/lipidomics



Conventional approaches for disease diagnosis using LC-MS



➤ Disadvantages of conventional methods

Identification of all metabolites is extremely difficult

1. Most of the peaks have no MS² spectra (Only 30%-40% peaks have MS² spectra).
2. Some MS² spectra quality are not high.
3. We don't have enough MS² spectra.
4. We can not avoid the false identification.

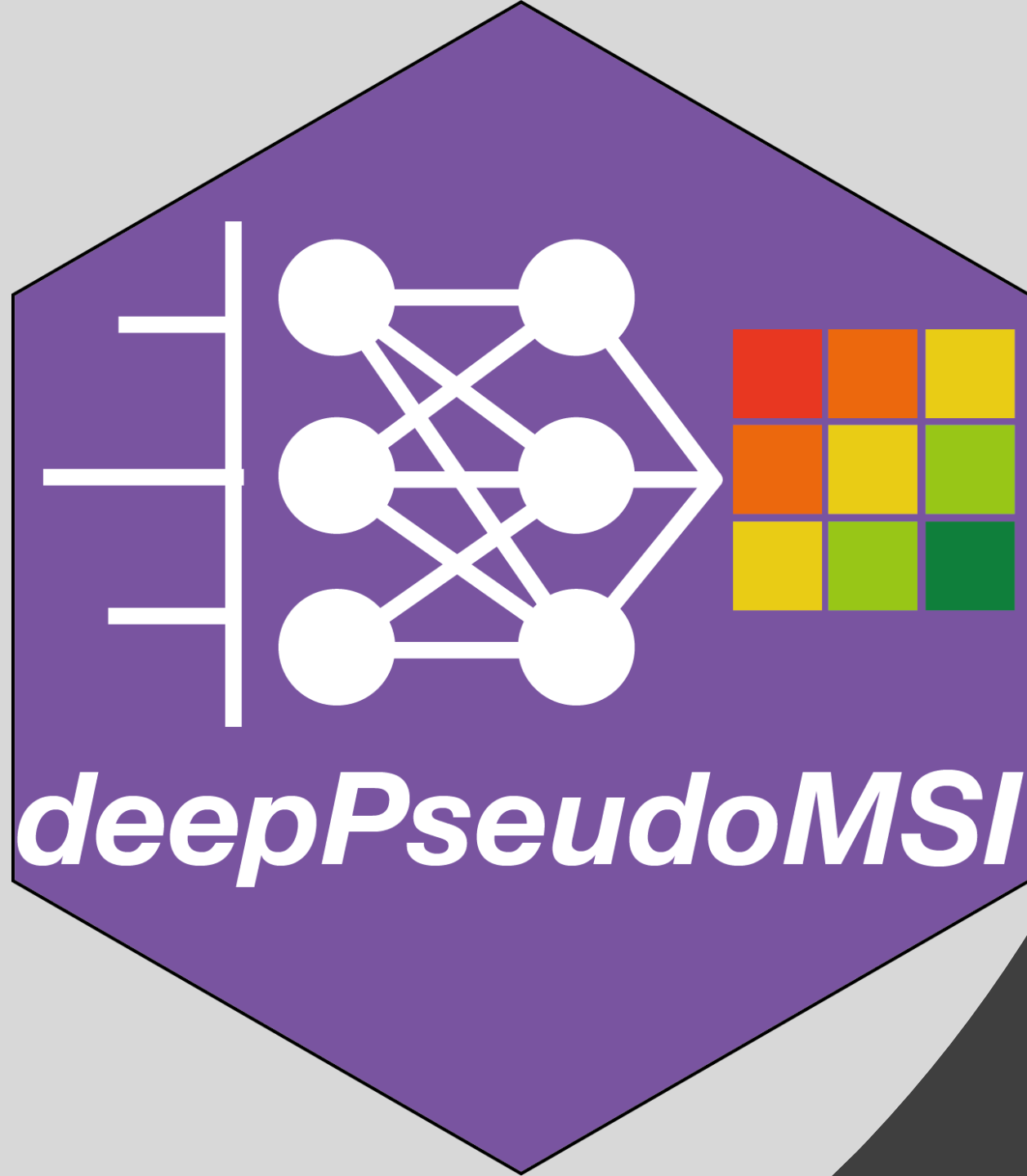
Prediction model with metabolites/lipids

1. **A lot of peaks with good prediction power will be lost.**
2. It is extremely time-costing to identify metabolites (level 1, standard confirmation).

Prediction model with peaks

1. Peaks are not robust.
2. It is difficult to use peaks in clinical application.

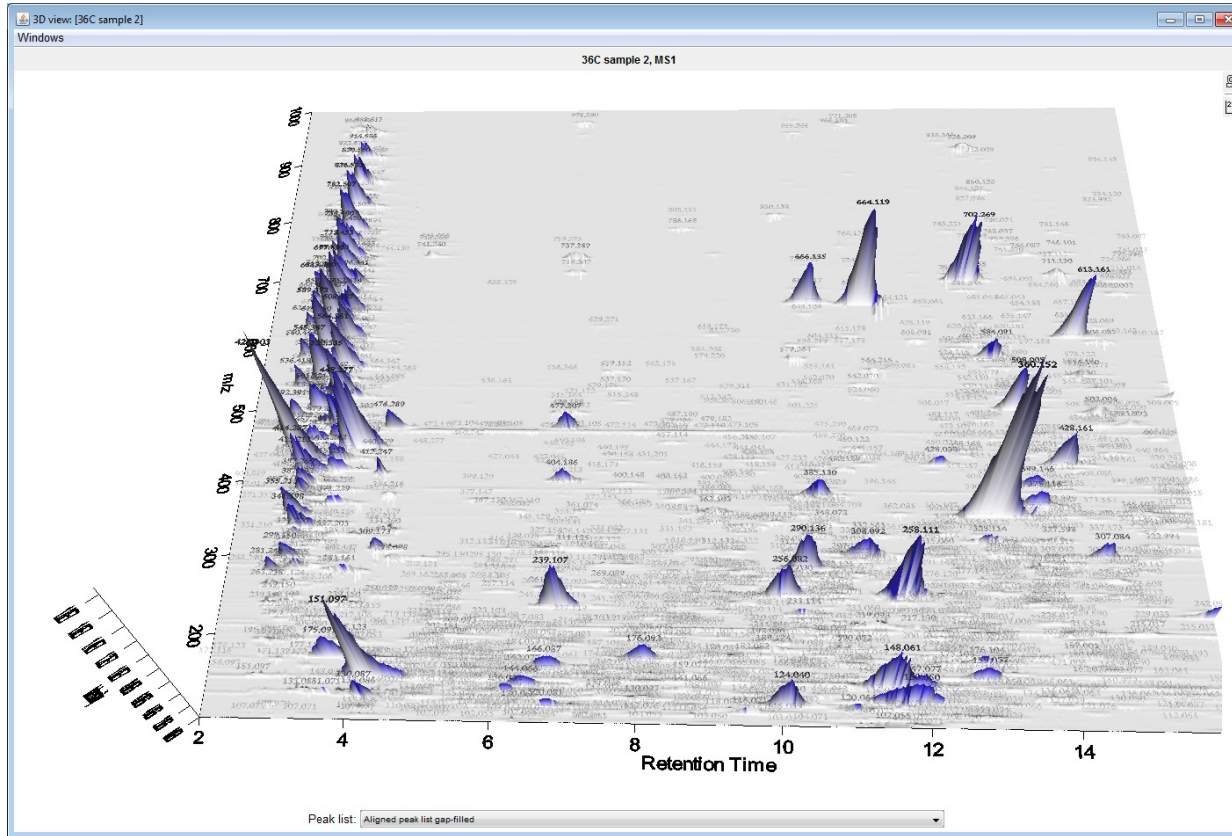
m/z, retention time and intensity (inter- intra-batch effect) shift



DeepPseudo MSI project

➤ LC-MS raw data: Image (pseudo Mass Spectrometry Image, pseudoMSI)

Raw MS data (Millions data points)



Metabolic image (pseudoMSI)



Intensity

Retention time

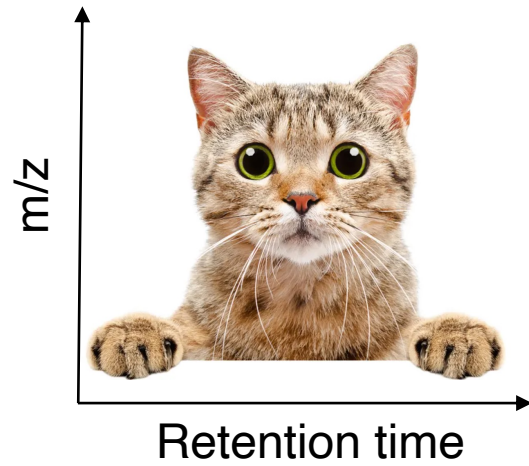
➤ Advantages of pseudoMS image

1. All the information are in the image



➤ Advantages of pseudoMS image

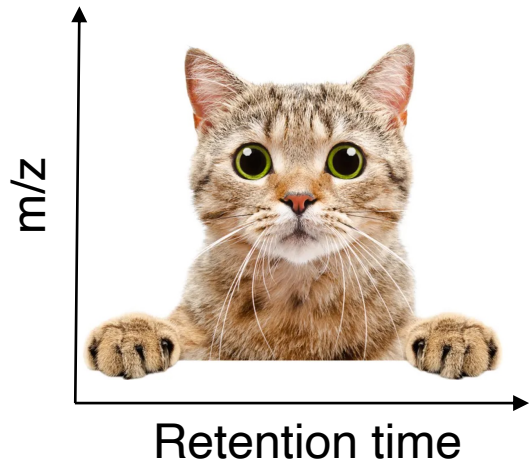
2. Handle batch effect



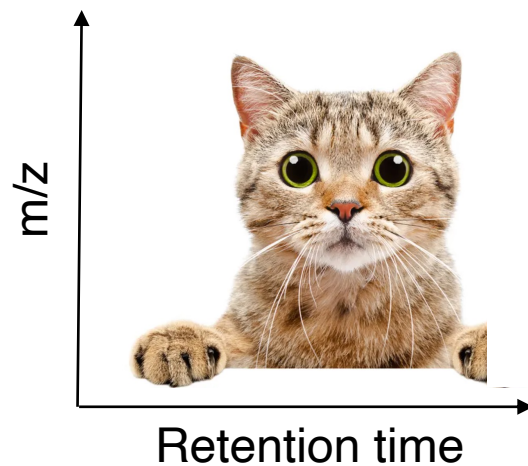
Raw data

➤ Advantages of pseudoMS image

2. Handle batch effect



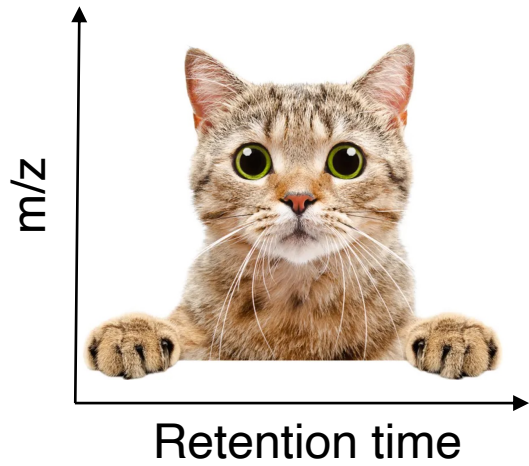
Raw data



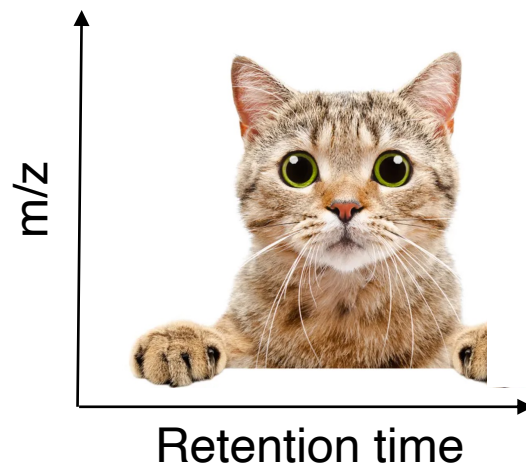
RT shift

➤ Advantages of pseudoMS image

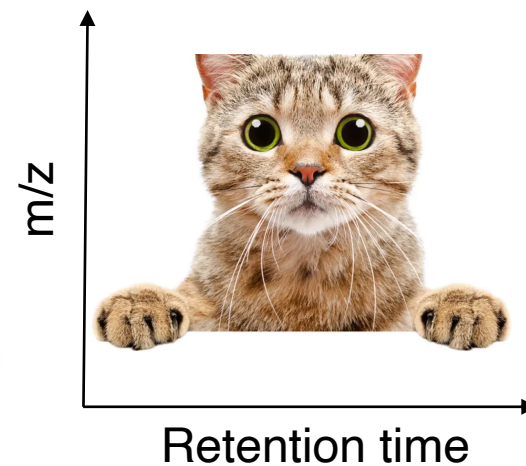
2. Handle batch effect



Raw data



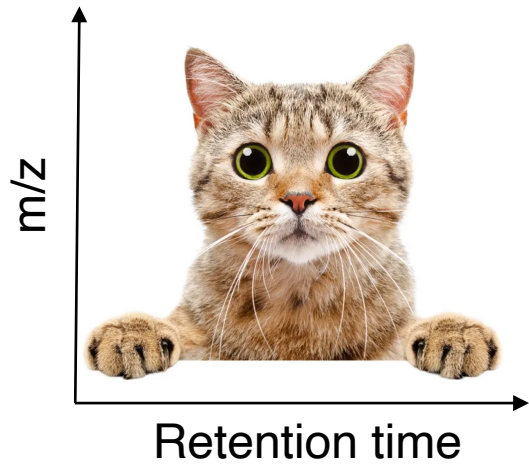
RT shift



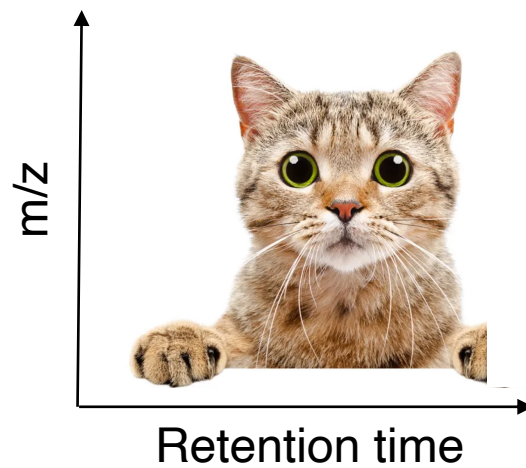
m/z shift

Advantages of pseudoMS image

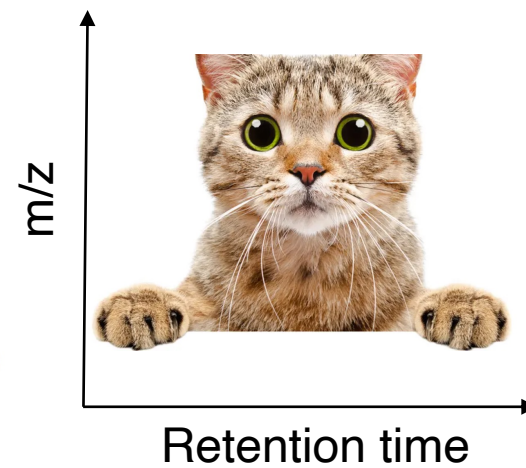
2. Handle batch effect



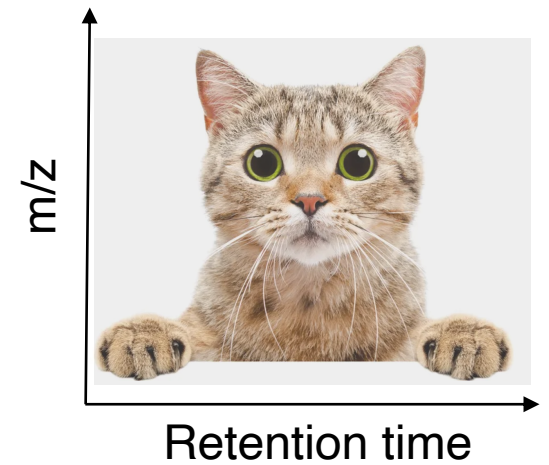
Raw data



RT shift

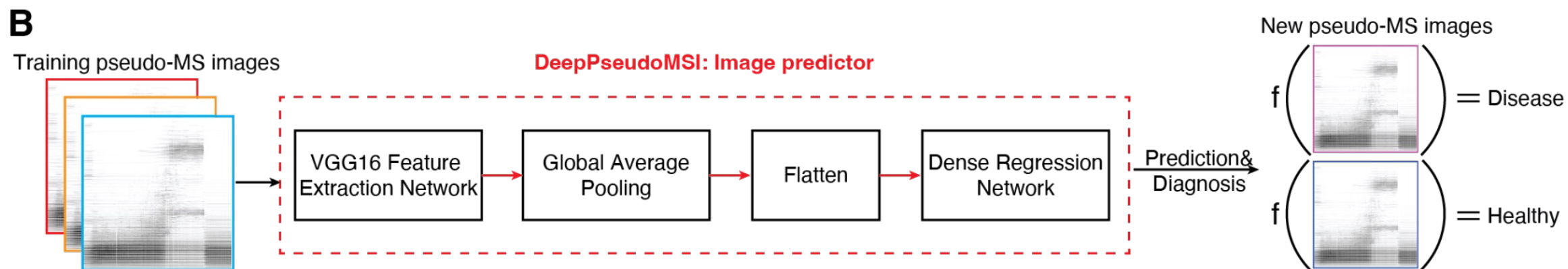
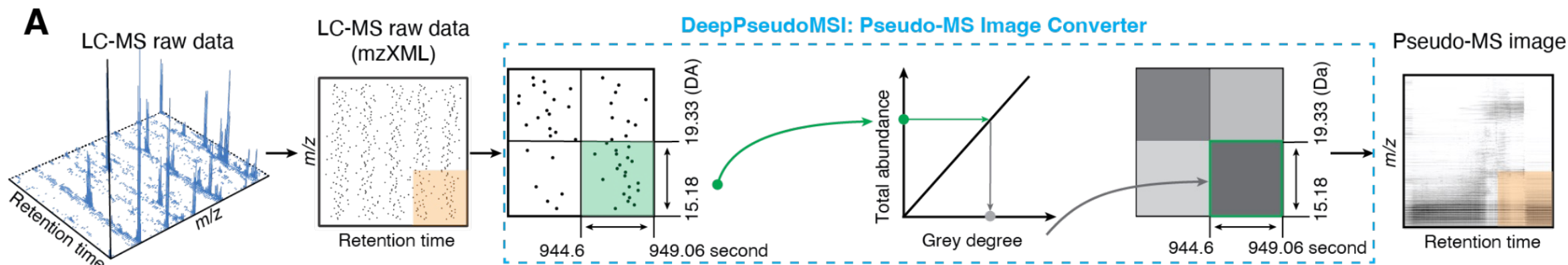


m/z shift

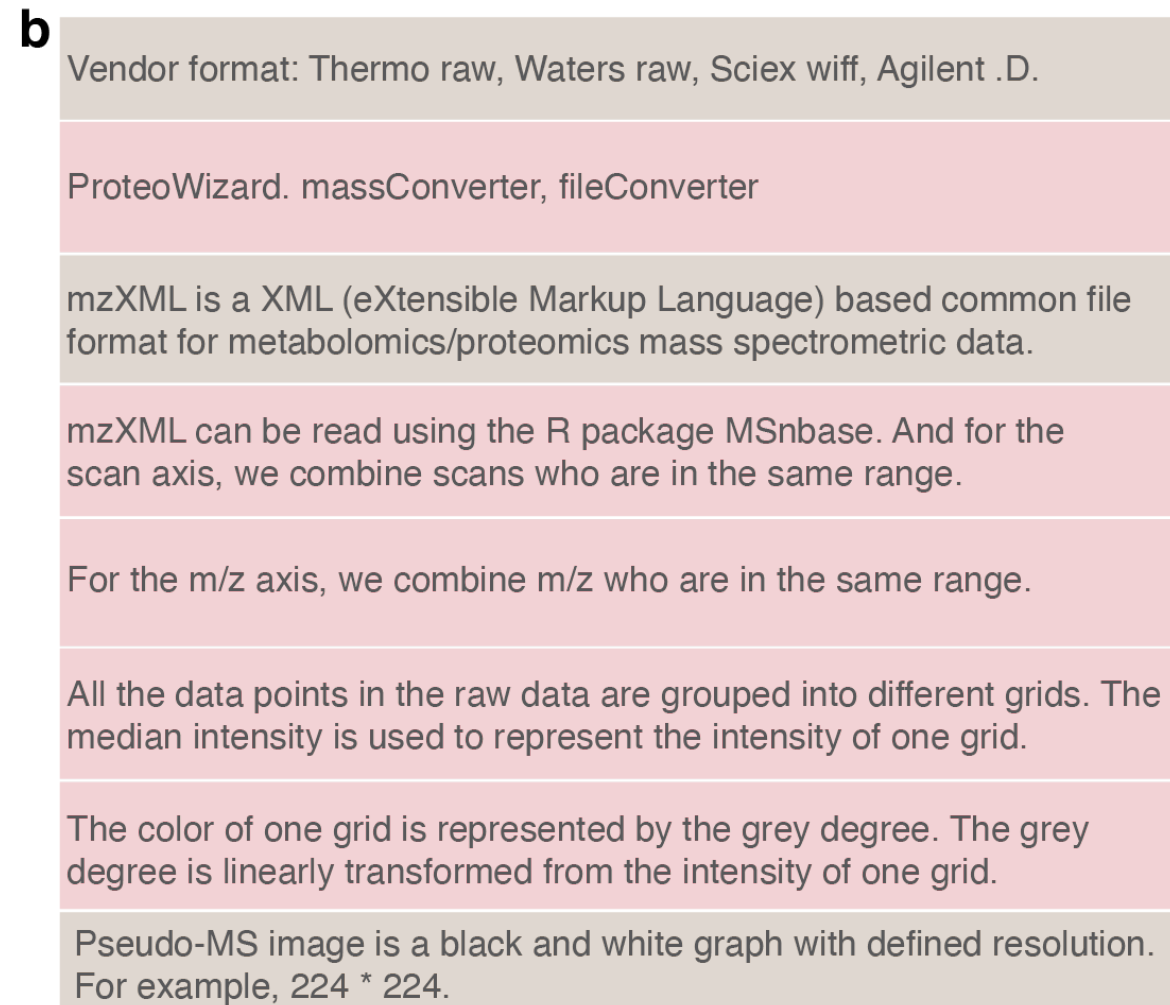
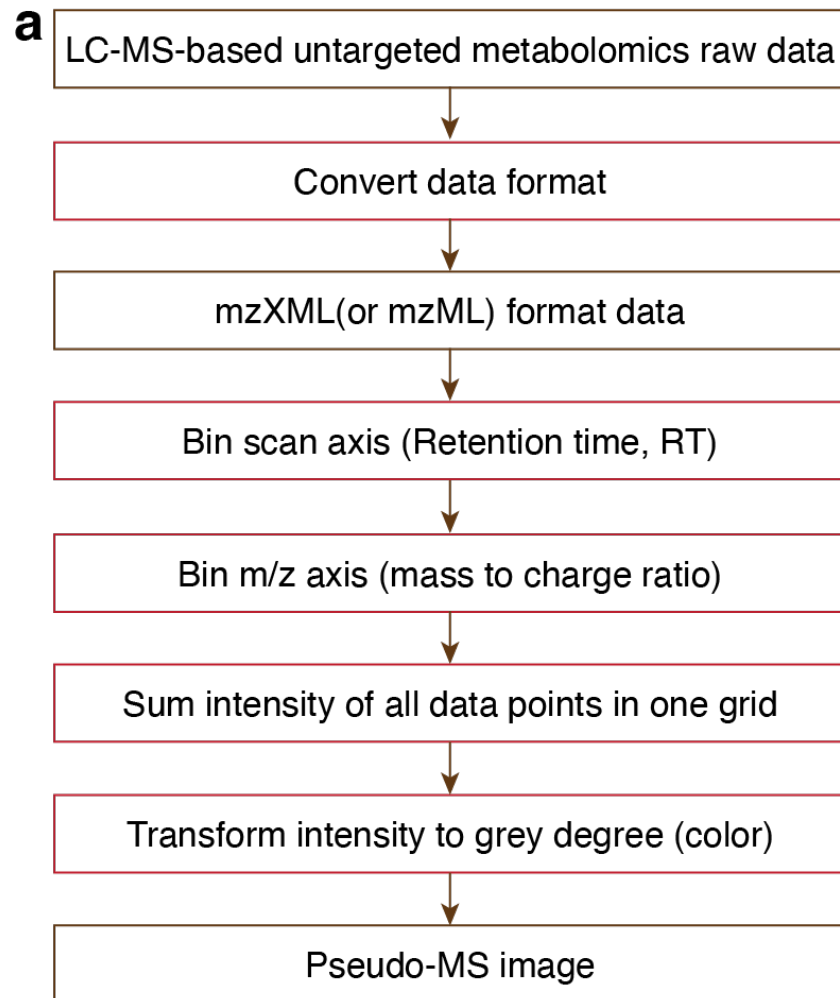


Intensity shift

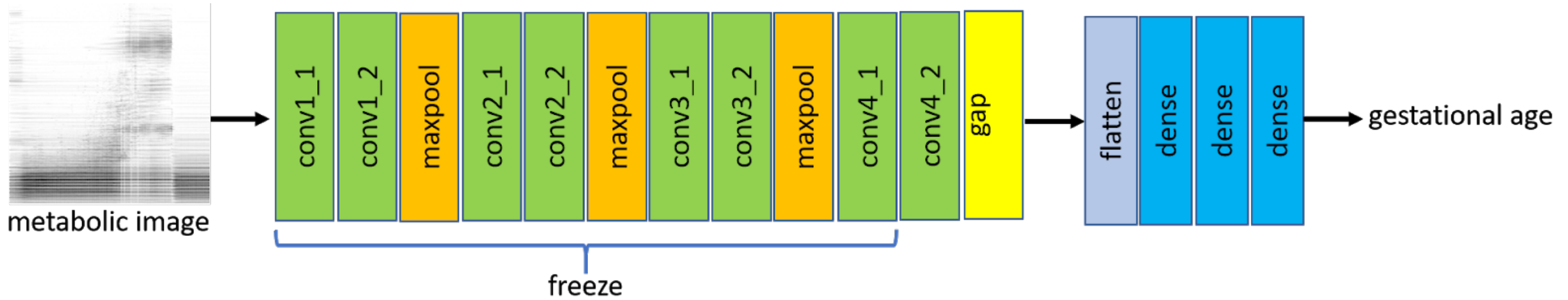
Workflow of deepPseudoMSI



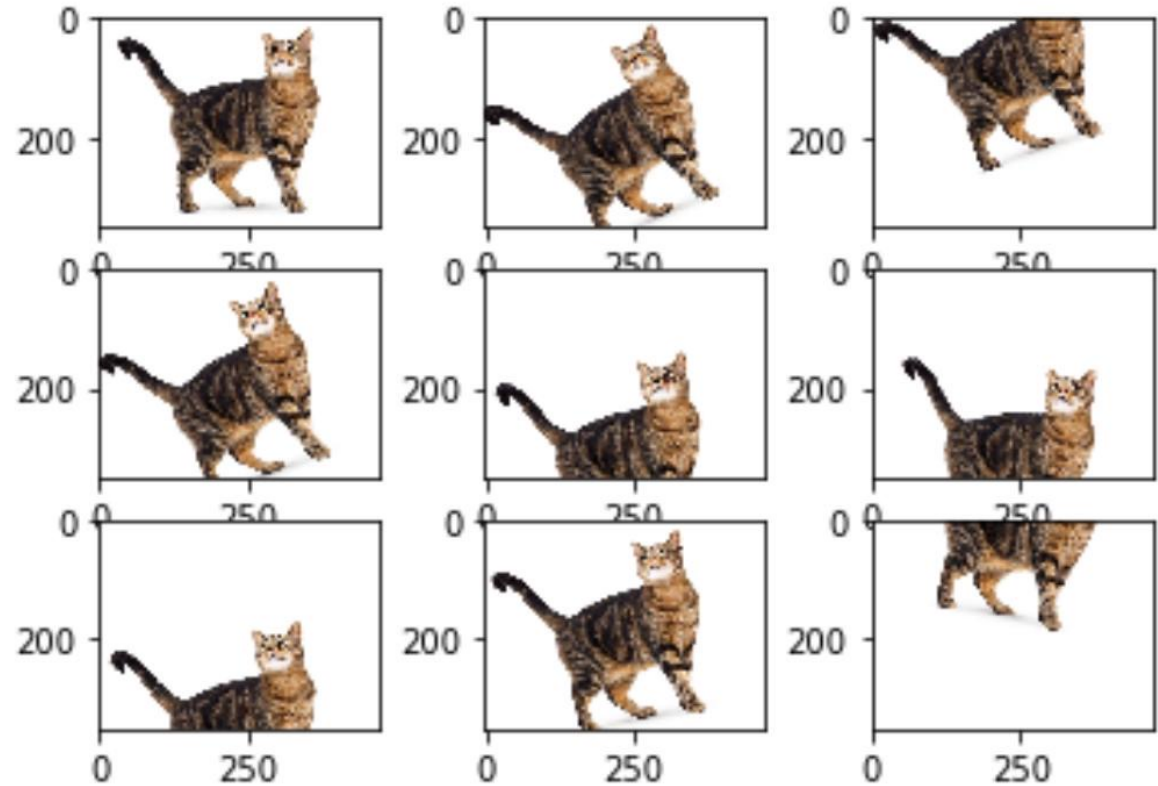
> PseudoMS Image converter



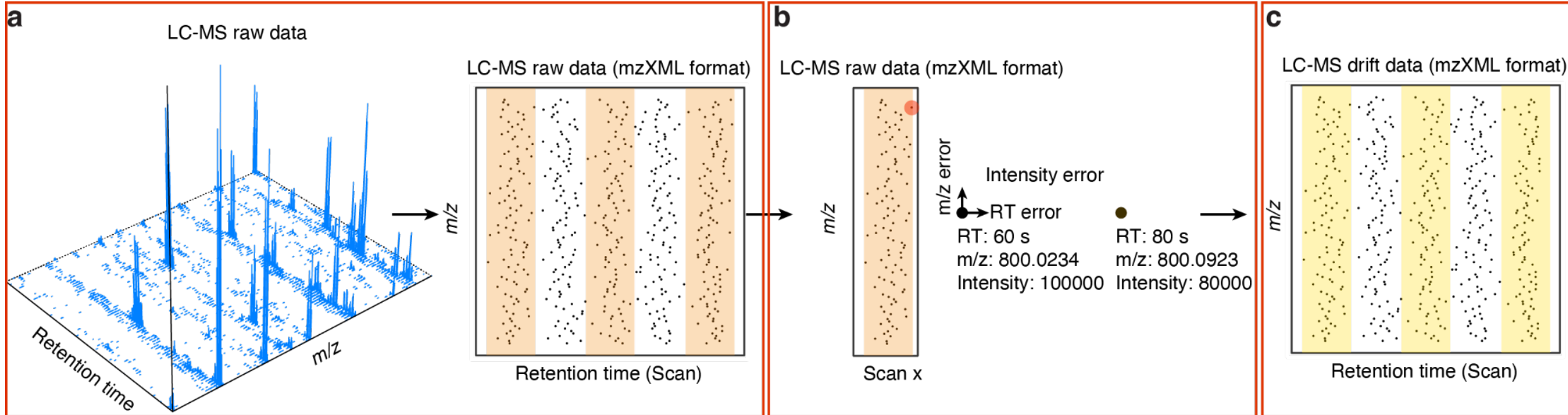
➤ PseudoMS Image predictor



➤ Workflow of pseudo-MS image augmentation



Workflow of pseudo-MS image augmentation



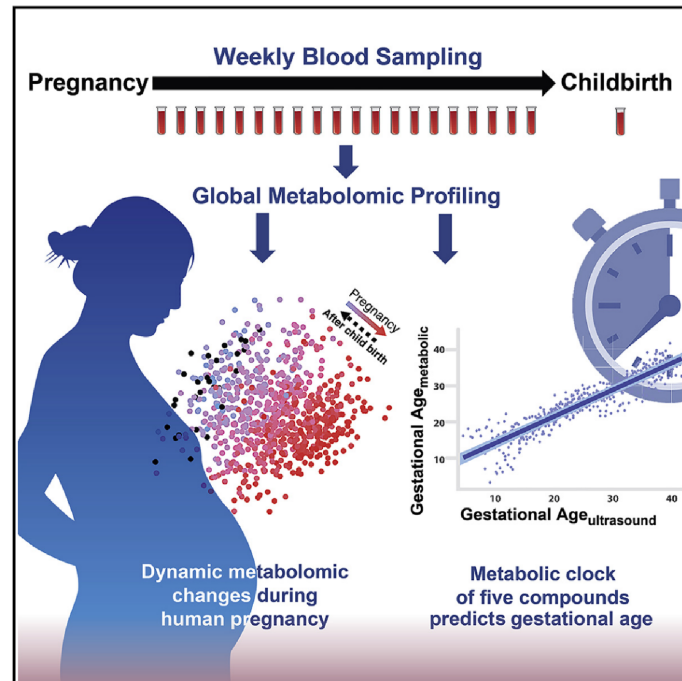
Applications of deepPseudoMSI

Cell

Resource

Metabolic Dynamics and Prediction of Gestational Age and Time to Delivery in Pregnant Women

Graphical Abstract



Authors

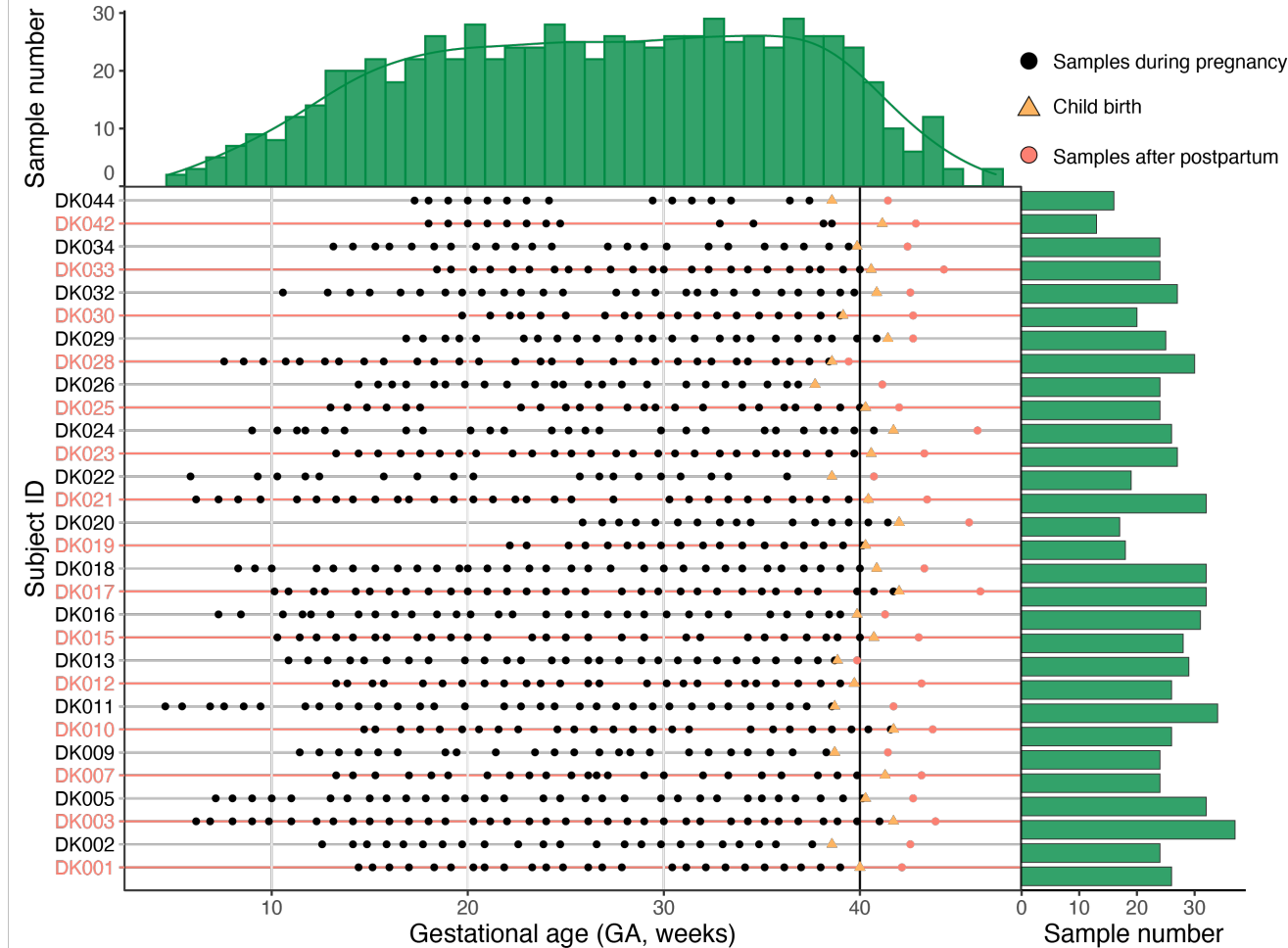
Liang Liang,
Marie-Louise Hee Rasmussen,
Brian Piening, ..., Hanyah Zackriah,
Michael Snyder, Mads Melbye

Correspondence

mpsnyder@stanford.edu (M.S.),
mmelbye@stanford.edu (M.M.)

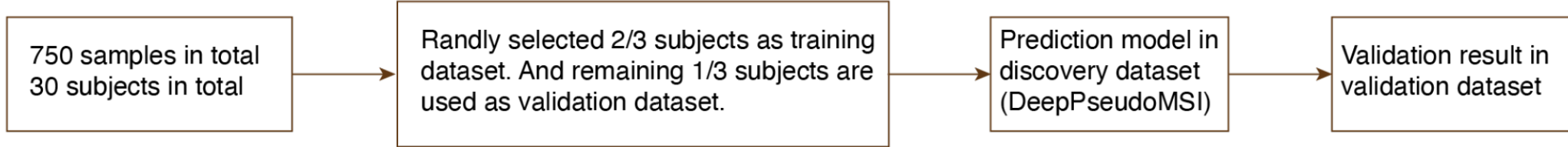
In Brief

Identification of blood metabolites in pregnant women that can accurately predict gestational age and provide insights into pregnancy variations undetected by ultrasound.

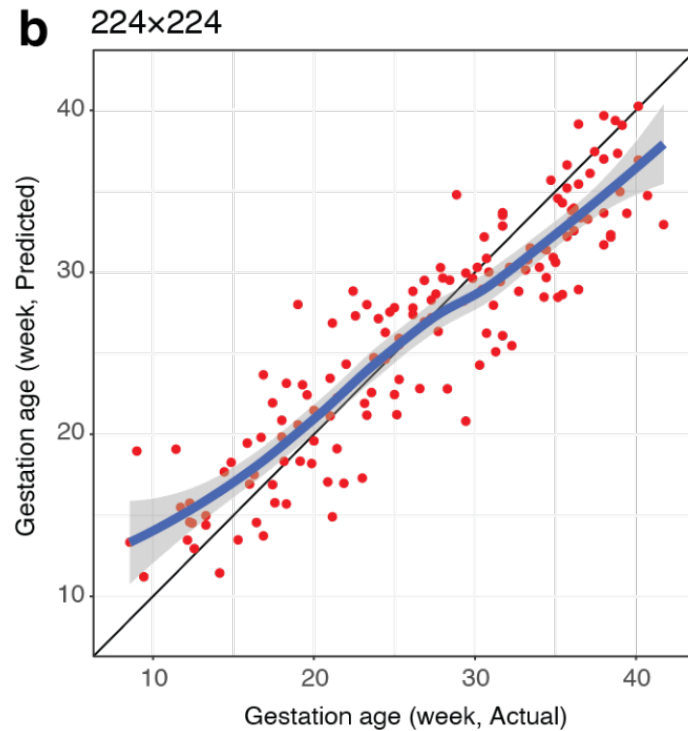


Resolution optimization

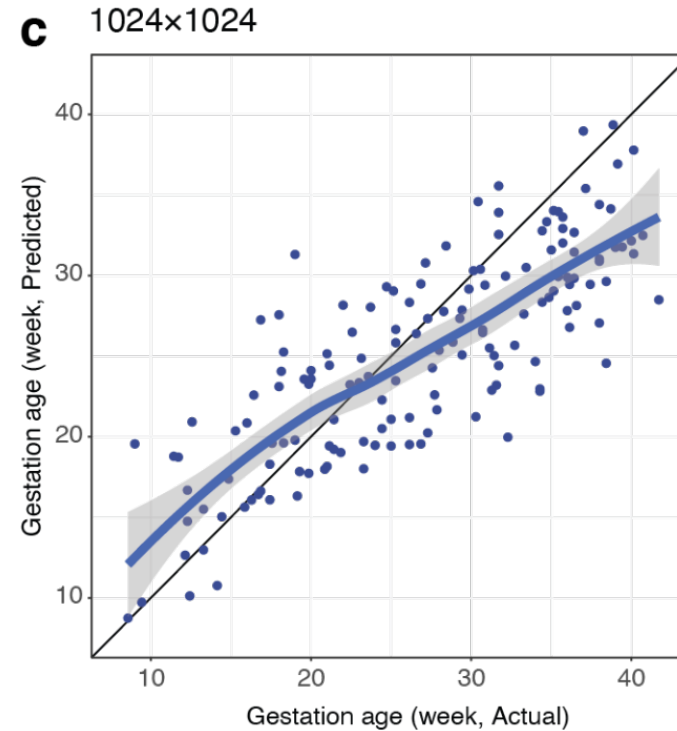
a



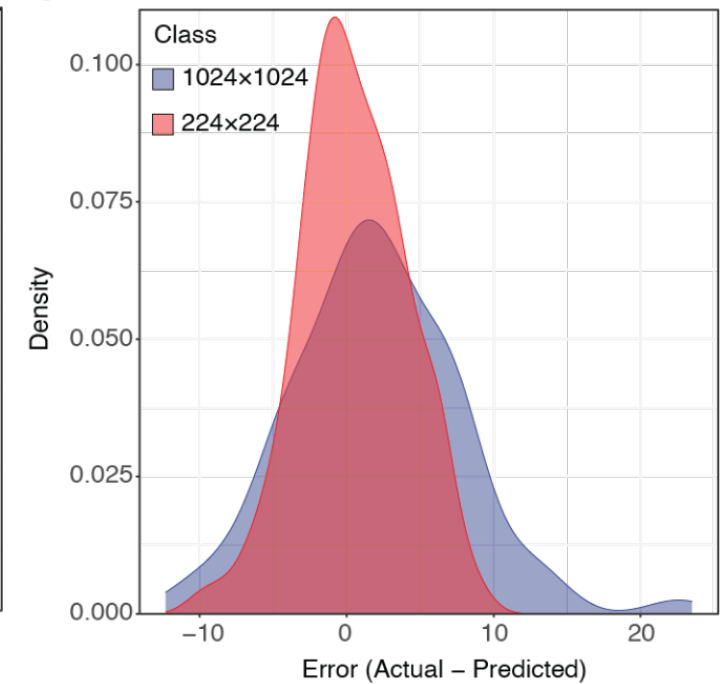
b



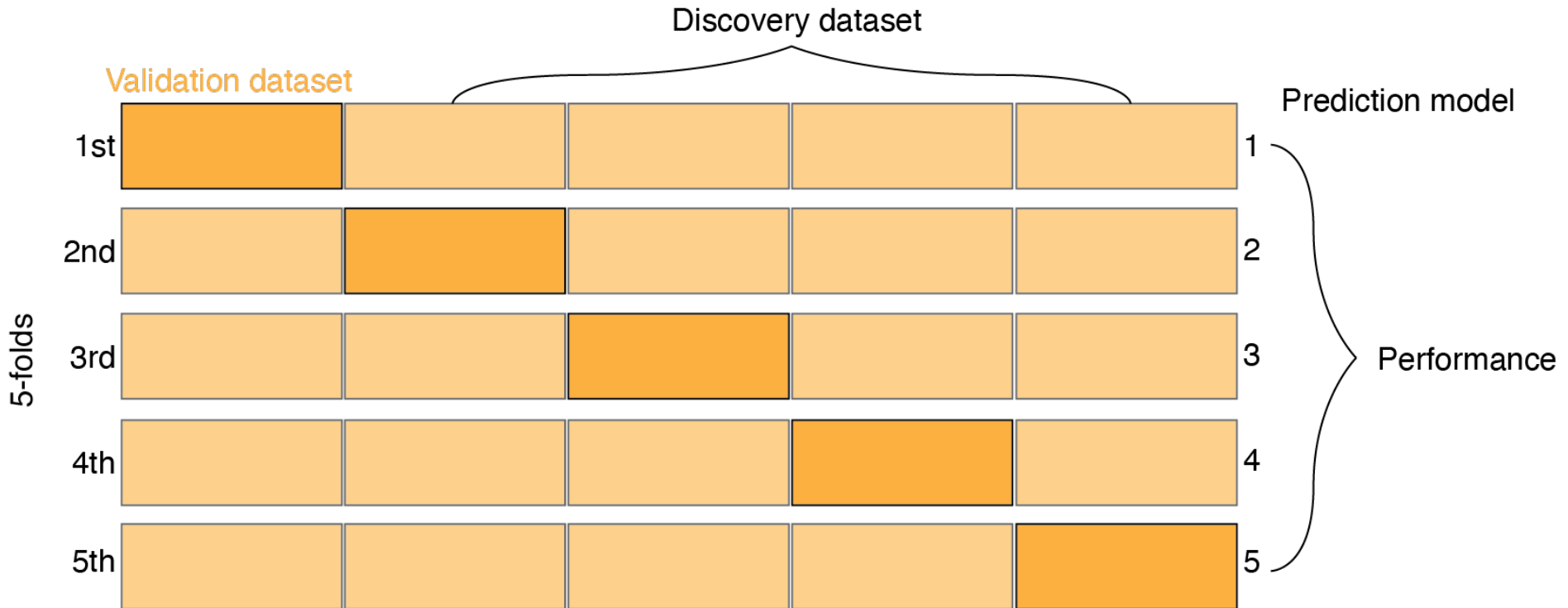
c



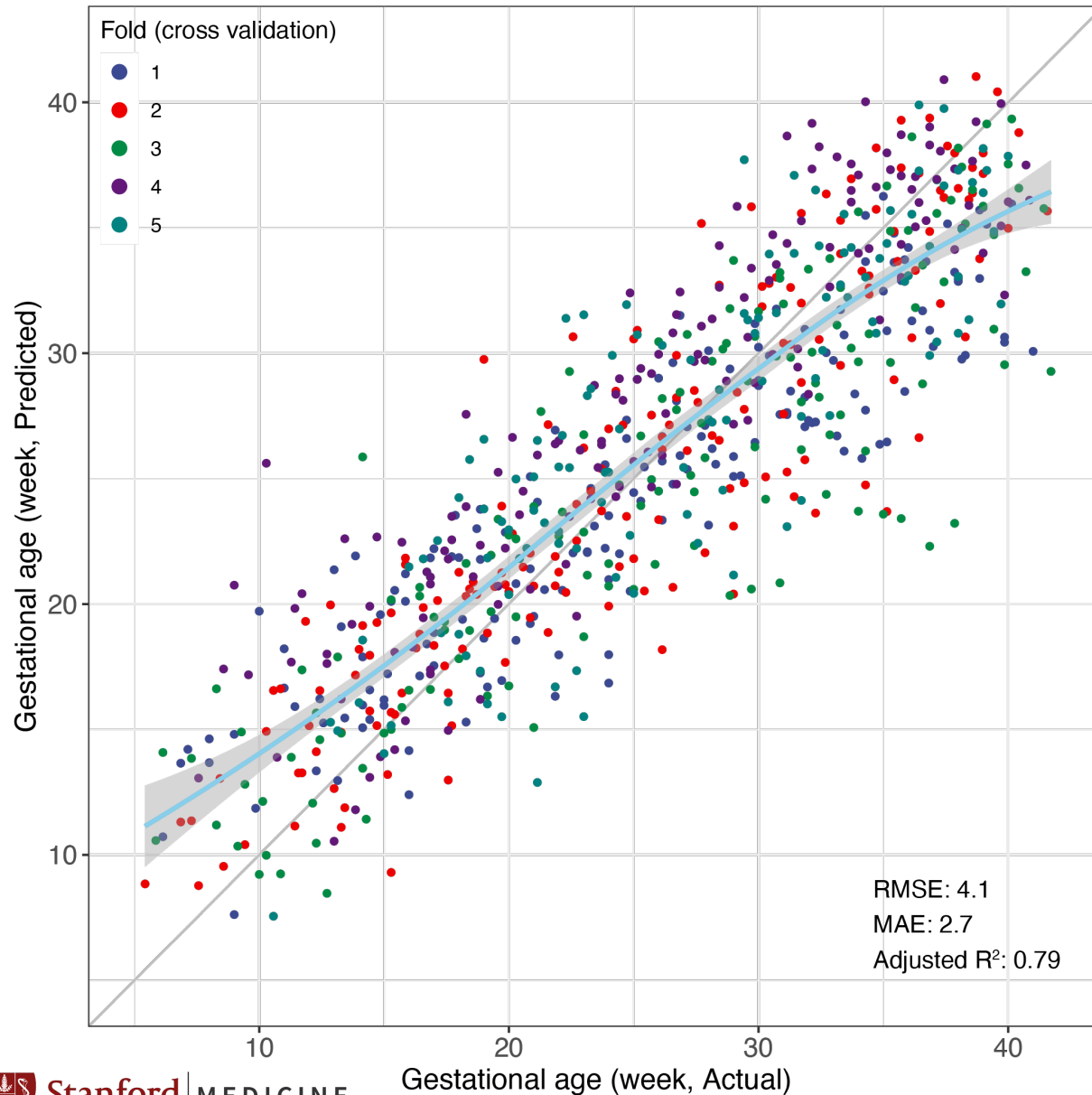
d



DeepPseudoMSI predicts gestation age



DeepPseudoMSI predicts gestation age

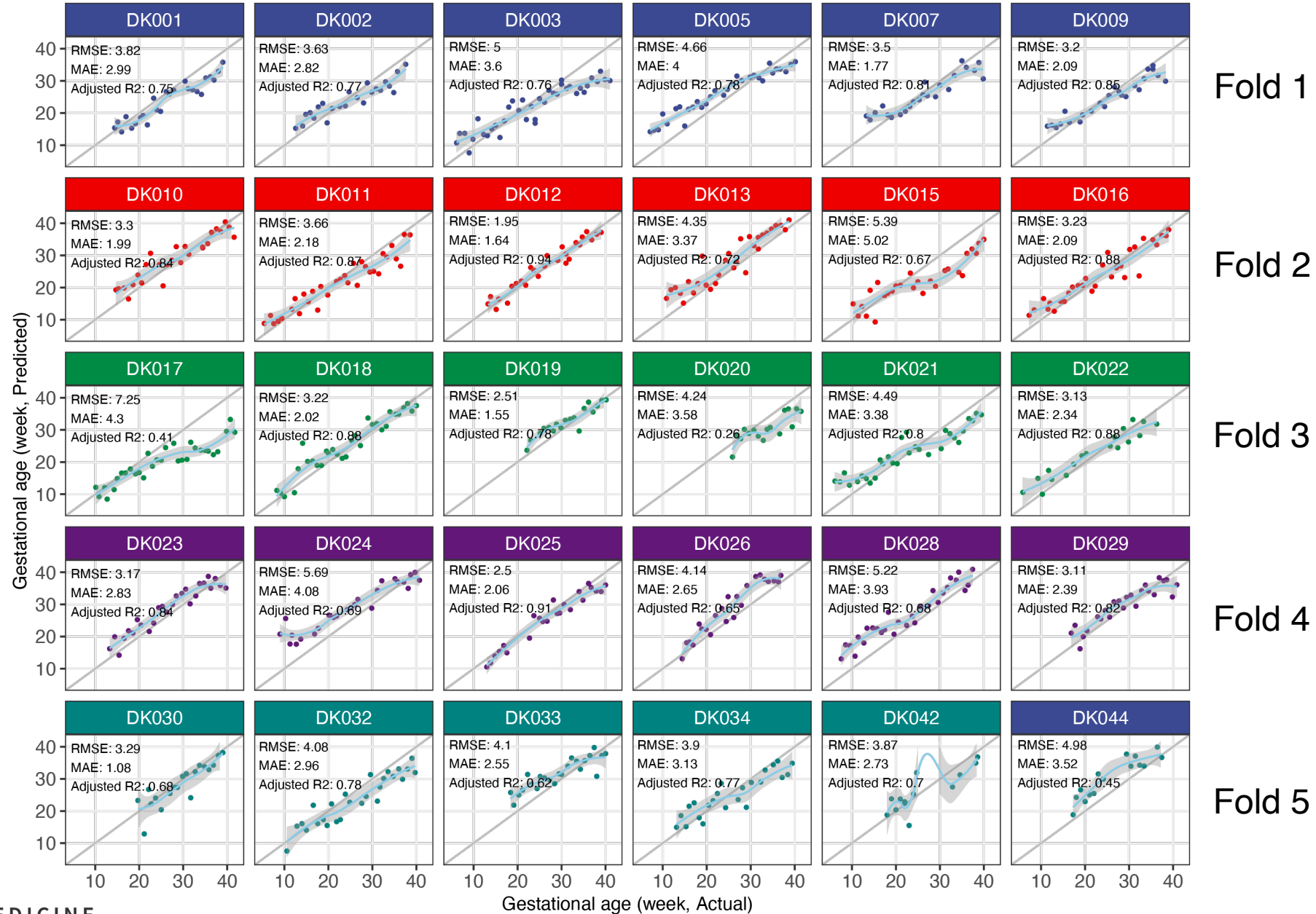


RMSE (Root Mean Squared Error): 4.1 weeks

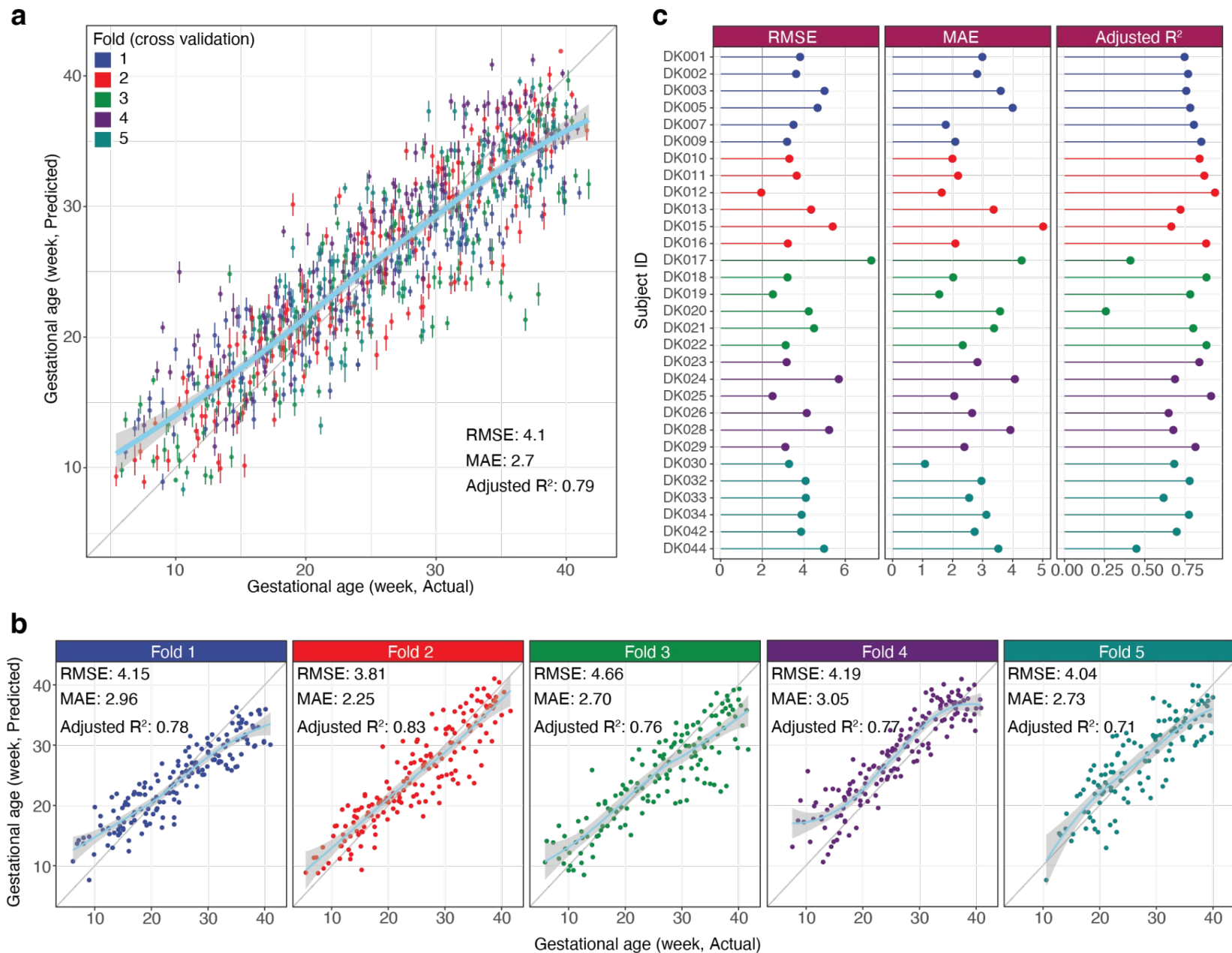
MAE (Mean Average Error): 2.7 weeks

Adjusted R²: 0.79

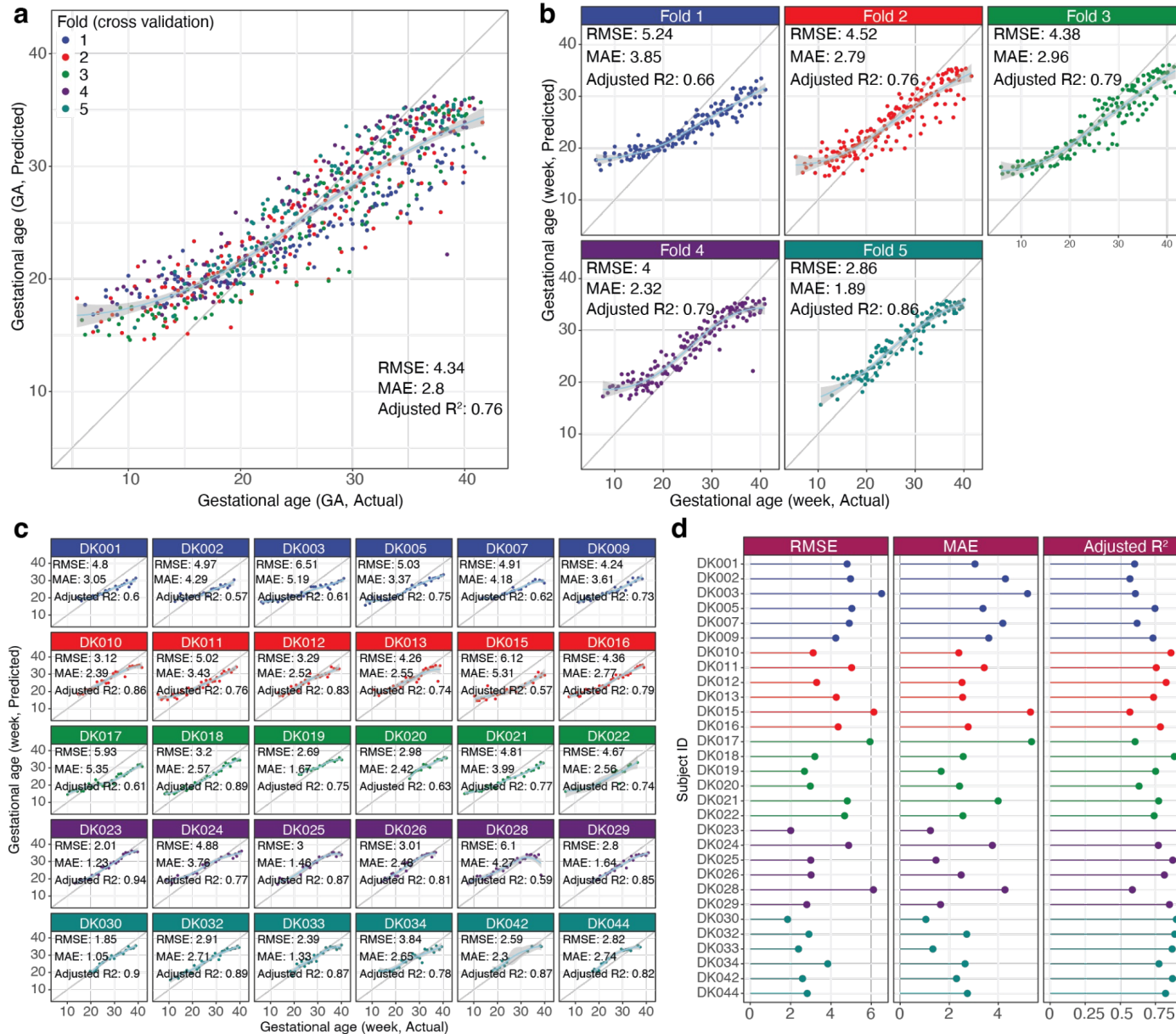
DeepPseudoMSI predicts gestation age (Individual)



DeepPseudoMSI predicts gestation age



> Conventional method



Other public databases

Received: 4 October 2021 | Revised: 9 January 2022 | Accepted: 13 January 2022
DOI: 10.1002/ijc.33943

TUMOR MARKERS AND SIGNATURES



A serum lipidomics study for the identification of specific biomarkers for endometrial polyps to distinguish them from endometrial cancer or hyperplasia

Xingxu Yan¹ | Wen Zhao² | Jinxia Wei¹ | Yaqi Yao¹ | Guijiang Sun³ | Lei Wang⁴ | Wenqing Zhang¹ | Siyu Chen¹ | Wenjie Zhou¹ | Huan Zhao¹ | Xiaomeng Li¹ | Yu Xiao⁵ | Yubo Li¹

scientific reports

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [scientific reports](#) > [articles](#) > [article](#)

Article | [Open Access](#) | [Published: 17 March 2020](#)

Sex Differences in Colon Cancer Metabolism Reveal A Novel Subphenotype

[Yuping Cai](#), [Nicholas J. W. Ratnayake](#), [Qian Zhang](#), [Varvara Mironova](#), [Alvaro Santos-Neto](#), [Kuo-Shun Hsu](#), [Zahra Ratnayake](#), [Justin R. Cross](#), [Yawei Zhang](#), [Philip B. Paty](#), [Sajid A. Khan](#) & [Caroline H. Johnson](#)

[Scientific Reports](#) **10**, Article number: 4905 (2020) | [Cite this article](#)

4498 Accesses | 19 Citations | 40 Altmetric | [Metrics](#)

A MTBLS3444

1	False positive 8	True positive 701
0	True negative 1324	False negative 49
	0	1

$$\text{Accuracy} = \frac{701 + 1324}{701 + 1324 + 8 + 49} = 97.3\%$$

$$\text{Specificity} = \frac{1324}{1324 + 49} = 96.4\%$$

$$\text{Sensitivity} = \frac{701}{701 + 8} = 98.9\%$$

B MTBLS1129

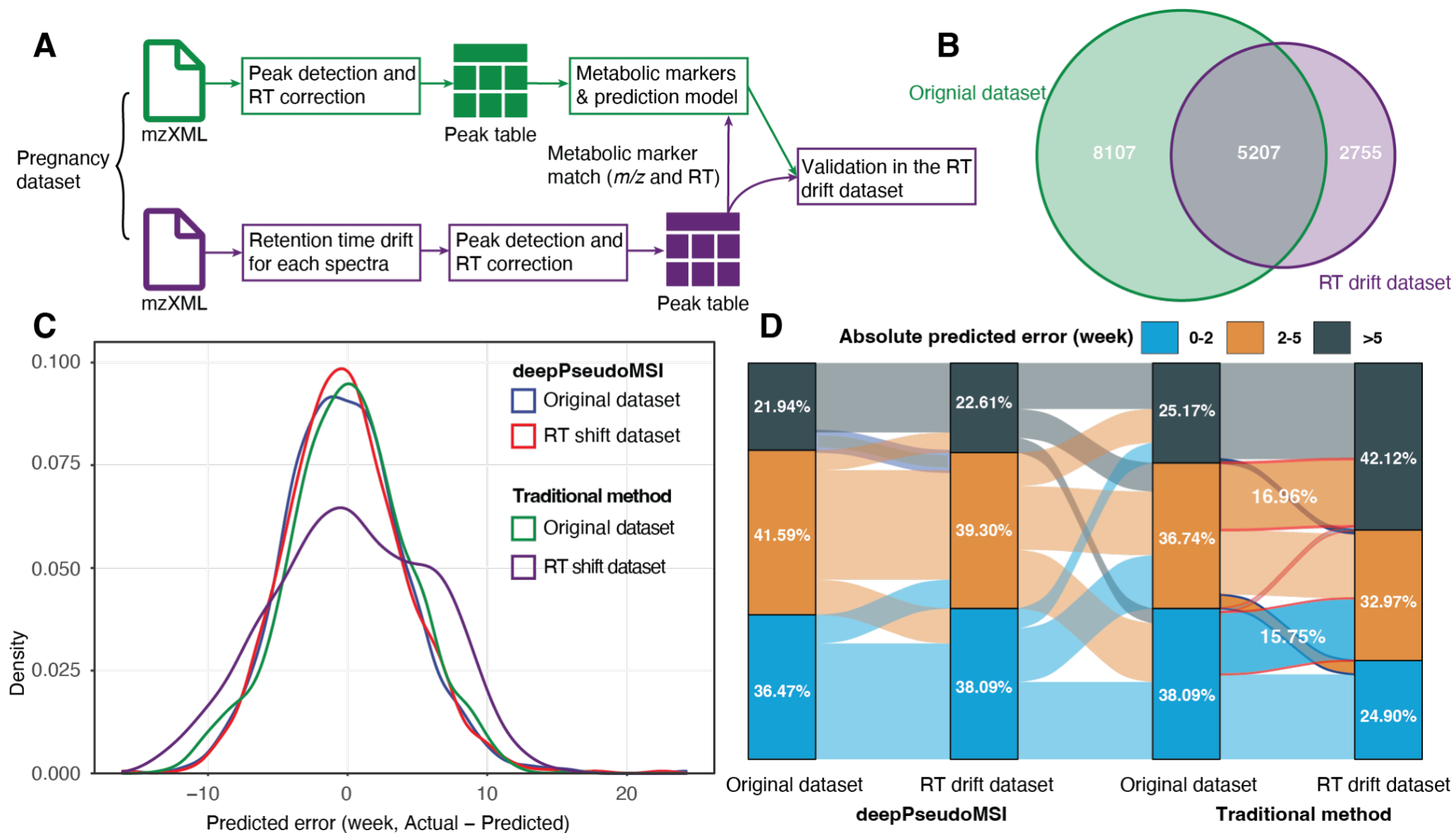
1	False positive 115	True positive 1122
0	True negative 119	False negative 58
	0	1

$$\text{Accuracy} = \frac{1122 + 119}{1122 + 119 + 115 + 58} = 87.8\%$$

$$\text{Specificity} = \frac{119}{119 + 58} = 67.2\%$$

$$\text{Sensitivity} = \frac{1122}{1122 + 115} = 90.7\%$$

> Evaluation of deepPseudoMSI





DeepPseudoMSI

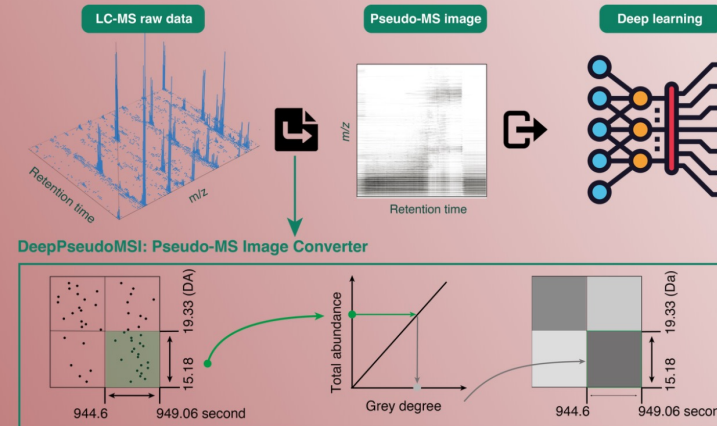
[About](#) [Case study](#) [Tutorial](#) [Team](#) [Tools](#) [Contact](#)

DeepPseudoMSI project

The deepPseudoMSI project is the first method that convert LC-MS raw data to “images” and then process them using deep learning method for diagnosis.

[Source code </>](#)

[Ask a question >](#)



DeepPseudoMSI

Deep learning-based pseudo-mass spectrometry imaging analysis for precision medicine

The **deepPseudoMSI** project is the first method that convert LC-MS raw data to “images” and then process them using deep learning method for diagnosis. It contains two parts.

1. Pseudo-MS image converter

One LC-MS raw data usually contains millions of data points, so we need to divide it into different pixels (or grids) based on the

Future works

pseudomsir · pseudomsir

pseudomsir.deepppseudomsir.org

Xiaotao Shen Boo... SelectorGadget LMGTFY stanford PubChem Identifie... my bookmarks Conservice SV SbmlViewer faculty

PSEUDOMSIR 0.99.3 Get started News Reference

Search...

GitHub Twitter

pseudomsir

pseudomsir is a part of deepPseudoMSI project.



About

pseudomsir is a R package which is used for deepPseudoMSI project.

Installation

You can install pseudomsir from GitLab

```
if(!require(remotes)){
  install.packages("remotes")
}
remotes::install_gitlab("jaspershen/pseudomsir")
```

or GitHub

```
remotes::install_github("deepPseudoMSI-project/pseudomsir")
```

Usage

Please see the Get started page to get the instruction of pseudomsir.

Need help?

If you have any quesitions about pseudomsir, please don't hesitate to email me (shenxt@stanford.edu).

shenzutao1990

shenxt@stanford.edu

Twitter

M339, Alway building, Cooper Lane, Palo Alto, CA 94304

Links

Browse source code

Find me

License

Full license

MIT + file LICENSE

Citation

Citing pseudomsir

Developers

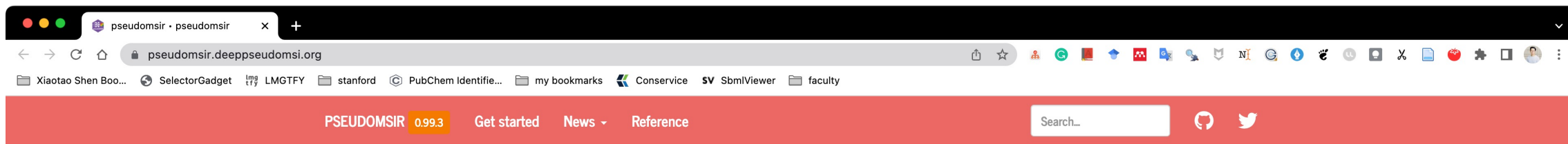
Xiaotao Shen

Author, maintainer

Dev status

CRAN	not published
code size	repo not found
dependencies	unknown
lifecycle	experimental

Future works



pseudomsir

pseudomsir is a part of [deepPseudoMSI project](#).

About

pseudomsir is a R package which is used for [deepPseudoMSI project](#).

Installation

You can install pseudomsir from [GitLab](#)

```
if(!require(remotes)){
  install.packages("remotes")
}
remotes::install_gitlab("jaspershen/pseudomsir")
```

or [GitHub](#)

```
remotes::install_github("deepPseudoMSI-project/pseudomsir")
```

Usage

Please see the [Get started](#) page to get the instruction of pseudomsir.

Need help?

If you have any questions about pseudomsir, please don't hesitate to email me (shenxt@stanford.edu).

[shenzutao1990](#)

shenxt@stanford.edu

[Twitter](#)

[M339, Alway building, Cooper Lane, Palo Alto, CA 94304](#)



Links

[Browse source code](#)

[Find me](#)

License

[Full license](#)

[MIT](#) + file [LICENSE](#)

Citation

[Citing pseudomsir](#)

Developers

[Xiaotao Shen](#)

Author, maintainer 

Dev status

CRAN [not published](#)

code size [repo not found](#)

dependencies [unknown](#)

lifecycle [experimental](#)

Future works

- ❖ **How to combine datasets of different chromatography and Electrospray ionization (ESI) modes to increase the prediction accuracy.**
- ❖ **A large-scale dataset for cancer diagnosis.**

> SUMMARY

- </> **DeepPseudoMSI is the first systematic study that converts the LC–MS-based untargeted metabolomics data to pseudo-MS images and then takes advantage of the power of deep learning in image processing for precision medicine.**
- </> **We also demonstrate that the deepPseudoMSI can overcome the limitations of the traditional method for LC-MS data in precision medicine.**

> ACKNOWLEDGEMENTS



Prof. Wei Shao
(University of Florida)



Dr. Chuchu Wang
(Stanford University)



Prof. Mirabela Rus
(Stanford University)



Prof. Michael Snyder
(Stanford University)

Stanford University

Liang Liang PhD

Songjie Chen PhD

Sai Zhang PhD



Stanford
MEDICINE



Thanks for your attention!

Q&A

Xiaotao Shen PhD

Stanford University School of Medicine
Department of Genetics



shenxt@stanford.edu



[@xiaotaoshen1990](https://twitter.com/xiaotaoshen1990)



shenxt.info



github.com/jaspershen