# Machine Learning for Better Medicine

Yuan Luo, PhD, FAMIA

Chief AI Officer

Northwestern University Clinical and Translational Sciences Institute

Institute for Augmented Intelligence in Medicine

Associate Professor

Department of Preventive Medicine

Northwestern University

yuan.luo@northwestern.edu

@yuanhypnosluo

NORTHWESTERN UNIVERSITY FEINBERG SCHOOL OF MEDICINE

10/9/2022

Northwestern University

# Lack of evidence in medical practice

- Only 11 percent of sampled American College of Cardiology and American Heart Association guidelines used evidence from RCTs or meta-analyses

JAMA Network

JAMA

Original Contribution

February 25, 2009

## Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines

Pierluigi Tricoci, MD, MHS, PhD; Joseph M. Allen, MA; Judith M. Kramer, MD, MS; et al

Author Affiliations | Article Information

*JAMA.* 2009;301(8):831-841. doi:10.1001/jama.2009.205

# Lack of evidence in medical practice

- Of more than 4000 recommendations for infectious diseases, 14 percent were based on RCTs

JAMA Network

JAMA Internal Medicine

**Original Investigation**

January 10, 2011

## Analysis of Overall Level of Evidence Behind Infectious Diseases Society of America Practice Guidelines

Dong Heun Lee, MD; Ole Vielemeyer, MD

» Author Affiliations | Article Information

*Arch Intern Med.* 2011;171(1):18-22. doi:10.1001/archinternmed.2010.482

# Lack of evidence in medical practice

- 88 publications of CPGs involving 3119 recommendations
- 9% based on PRCTs and percentage decreased over time

*Intensive Care Med* (2018) 44:1189–1191
https://doi.org/10.1007/s00134-018-5142-8

LETTER

CrossMark

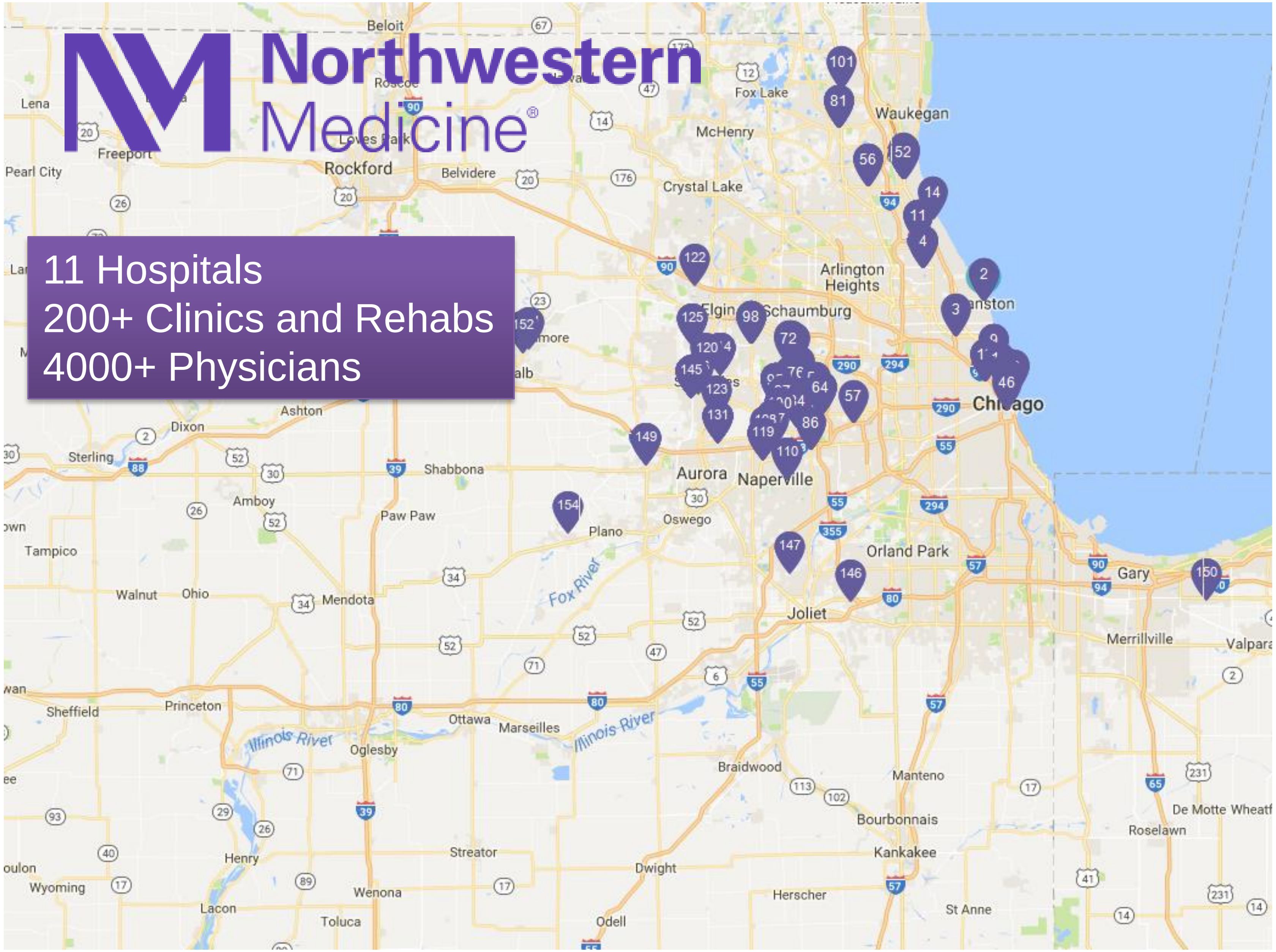
## Scientific evidence underlying the recommendations of critical care clinical practice guidelines: a lack of high level evidence

Zhongheng Zhang*, Yucai Hong and Ning Liu

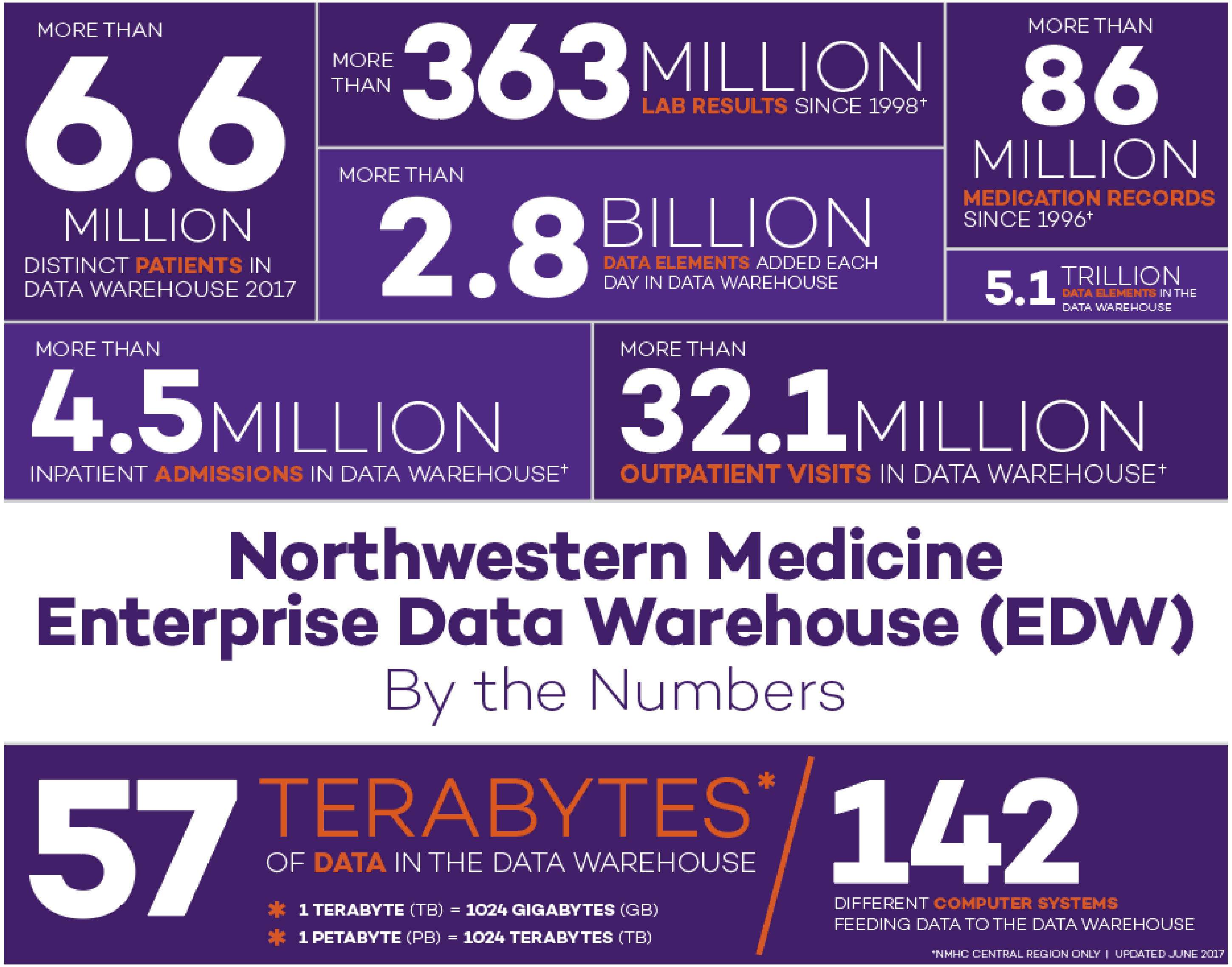
# Evidence-based medicine or expert opinion

Northwestern Medicine®

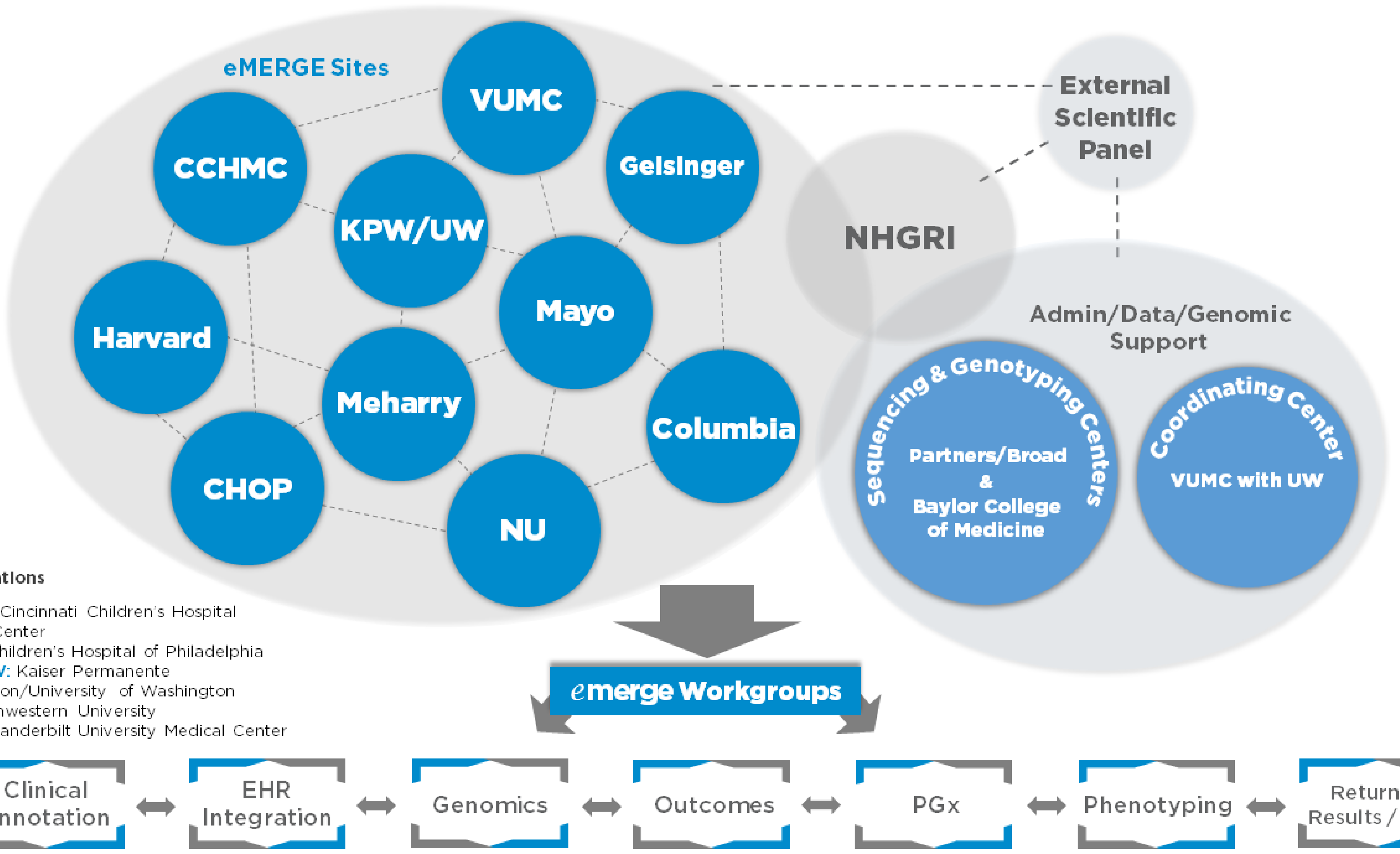
11 Hospitals
200+ Clinics and Rehabs
4000+ Physicians



# Northwestern Medicine Enterprise Data Warehouse



# Electronic Medical Records and Genomics (eMERGE) Network




abbreviations

CCHMC: Cincinnati Children's Hospital Medical Center
CHOP: Children's Hospital of Philadelphia
KPW/UW: Kaiser Permanente Washington/University of Washington
NU: Northwestern University
VUMC: Vanderbilt University Medical Center

# Machine learning to integrate multi-modal healthcare data

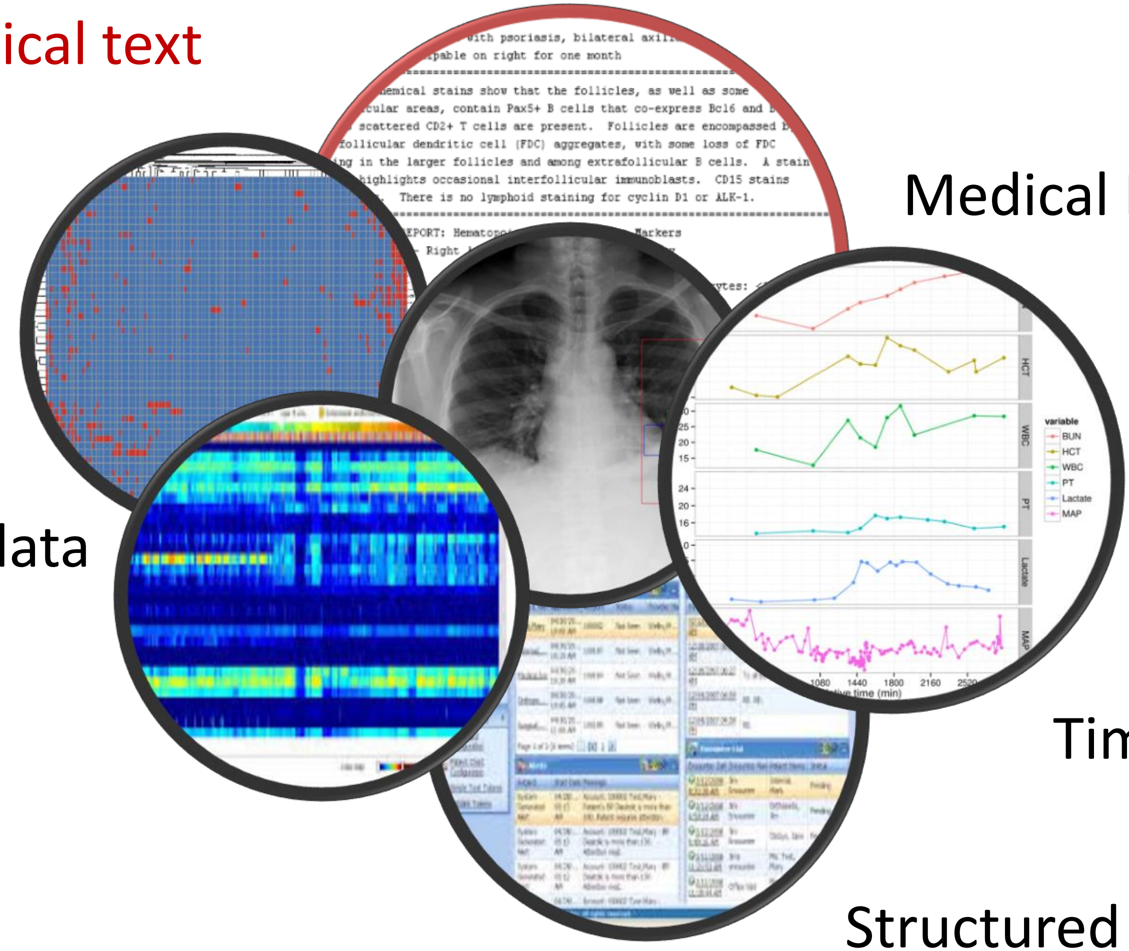
Biomedical text

Medical Imaging

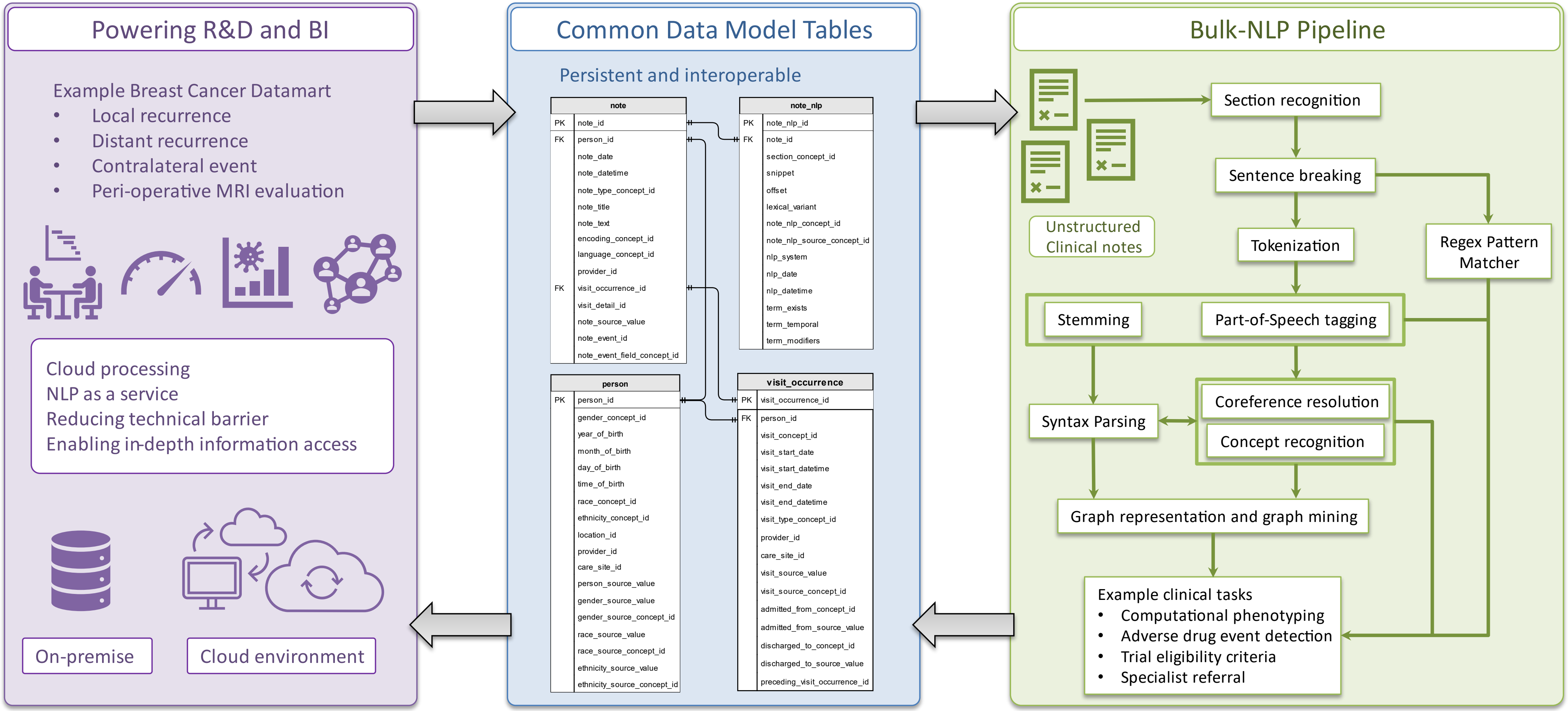
Omics data

Time series

Structured data



# Bulk NLP to power R&D and BI




NUCATS Clinical and Translational Sciences Institute · Northwestern · Northwestern Medicine · Ann & Robert H. Lurie Children's Hospital of Chicago · Northwestern Medicine Feinberg School of Medicine · Shirley Ryan Abilitylab

# Computational phenotyping to ascertain breast cancer recurrence

## After 48 Years, NCI Aims to Track Breast Cancer Recurrences

**Change to SEER Eventually Planned**

Nick Mulcahy
January 08, 2021

1 Read Comment ADD TO EMAIL ALERTS


The problem we tried to solve: Cancer Recurrence data are hard to come by

# Computational phenotyping to ascertain breast cancer recurrence

- Progress notes and pathology report contain rich information on BC local, distant recurrences and contralateral events


Springer Link

Research Article | Published: 08 April 2019

**Identifying Breast Cancer Distant Recurrences from Electronic Health Records Using Machine Learning**

Zexian Zeng, Liang Yao, Ankita Roy, Xiaoyu Li, Sasa Espino, Susan E Clare, Seema A Khan & Yuan Luo

*Journal of Healthcare Informatics Research* **3**, 283–299(2019) | Cite this article

AMIA Annual Symposium Proceedings Archive

AMIA Annu Symp Proc. 2017; 2017: 1885–1892. PMCID: PMC5977664

Published online 2018 Apr 16. PMID: 29854260

**Contralateral Breast Cancer Event Detection Using Nature Language Processing**

Zexian Zeng,[1] Xiaoyu Li,[2] Sasa Espino,[3] Ankita Roy,[3] Kristen Kitsch,[3] Susan Clare,[3] Seema Khan,[3] and Yuan Luo[1,*]

Zeng *et al. BMC Bioinformatics* 2018, **19**(Suppl 17):498
https://doi.org/10.1186/s12859-018-2466-x

BMC Bioinformatics

RESEARCH Open Access

**Using natural language processing and machine learning to identify breast cancer local recurrence**

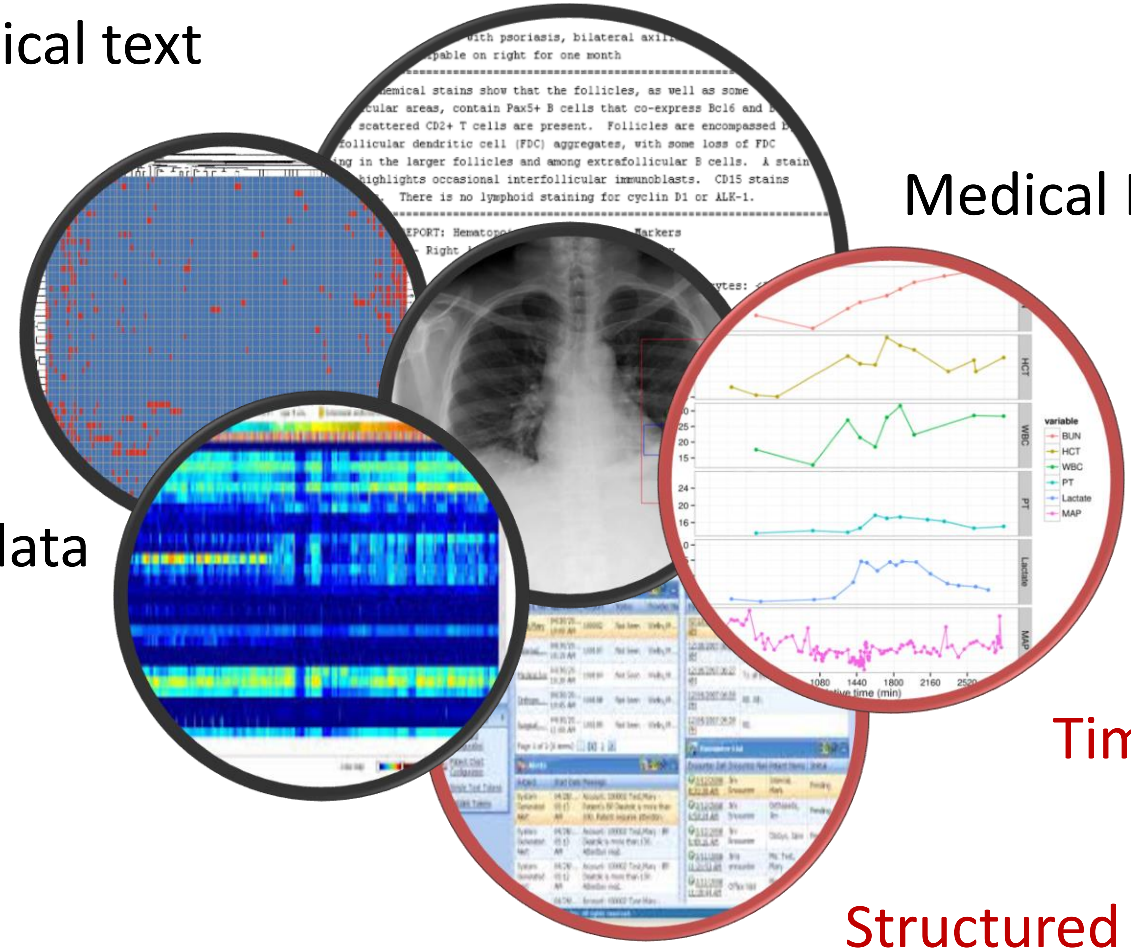
Zexian Zeng[1], Sasa Espino[2], Ankita Roy[2], Xiaoyu Li[3], Seema A. Khan[2], Susan E. Clare[2], Xia Jiang[4], Richard Neapolitan[1] and Yuan Luo[1*]

# Machine learning to integrate multi-modal healthcare data

Biomedical text

Medical Imaging

Omics data




Time series

Structured data

# Simulation of Ventilator Allocation in Critically Ill Patients with COVID-19

American Journal of Respiratory and Critical Care Medicine

Home > American Journal of Respiratory and Critical Care Medicine > List of Issues > Just Accepted

Article Tools

## Simulation of Ventilator Allocation in Critically Ill Patients with COVID-19

Previous Article Next Article

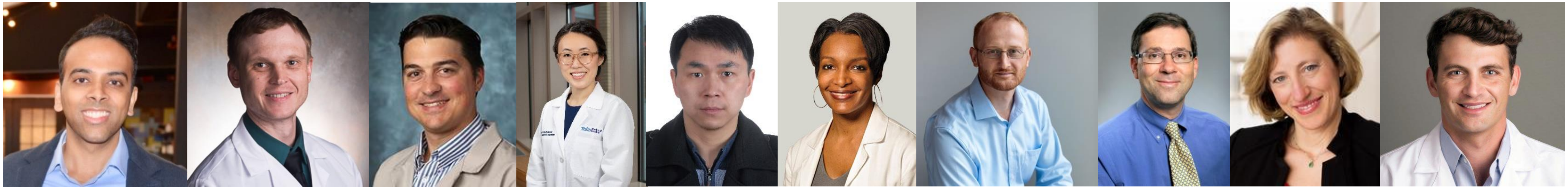
**Sivasubramanium V Bhavani** * **Yuan Luo** * **William D. Miller** , **Lazaro N. Sanchez-Pinto** ; **Xuan Han** , **Chengsheng Mao** , **Burhaneddin Sandikci** , **Monica E. Peek** , **Craig M. Coopersmith** , **Kelly N. Michelson** , and **William F Parker**

+ Author Information



https://doi.org/10.1164/rccm.202106-1453LE **PubMed:** 34499587

Received: June 17, 2021 Accepted: September 07, 2021

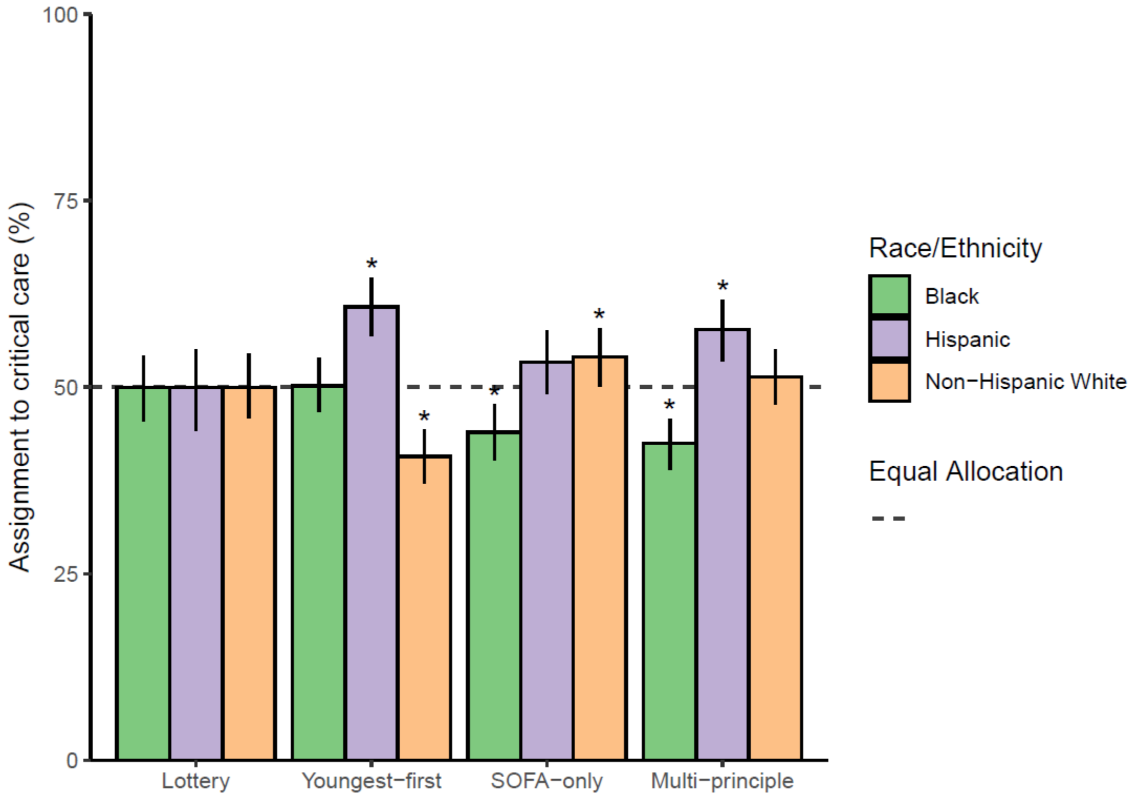


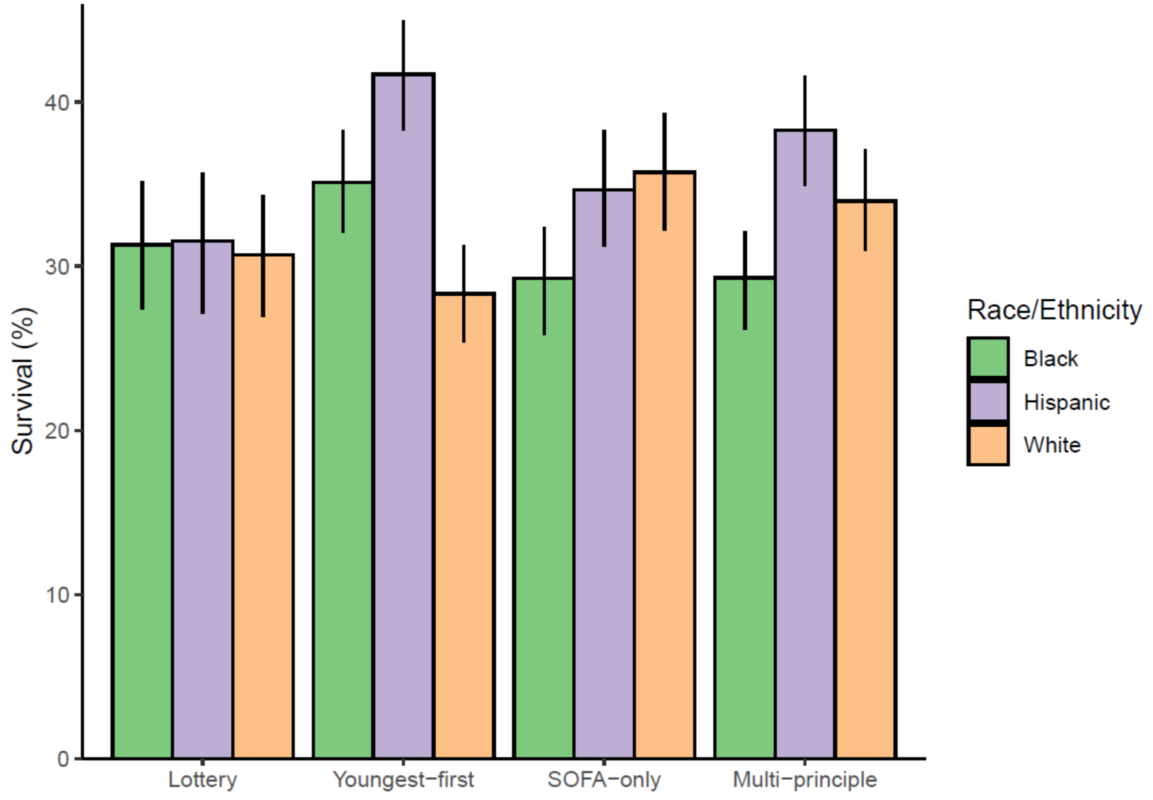
| | Total | Tertiary care hospital A | Tertiary care hospital B | Community hospital network |
|---|---|---|---|---|
| N | 2,363 | 643 | 687 | 1,033 |
| Age, years | 64 (52-75) | 63 (50-73) | 64 (51-74) | 65 (53-76) |
| Sex, male | 1,398 (59.2) | 364 (56.6) | 407 (59.2) | 627 (60.7) |
| Race/Ethnicity | | | | |
| Non-Hispanic Black | 760 (32.2) | 489 (76.0) | 204 (29.7) | 67 (6.5) |
| Non-Hispanic White | 828 (35.0) | 55 (8.6) | 209 (30.4) | 564 (54.6) |
| Hispanic | 582 (24.6) | 65 (10.1) | 202 (29.4) | 315 (30.5) |
| Other | 193 (8.2) | 34 (5.3) | 72 (10.5) | 87 (8.4) |
| Insurance | | | | |
| Medicare | 1,227 (51.9) | 341 (53) | 342 (49.8) | 544 (52.7) |
| Medicaid | 480 (20.3) | 188 (29.2) | 121 (17.6) | 171 (16.6) |
| Private | 613 (25.9) | 112 (17.4) | 211 (30.7) | 290 (28.1) |
| Unknown | 43 (1.8) | 2 (0.3) | 13 (1.9) | 28 (2.7) |
| SOFA score | 5 (3-8) | 6 (4-8) | 5 (3-8) | 5 (3-7) |
| Chronic Disease | | | | |
| Non-major (Score<16) | 1,730 (73.2) | 331 (51.5) | 480 (69.9) | 919 (89) |
| Major (16 ≥ score < 28) | 387 (16.4) | 187 (29.1) | 116 (16.9) | 84 (8.1) |
| Severe (score ≥ 28) | 246 (10.4) | 125 (19.4) | 91 (13.2) | 30 (2.9) |
| Critical care resources | | | | |
| Hemodialysis | 388 (16.4) | 171 (26.6) | 112 (16.3) | 105 (10.2) |
| Vasopressor use | 843 (35.7) | 259 (40.3) | 294 (42.8) | 290 (28.1) |
| High flow nasal cannula | 1,313 (55.6) | 293 (45.6) | 415 (60.4) | 605 (58.6) |
| Non-invasive ventilation | 621 (26.3) | 89 (13.8) | 186 (27.1) | 346 (33.5) |
| Mechanical ventilation | 998 (42.2) | 247 (38.4) | 353 (51.4) | 398 (38.5) |
| Outcomes | | | | |
| LOS, d | 11.7 (6.0-21.1) | 10.8 (6.3-18.4) | 15.1 (7.4-27.3) | 10.2 (4.9-18.7) |
| Mortality rate | 475 (20.1) | 162 (25.2) | 111 (16.2) | 202 (19.6) |

# Ventilator allocation protocols

| Protocol | Ethical Principles | Rules | Countries and US States with Similar Protocols | Lives saved (%) |
|---|---|---|---|---|
| Lottery | Treating people equally | Random assignment | None | 31.1 (29.7-32.6) |
| Youngest-first | Save life-years<br><br>Prudential Lifespan Equity | Rank by age | Italy | 34.7 (33.5-35.8) |
| SOFA-only* | Save lives | Three SOFA tiers:<br>Red: ≤ 7<br>Yellow: 8-11<br>Blue: >11<br><br>Lottery tiebreaker | Arizona, California, Indiana, Iowa, Kansas, Louisiana, Michigan, Nevada, New Mexico,<br>New York,<br>South Carolina, Tennessee, Vermont, Washington | 32.9 (31.6-34.1) |
| Multi-principle† | Save lives<br><br>Save life-years | SOFA category points:<br>1. ≤8<br>2. 9-11<br>3. 12-14<br>4. >14<br><br>Chronic conditions:<br>+ 3 points if “severe”<br><br>Age tiebreaker‡ | Maryland, Massachusetts, Pennsylvania, Oklahoma | 33.5 (32.3-34.8) |

# Allocation and survival by race



## Unintended Disparities in Scarce Resource Allocation in Critically Ill Patients with COVID-19

- Black patients had higher SOFA scores and higher prevalence of comorbidities
- Leading to lower priority tiers and significantly less allocation of ventilators
- In turn leading to significantly lower survival
- White patients have increased vents in the SOFA-only protocol, but did not translate into higher survival

# International shared-task challenge for clinical missing data imputation

All admissions where each of the 13 tests was recorded at least once
- •47,862 unique admissions
- •696,673 time points
- •7,601,216 test results

Exclude time points with more than half of the 13 laboratory tests missing

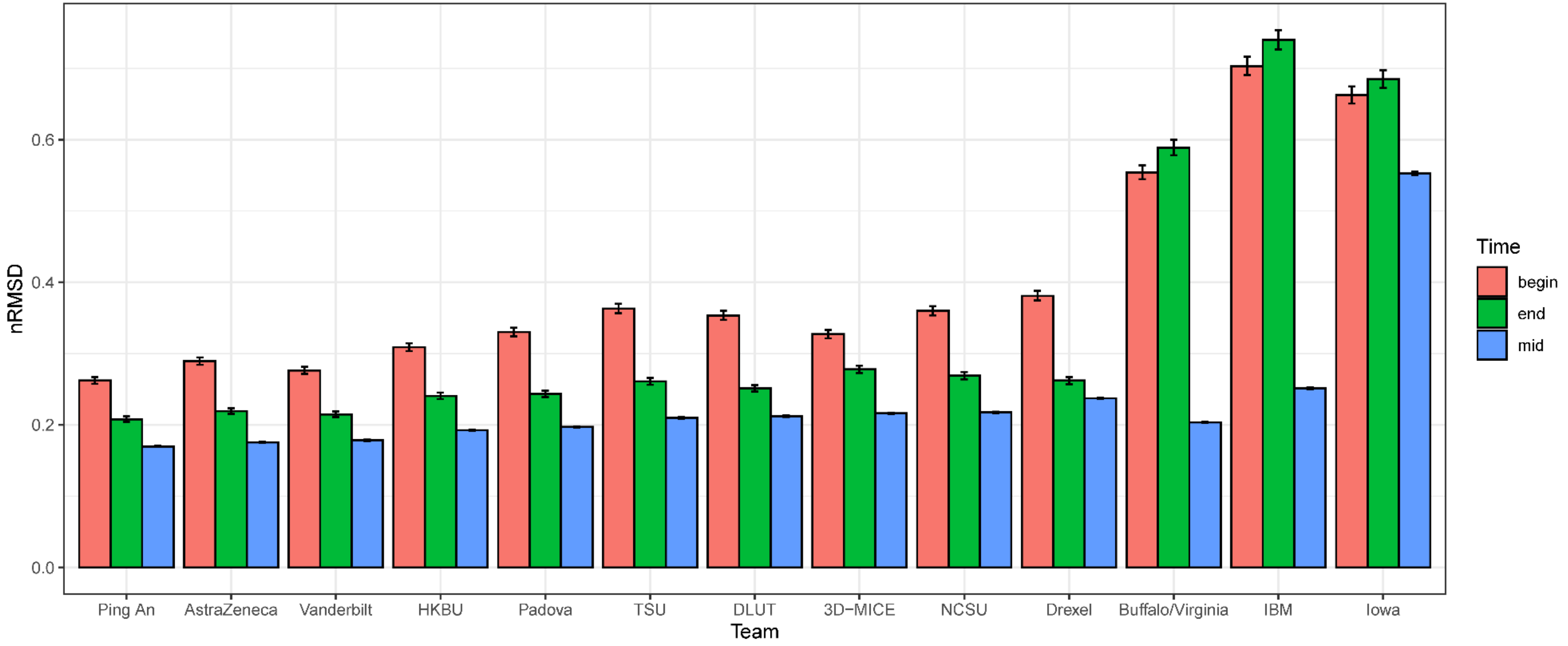
- •47,862 unique admissions
- •594,187 time points
- •7,143,122 test results

Exclude admissions without at least 10 remaining time points for each variable

- •16,537 unique admissions
- •396,668 time points
- •4,774,228 test results

Exclude admissions with a constant value across time points for a test

- •16,534 unique admissions
- •396,631 time points
- •4,773,769 test results



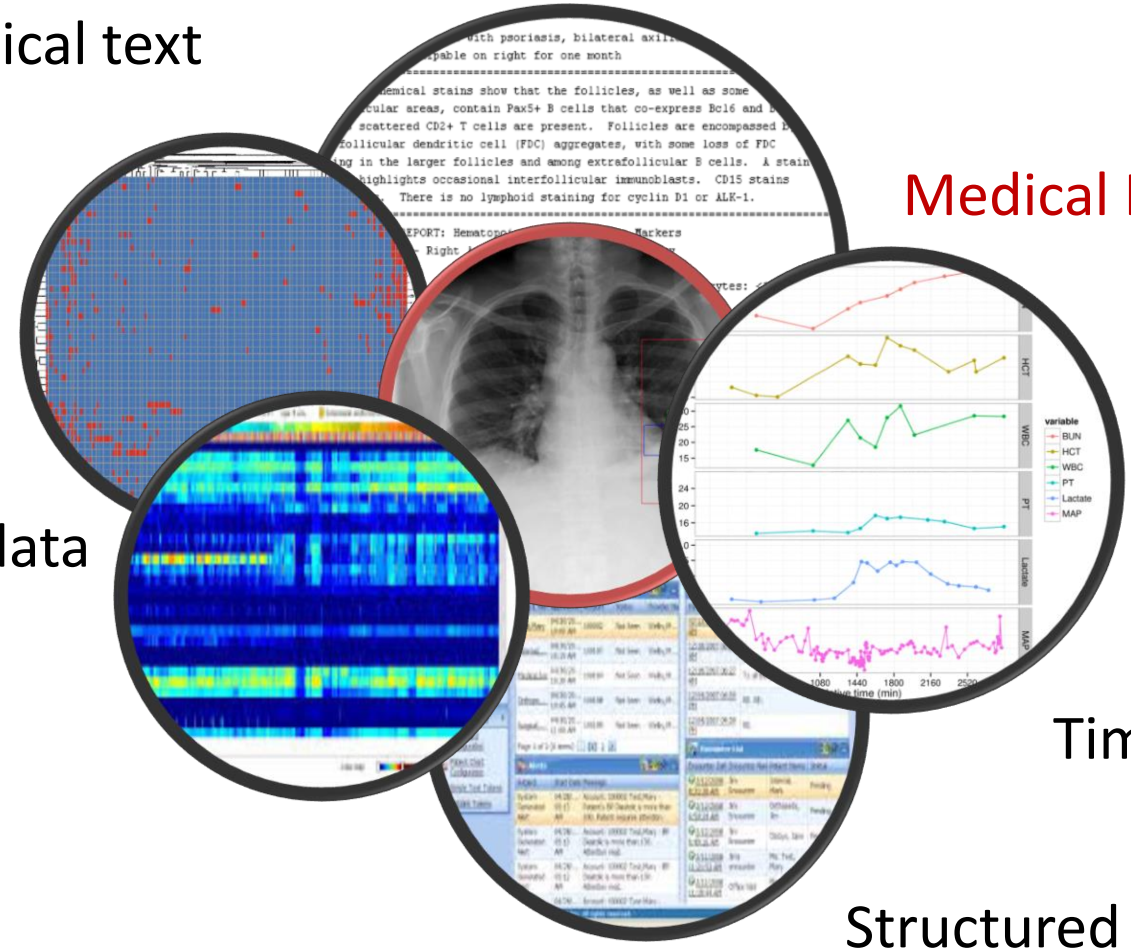

Luo Y, Evaluating the state-of-the-art in missing data imputation for clinical data, *Briefings in Bioinformatics* 2021 [https://doi.org/10.1093/bib/bbab489](https://doi.org/10.1093/bib/bbab489)

# Machine learning to integrate multi-modal healthcare data

Biomedical text

Medical Imaging

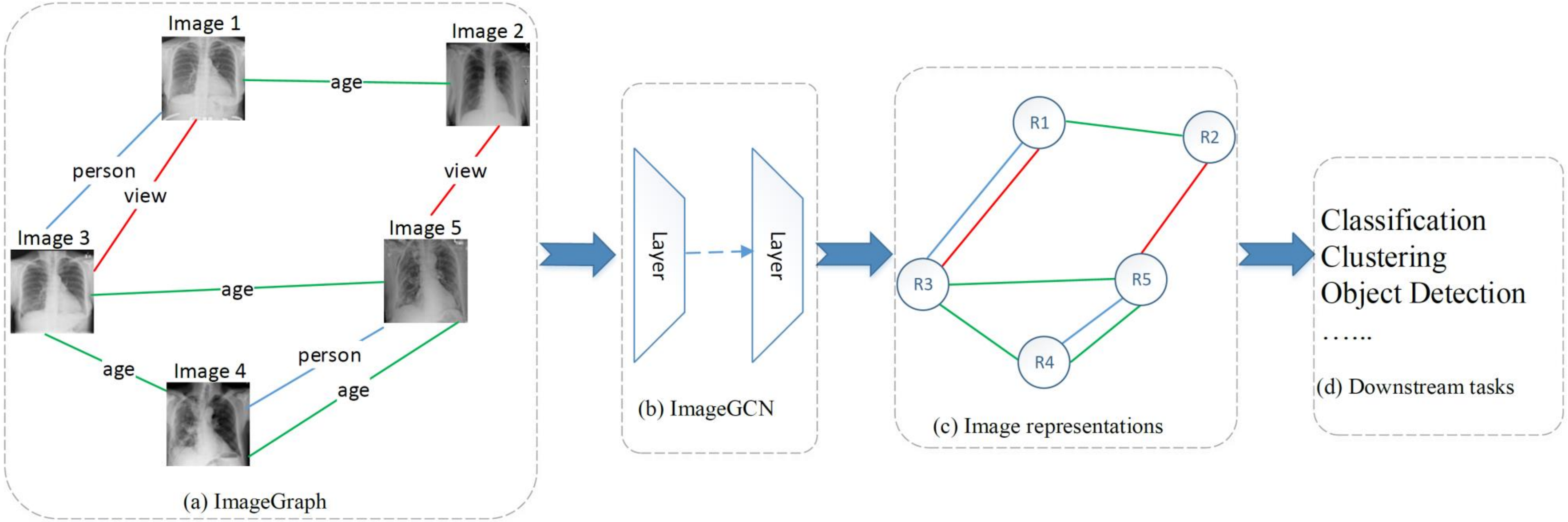
Omics data




Time series

Structured data

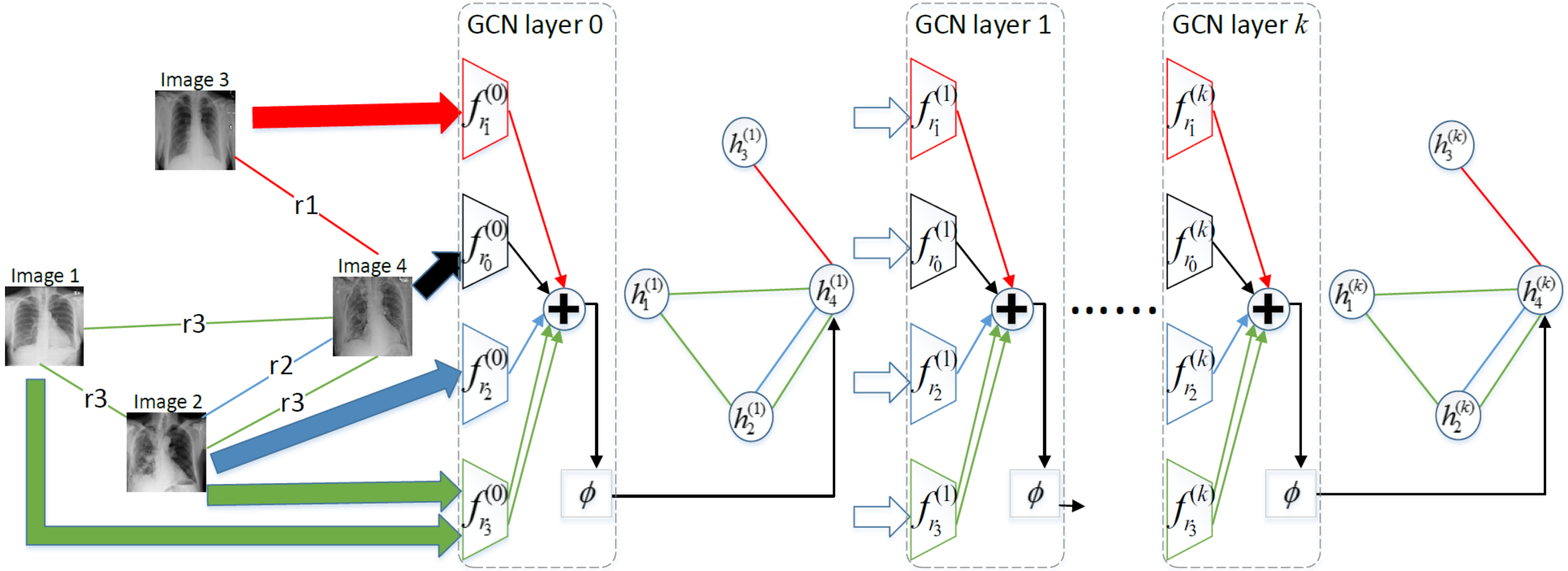
# ImageGCN: Graph Neural Networks to model relations between images




C Mao, L Yao, Y Luo, ImageGCN: Multi-Relational Image Graph Convolutional Networks for Disease Identification with Chest X-rays. *IEEE Transactions on Medical Imaging 2022 10.1109/TMI.2022.3153322*.

# ImageGCN: Graph Neural Networks to Model Relations Between Images

Multi-dimensional relations between medical images =>
deep graph neural networks =>
Better interpretability and better accuracy



# ImageGCN: Graph Neural Networks to Model Relations Between Images

| | Atel | Card | Effu | Infi | Mass | Nodu | Pneu1 | Pneu2 | Cons | Edem | Emph | Fibr | PT | Hern | mean | MR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A-GCN-PPS (ours) | 0.781 | 0.899 | 0.865 | 0.701 | 0.813 | 0.721 | 0.718 | 0.881 | 0.788 | 0.888 | 0.882 | 0.804 | 0.778 | 0.904 | 0.816 | 6.73 |
| A-GCN-APS | 0.739 | 0.876 | 0.815 | 0.671 | 0.799 | 0.704 | 0.679 | 0.857 | 0.762 | 0.846 | 0.863 | 0.792 | 0.765 | 0.910 | 0.791 | 13.73 |
| AlexNet | 0.782 | 0.895 | 0.863 | 0.705 | 0.781 | 0.714 | 0.716 | 0.869 | 0.790 | **0.889** | 0.876 | 0.799 | 0.773 | 0.899 | 0.811 | 7.93 |
| R-GCN-PPS (ours) | 0.785 | 0.890 | 0.868 | 0.699 | 0.824 | 0.739 | 0.723 | 0.895 | 0.790 | 0.887 | 0.911 | 0.819 | 0.786 | 0.941 | 0.826 | 4.47 |
| R-GCN-APS | 0.741 | 0.861 | 0.822 | 0.680 | 0.819 | 0.728 | 0.684 | 0.873 | 0.768 | 0.852 | 0.889 | 0.790 | 0.751 | 0.908 | 0.798 | 12.27 |
| ResNet50 | 0.789 | 0.889 | 0.863 | 0.698 | 0.807 | 0.723 | 0.714 | 0.876 | 0.791 | 0.888 | 0.899 | 0.799 | 0.772 | 0.933 | 0.817 | 6.93 |
| V-GCN-PPS (ours) | **0.802** | 0.894 | **0.874** | 0.702 | **0.843** | 0.768 | 0.715 | **0.900** | 0.796 | 0.883 | 0.915 | 0.825 | **0.791** | **0.943** | **0.832** | **2.60** |
| V-GCN-APS | 0.754 | 0.871 | 0.826 | 0.676 | 0.820 | 0.737 | 0.688 | 0.872 | 0.769 | 0.839 | 0.894 | 0.789 | 0.770 | 0.926 | 0.802 | 11.07 |
| VGGNet16BN | 0.796 | 0.893 | 0.872 | 0.700 | 0.831 | 0.756 | 0.717 | 0.882 | 0.794 | 0.878 | 0.909 | 0.799 | 0.785 | 0.923 | 0.824 | 4.60 |
| Wang et al. [8] | 0.716 | 0.807 | 0.784 | 0.609 | 0.706 | 0.671 | 0.633 | 0.806 | 0.708 | 0.835 | 0.815 | 0.769 | 0.708 | 0.767 | 0.738 | 17.33 |
| Yao et al. [11] | 0.772 | 0.904 | 0.859 | 0.695 | 0.792 | 0.717 | 0.713 | 0.841 | 0.788 | 0.882 | 0.829 | 0.767 | 0.765 | 0.914 | 0.803 | 10.80 |
| Li et al. [44] | 0.800 | 0.870 | 0.870 | 0.700 | 0.830 | 0.750 | 0.670 | 0.870 | **0.800** | 0.880 | 0.910 | 0.780 | 0.760 | 0.770 | 0.804 | 8.40 |
| Kumar et al. [12] | 0.762 | **0.913** | 0.864 | 0.692 | 0.750 | 0.666 | 0.715 | 0.859 | 0.784 | 0.888 | 0.898 | 0.756 | 0.774 | 0.802 | 0.794 | 10.67 |
| Tang et al. [45] | 0.756 | 0.887 | 0.819 | 0.689 | 0.814 | 0.755 | 0.729 | 0.85 | 0.728 | 0.848 | 0.906 | 0.818 | 0.765 | 0.875 | 0.803 | 10.33 |
| Shen et al. [46] | 0.766 | 0.801 | 0.797 | **0.751** | 0.76 | 0.741 | **0.778** | 0.800 | 0.787 | 0.82 | 0.773 | 0.765 | 0.759 | 0.748 | 0.775 | 13.40 |
| Mao et al. [13] | 0.750 | 0.869 | 0.810 | 0.687 | 0.782 | 0.726 | 0.695 | 0.845 | 0.728 | 0.834 | 0.870 | 0.798 | 0.758 | 0.877 | 0.788 | 14.33 |
| Guan et al. [14] | 0.781 | 0.883 | 0.831 | 0.697 | 0.83 | 0.764 | 0.725 | 0.866 | 0.758 | 0.853 | 0.911 | 0.826 | 0.78 | 0.918 | 0.816 | 7.00 |
| Liu et al. [31] | 0.773 | 0.889 | 0.821 | 0.710 | 0.829 | **0.770** | 0.713 | 0.869 | 0.749 | 0.847 | **0.934** | **0.845** | 0.773 | 0.925 | 0.818 | 6.93 |
| p-val (RM-ANOVA) | *** | *** | *** | *** | * | * | *** | *** | *** | *** | *** | ** | ** | | *** | |
| PPS > APS | *** | *** | *** | *** | * | ** | *** | *** | *** | *** | *** | ** | ** | | *** | |
| PPS > base | | * | ** | | ** | ** | | *** | | | ** | * | ** | * | ** | |

Atel: Atelectasis; Card: Cardiomegaly; Effu: Effusion; Infi: Infiltration; Nodu: Nodule; Pneu1: Pneumonia; Pneu2:Pneumothorax; Cons: Consolidation; Edem: Edema; Emph: Emphysema; Fibr: Fibrosis; PT:Pleural Thickening; Hern: Hernia. MR=mean rank

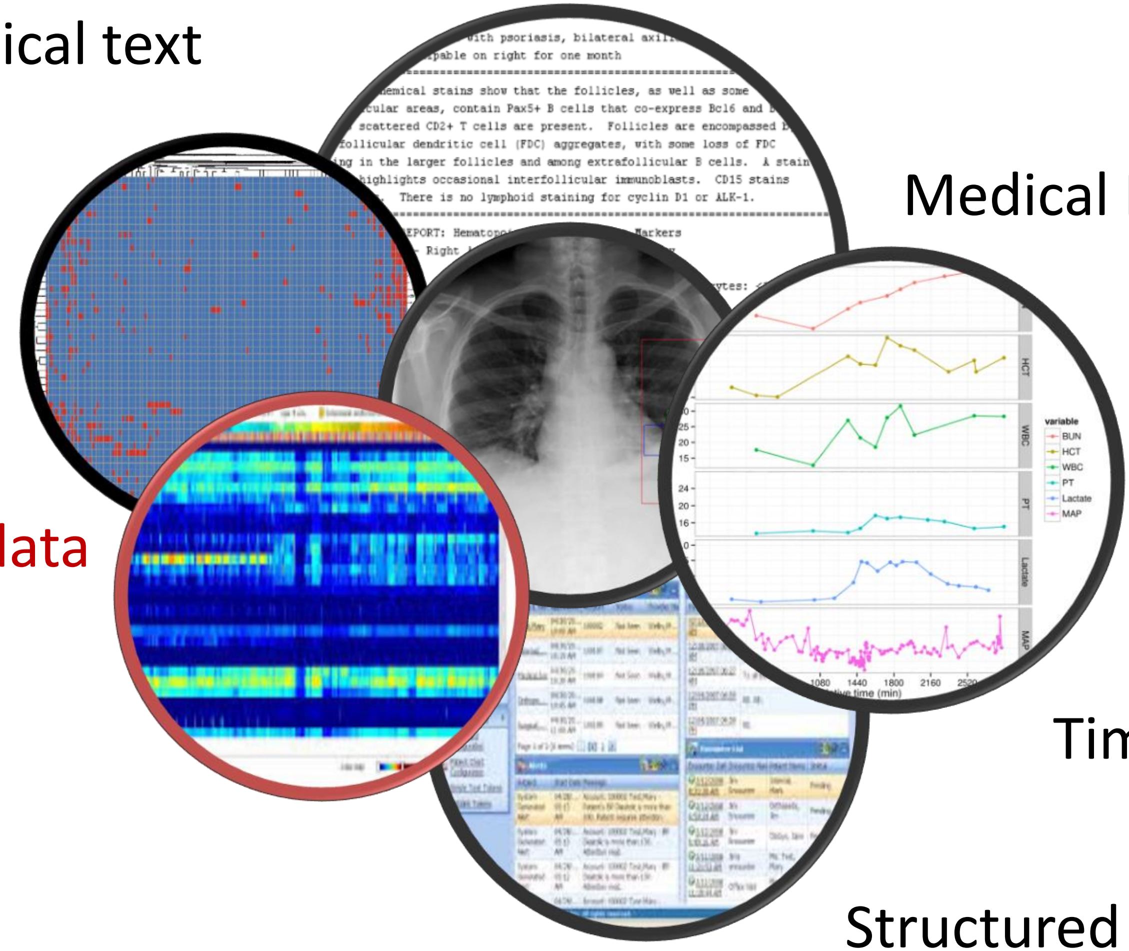
https://github.com/mocherson/ImageGCN

# Machine learning to integrate multi-modal healthcare data

Biomedical text

Medical Imaging

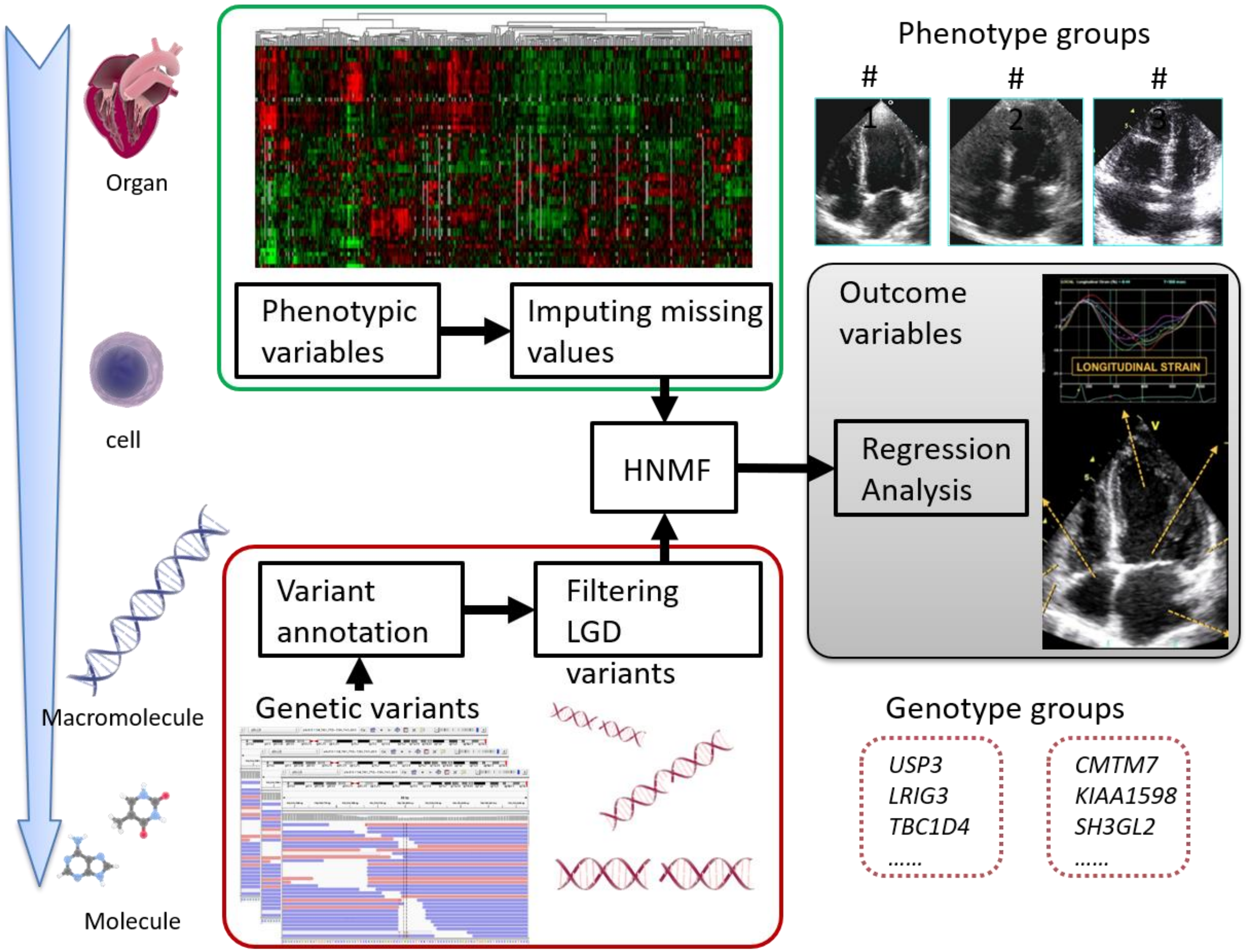
Omics data



Time series

Structured data

# Integrative omics: CVD precision medicine




Integrating phenotype and genotype to better subtype heterogeneous patient cohort and inform targeted therapy

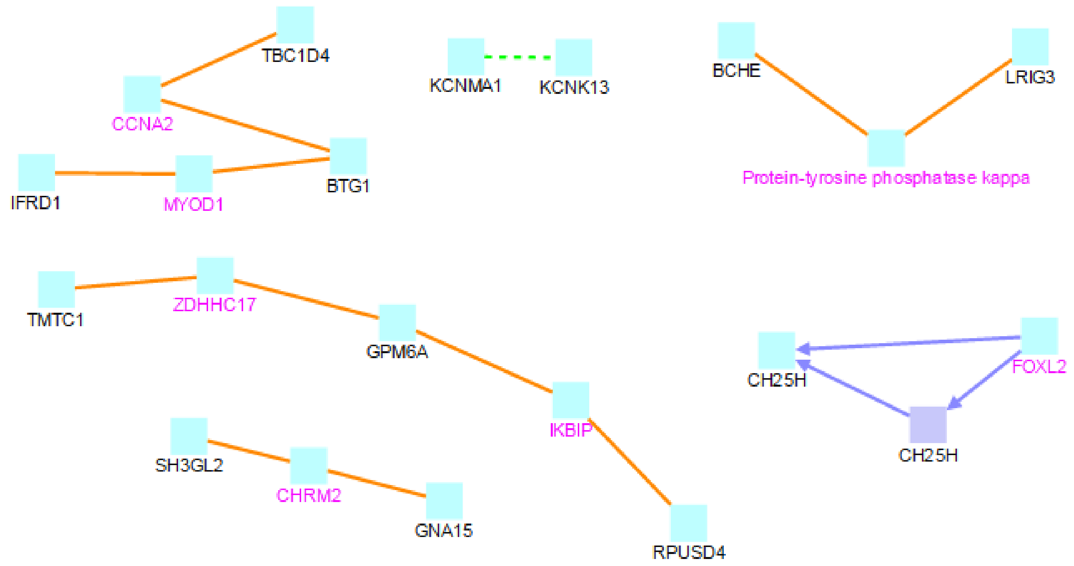
Y Luo, C Mao, Y Yang, F Wang, FS Ahmad, D Arnett, MR Irvin, SJ Shah. Integrating Hypertension Phenotype and Genotype with Hybrid Non-negative Matrix Factorization. *Bioinformatics 2018 PMID: 30239588*

# PANTHER: Pathway Augmented Nonnegative Tensor factorization for HighER-order feature learning in genetic medicine

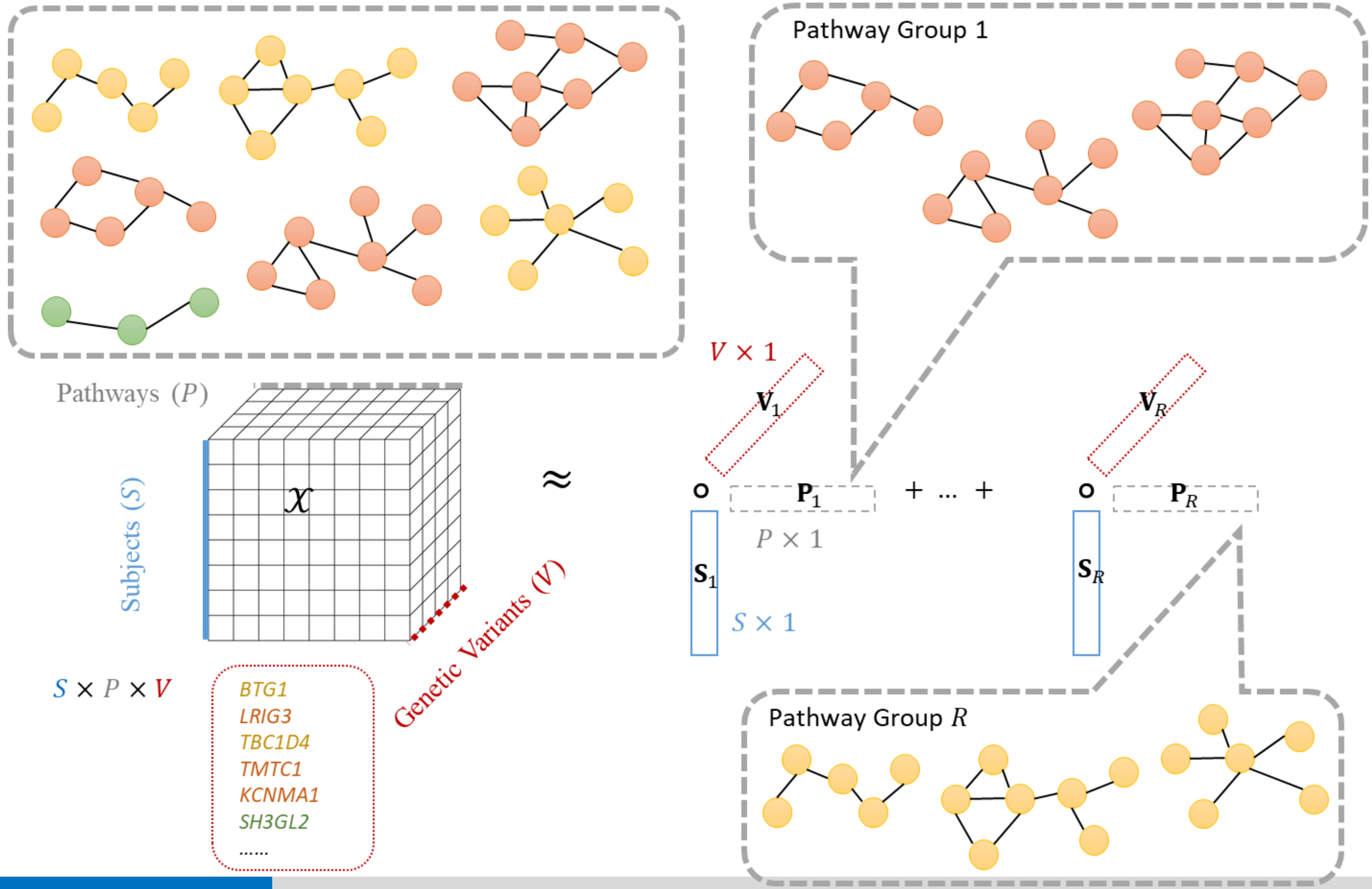
- Genetic pathways are important in understanding the molecular mechanisms
- Individual pathways are part of the entire biological system and interact with each other
- It is important to model co-functioning molecular mechanisms by grouping pathways together

*BTG1*
*LRIG3*
*TBC1D4*
*TMTC1*
*KCNMA1*
*SH3GL2*
……

v.s.

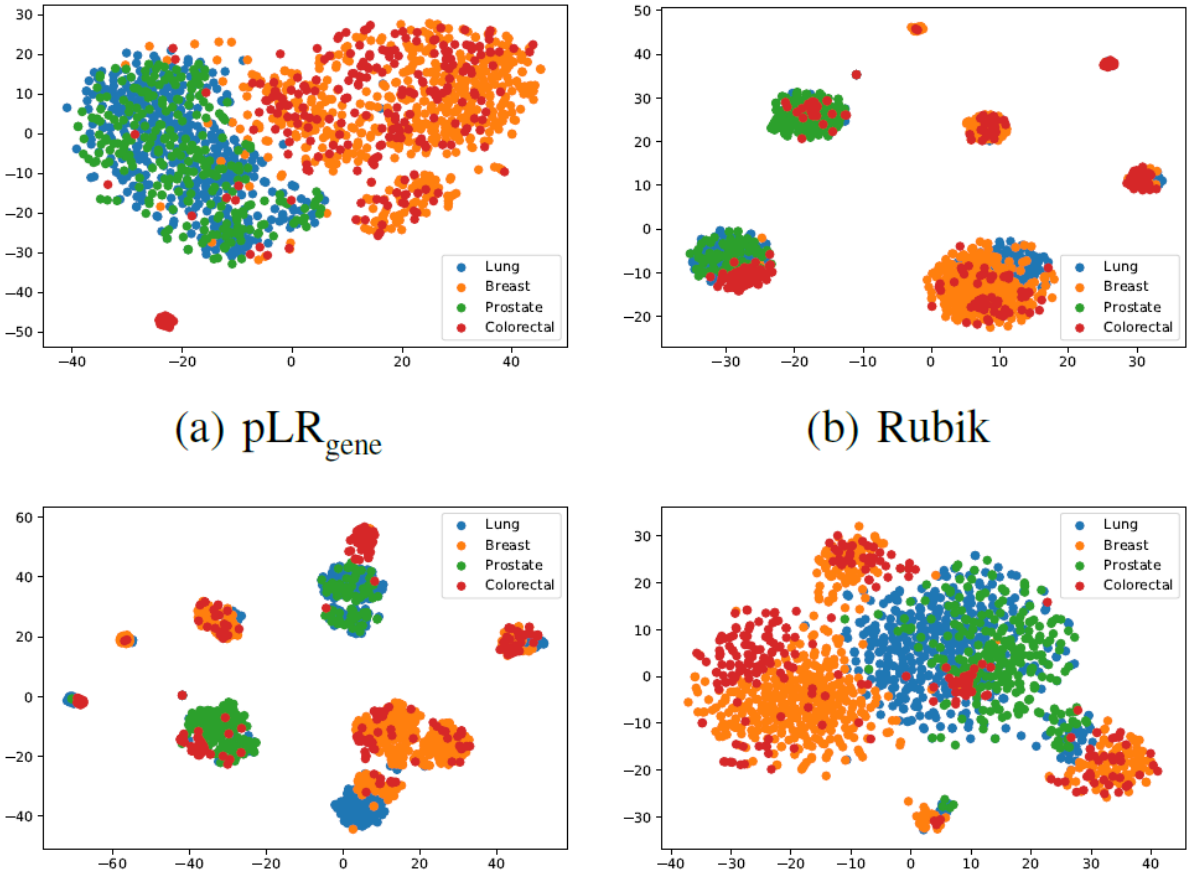


# PANTHER factorization scheme



# Results and t-SNE visualization of the learned train set subject features

| Model | $R$ | Test Accuracy |
|---|---|---|
| $\mathrm{pLR}_{\mathrm{gene}}$ | - | 0.8016 |
| $\mathrm{pLR}_{\mathrm{pathway}}$ | - | 0.7701 |
| $\mathrm{pLR}_{\mathrm{gene+pathway}}$ | - | 0.7682 |
| $\mathrm{brms}_{\mathrm{gene}}$ | - | 0.8016 |
| $\mathrm{brms}_{\mathrm{pathway}}$ | - | 0.7642 |
| $\mathrm{brms}_{\mathrm{gene+pathway}}$ | - | 0.8173 |
| $\mathrm{NMF}_{\mathrm{gene}}$ | 400 | 0.8173 |
| $\mathrm{NMF}_{\mathrm{pathway}}$ | 400 | 0.7819 |
| $\mathrm{NMF}_{\mathrm{gene+pathway}}$ | 100 | 0.7957 |
| $\mathrm{grpreg}_{\mathrm{gene}}$ | 66 | 0.6306 |
| $\mathrm{grpreg}_{\mathrm{pathway}}$ | 66 | 0.8153 |
| $\mathrm{grpreg}_{\mathrm{gene+pathway}}$ | 131 | 0.8016 |
| Rubik | 100 | 0.8035 |
| SUSTain | 450 | 0.7839 |
| SURF | 50 | 0.7466 |
| LogPar | 350 | 0.6031 |
| TASTE | 200 | 0.8369 |
| LOM | 100 | 0.7505 |
| PANTHER | 200 | **0.8644** |




(a) $\mathrm{pLR}_{\mathrm{gene}}$ (b) Rubik (c) TASTE (d) PANTHER

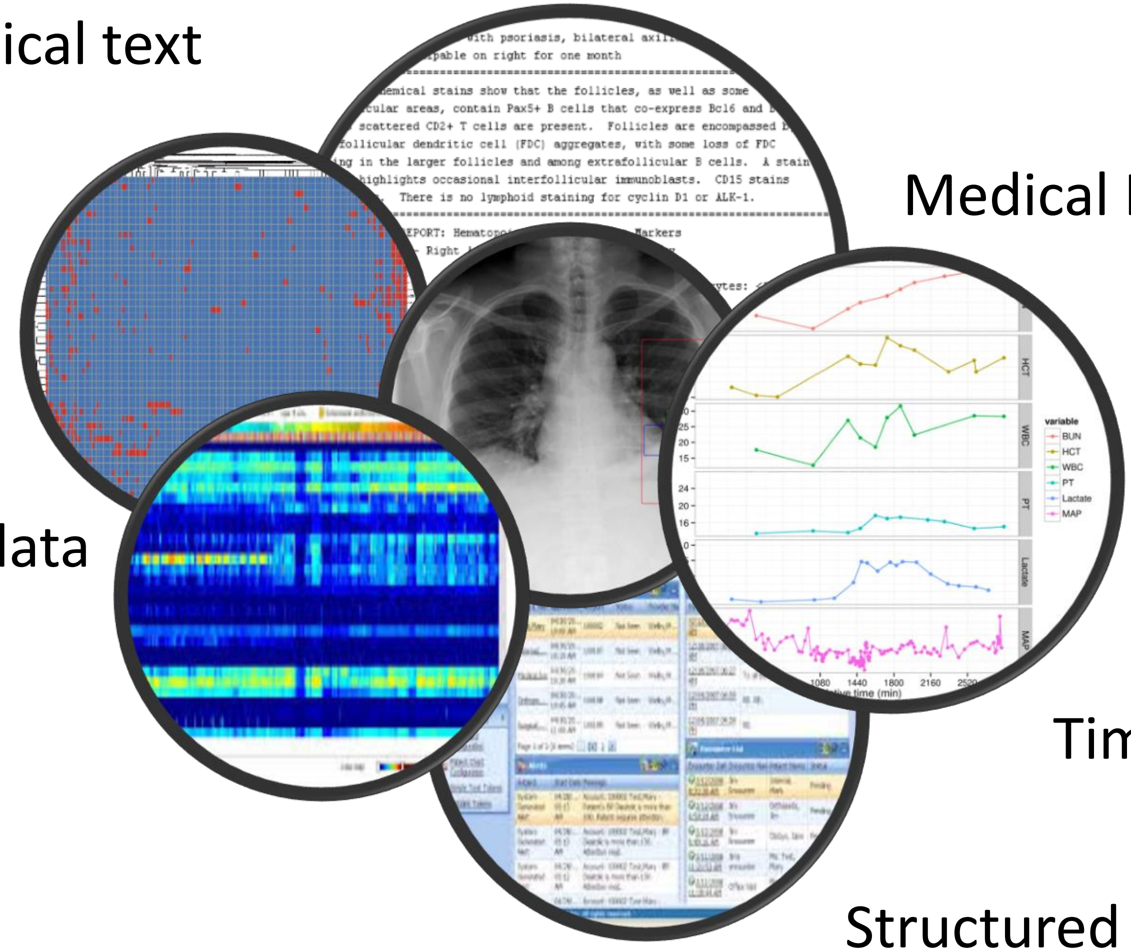
Y Luo, C Mao. PANTHER: Pathway Augmented Nonnegative Tensor factorization for HighER-order feature learning. *AAAI 2021*, *371-380*.

# Machine learning to integrate multi-modal healthcare data

Biomedical text

Medical Imaging

Omics data




Time series

Structured data

nature medicine

BRIEF COMMUNICATION

https://doi.org/10.1038/s41591-020-1007-0

Check for updates

# A multidimensional precision medicine approach identifies an autism subtype characterized by dyslipidemia

Yuan Luo[1,2,3,4,10], Alal Eran[5,6,7,10], Nathan Palmer[6], Paul Avillach[6], Ami Levy-Moonshine[8], Peter Szolovits[9] and Isaac S. Kohane[6] ✉






# Autism Spectrum Disorder (ASD)

- Common neurodevelopmental disorder
- Prevalence ~1%, male centric, early onset in childhood
- A strong genetic component
- Early molecular diagnosis is a promising path
- Unknown specific genetic cause in most cases

Photo credit: The Kobal Collection

# Identification of 33 neurodevelopmentally co-regulated, sex-differentially expressed exon clusters with ASD segregating deleterious variation

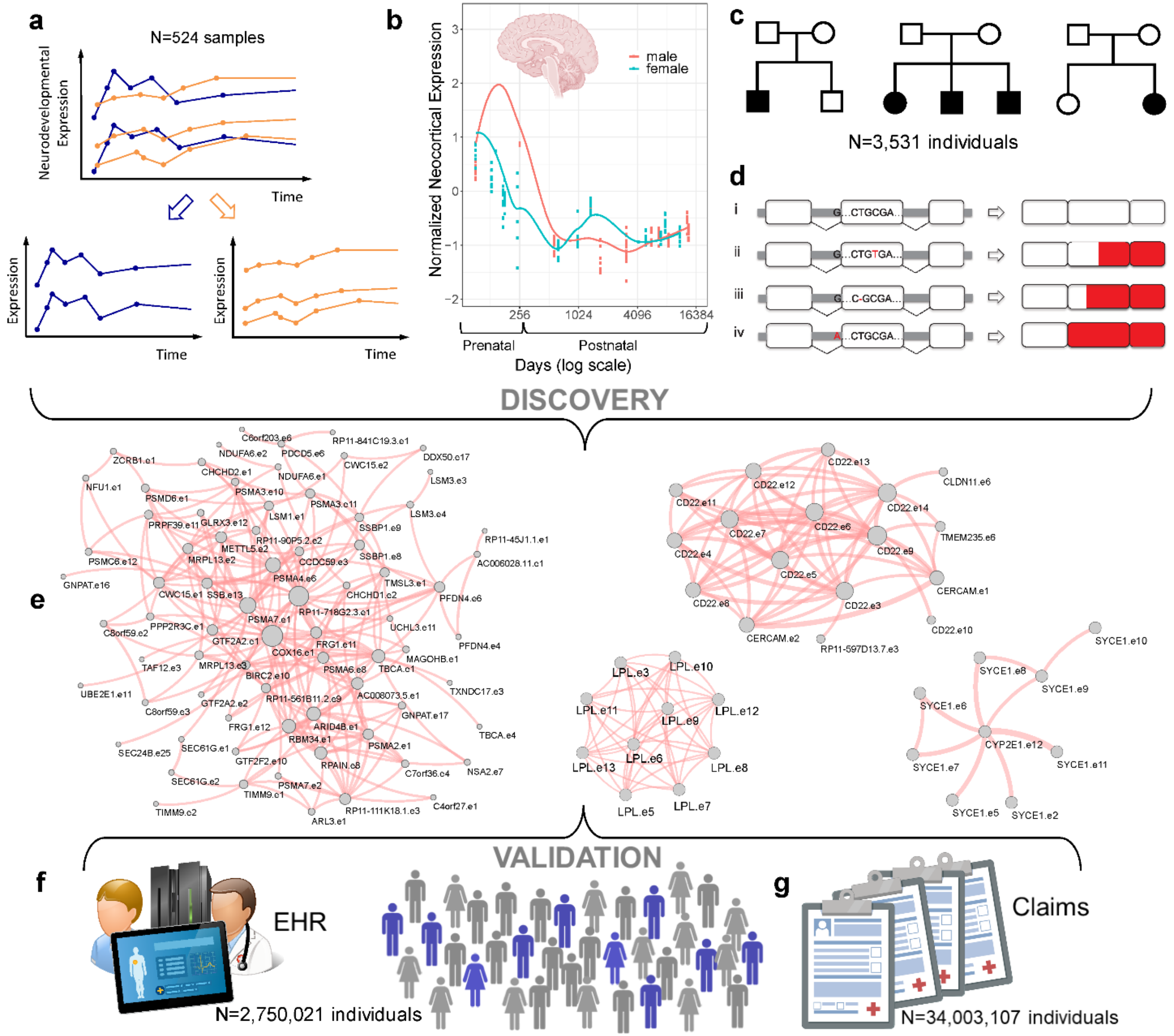



Figure.1

# One large cluster (84 genes) associated with lipid regulation functions

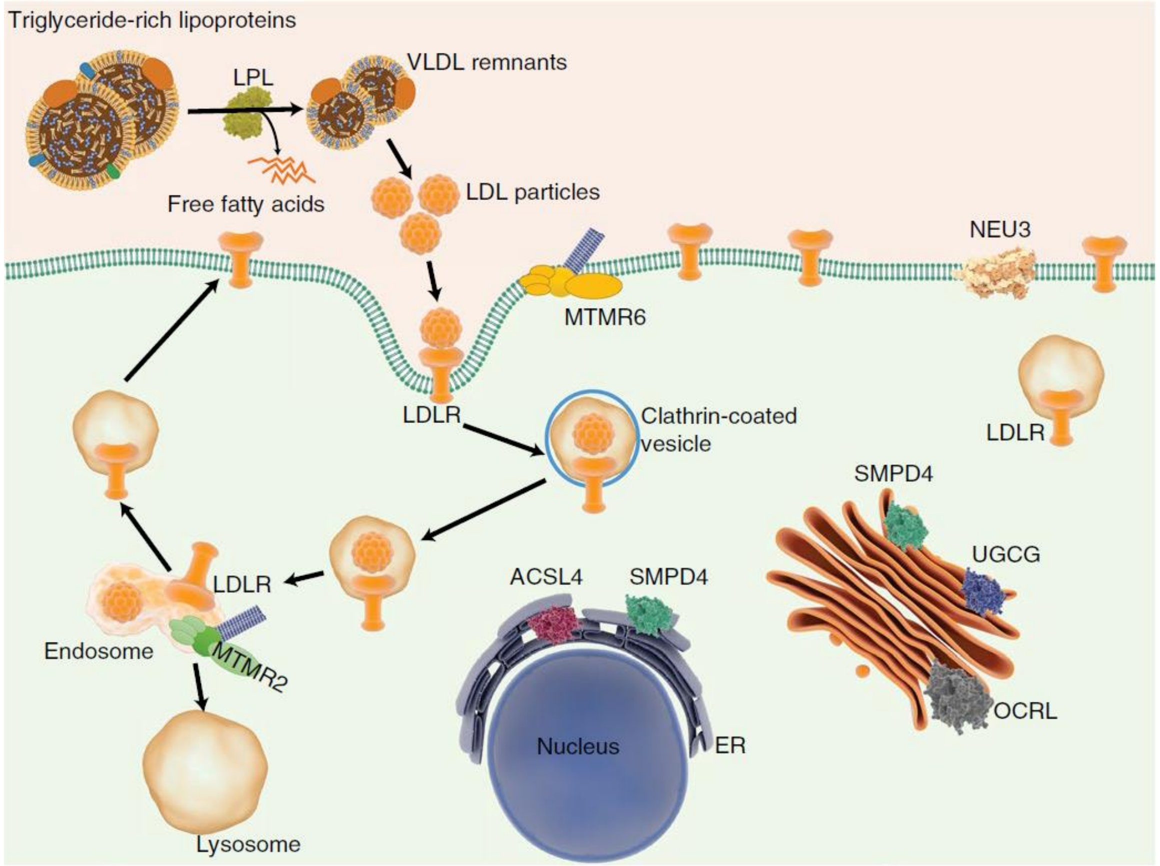



Figure.2c

| Cluster ID | Functional enrichment |
|---|---|
| 5965 | Reactome metabolism of lipids and lipoproteins |

# One large cluster (84 genes) associated with lipid regulation functions

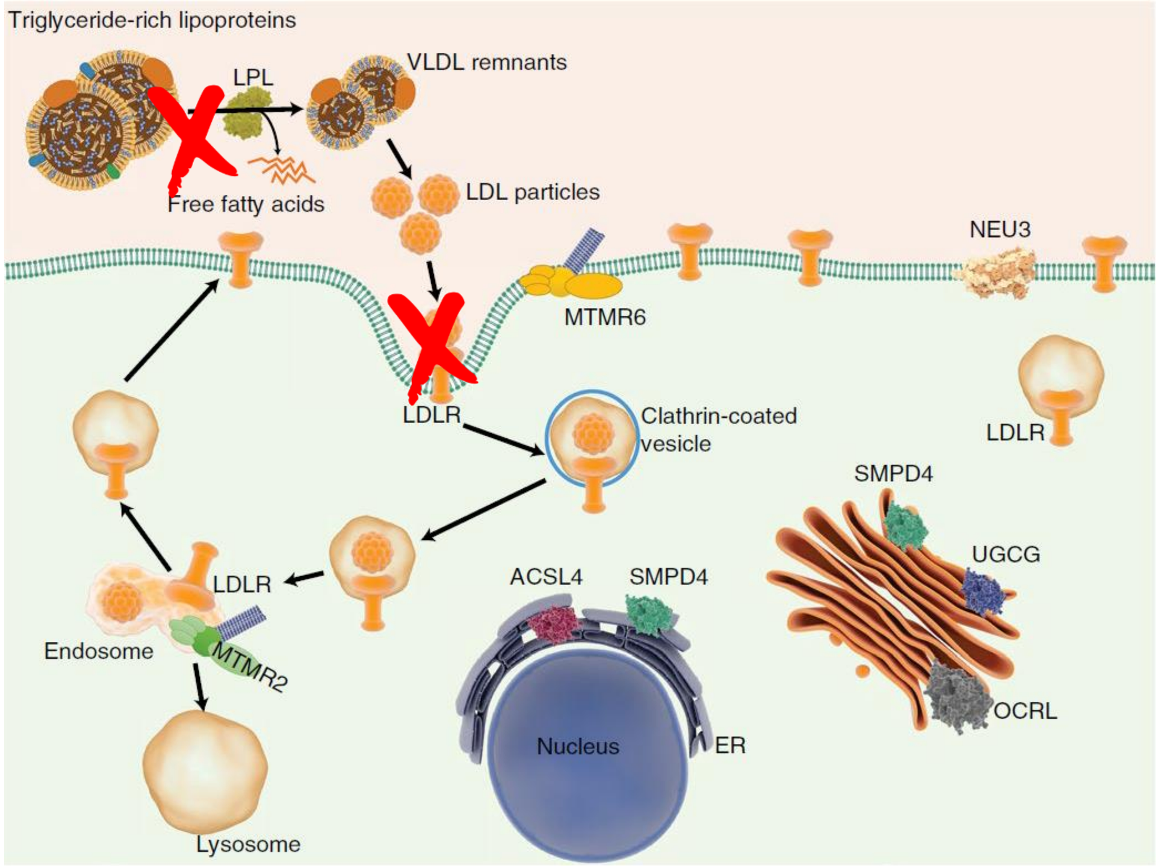



Figure.2c

| Cluster ID | Functional enrichment |
|---|---|
| 5965 | Reactome metabolism of lipids and lipoproteins |

# ASD have blood lipid profiles that are outside the normal range – Boston Children's Hospital (BCH)

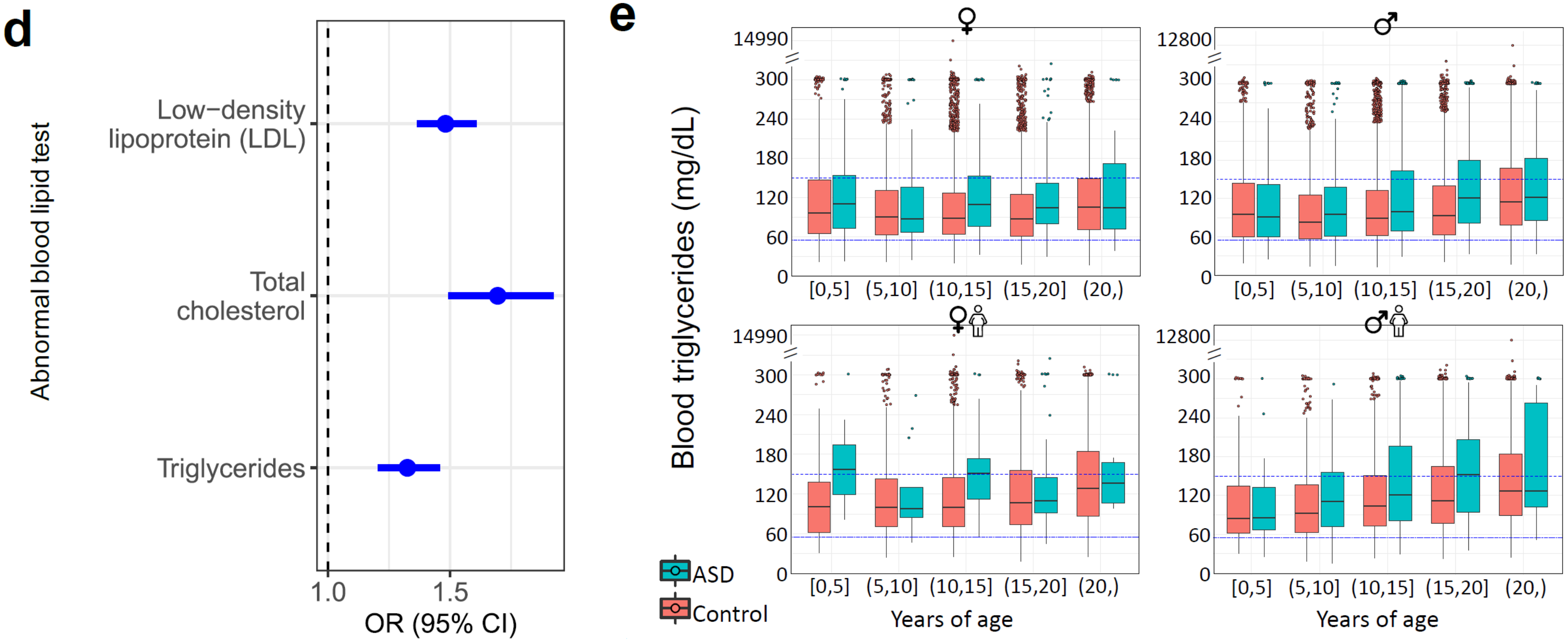



Figure 2e: Results from individuals with metabolic dysregulation, including obesity, diabetes and metabolic syndrome X, are shown at the bottom.

Figure 2d and 2e

# Enrichment of dyslipidemia diagnoses in individuals with ASD – Aetna healthcare claims

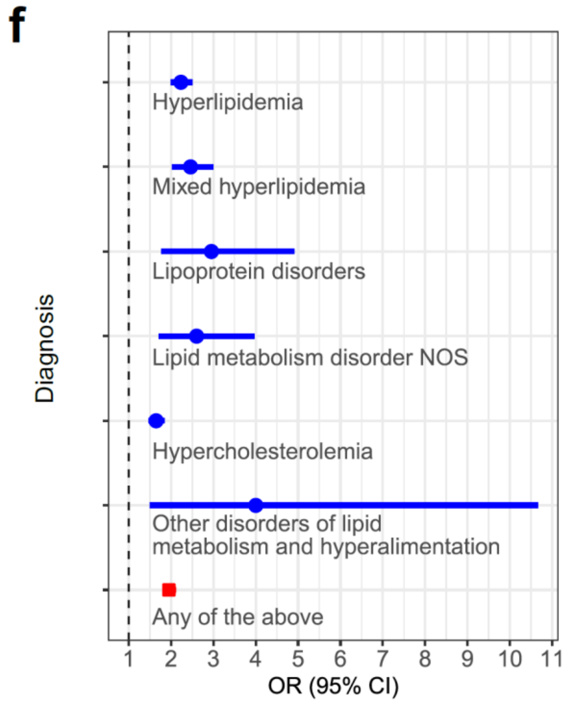
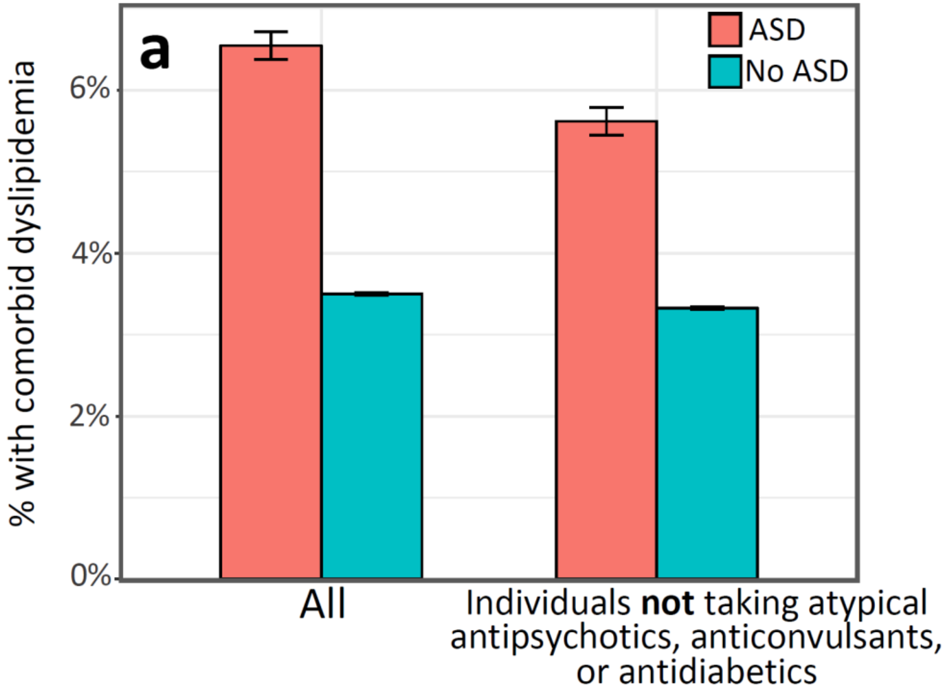
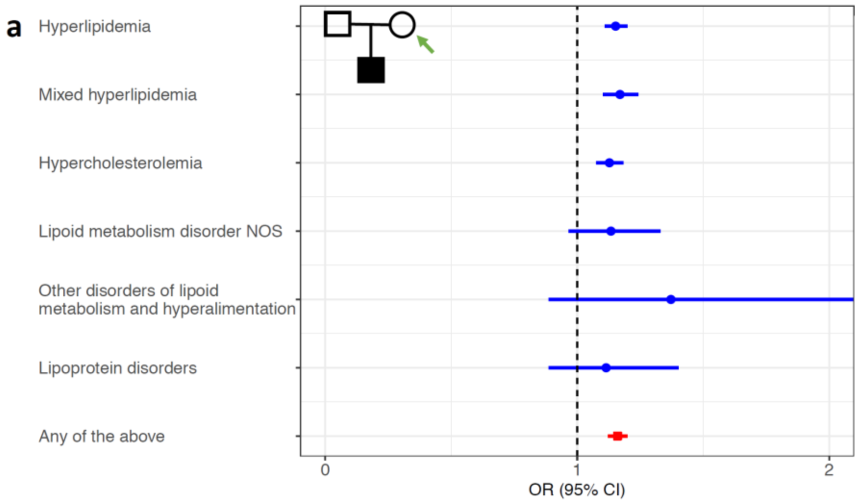


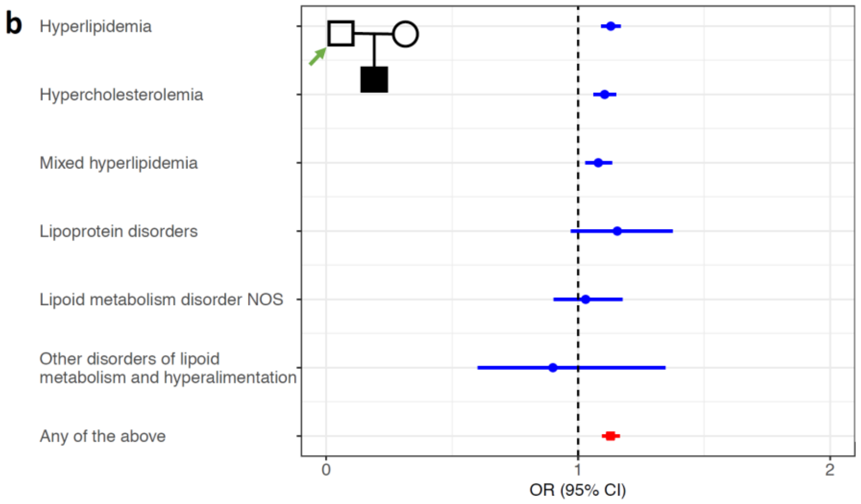

Figure 2f




Extended figure 2a

# Enrichment of dyslipidemia diagnoses in parents of children with ASD – Aetna healthcare claims

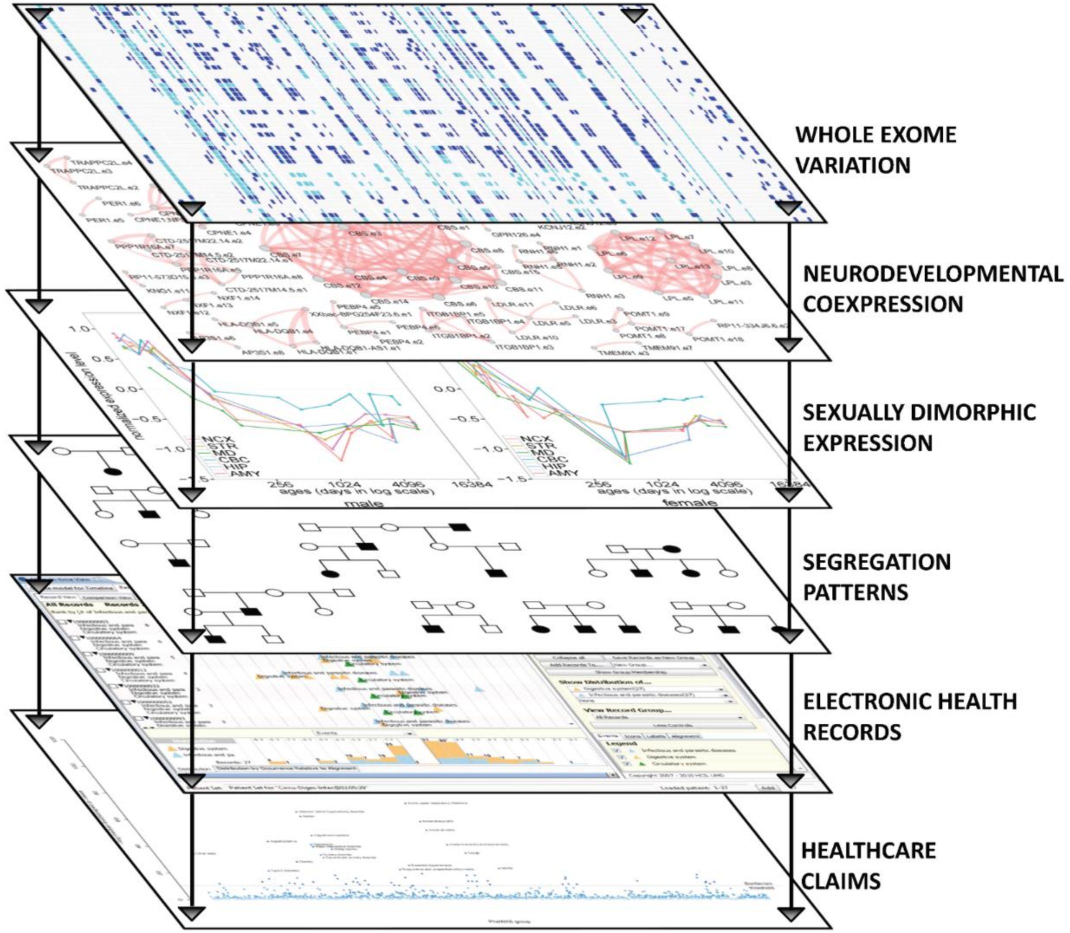



Extended figure 3

# ASD study conclusion

- New AI-driven precision medicine method to make sense out of massive genomic, transcriptomic, EHR and claims data etc.
- Identified a previously unrecognized risk factor for ASD which was verified in large clinical cohorts
- Opens avenue for early detection, intervention, and better outcomes for individuals with ASD and their family members




**Source code**: https://github.com/luoyuanlab/autism_precision_medicine

# The pandemic as a stress test for machine learning in healthcare

REVIEW ARTICLE | FOCUS

nature medicine

https://doi.org/10.1038/s41591-018-0300-7

**High-performance medicine: the convergence of human and artificial intelligence**

Eric J. Topol

nature machine intelligence

ANALYSIS

https://doi.org/10.1038/s42256-021-00307-0

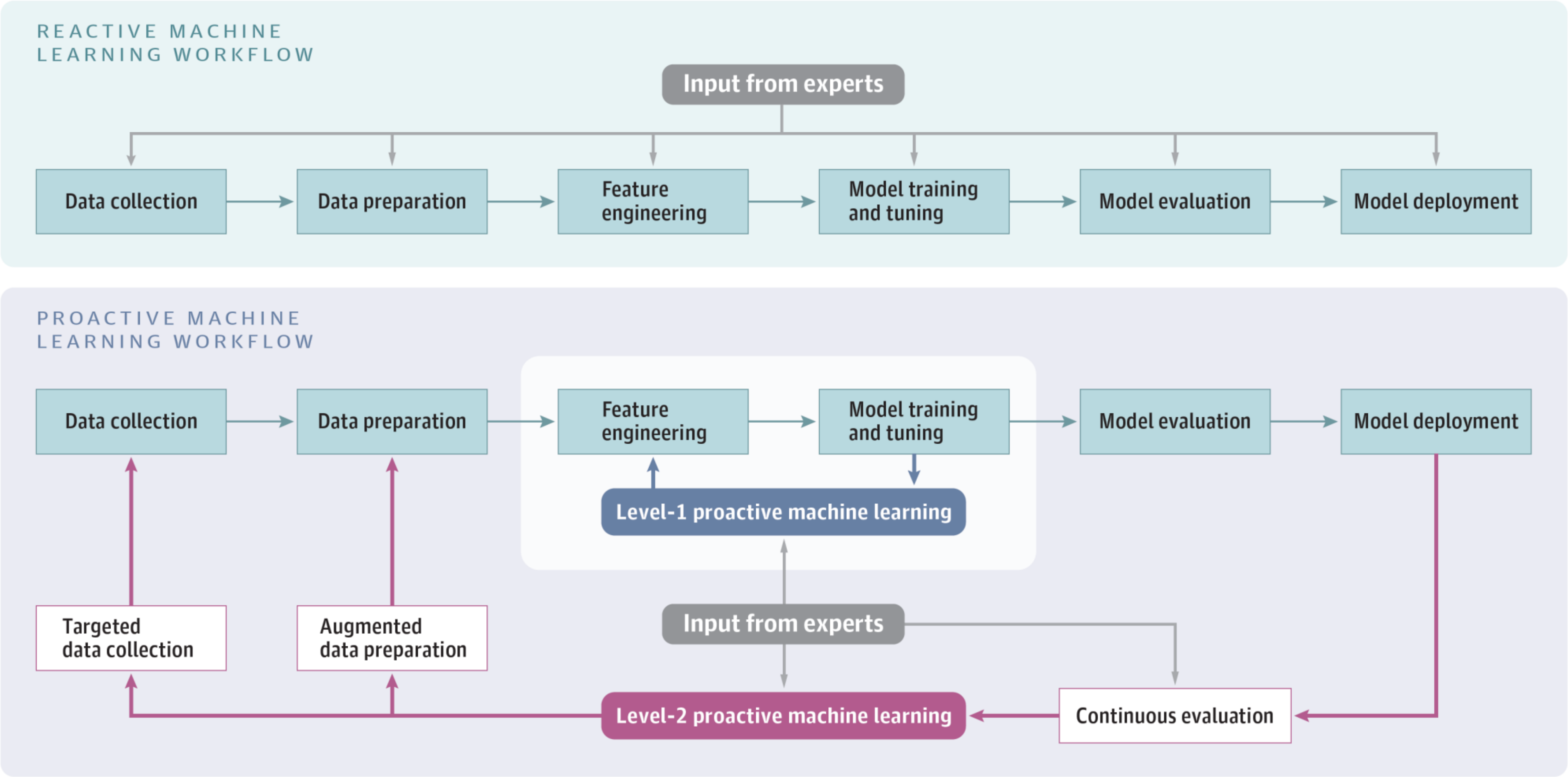
Check for updates

OPEN

**Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans**

Michael Roberts[1,2] ✉, Derek Driggs[1], Matthew Thorpe[3], Julian Gilbey[1], Michael Yeung[4], Stephan Ursprung[4,5], Angelica I. Aviles-Rivero[1], Christian Etmann[1], Cathal McCague[4,5], Lucian Beer[4], Jonathan R. Weir-McCall[4,6], Zhongzhao Teng[4], Effrossyni Gkrania-Klotsas[7], AIX-COVNET*, James H. F. Rudd[8,36], Evis Sala[4,5,36] and Carola-Bibiane Schönlieb[1,36]

# Moving from reactive to proactive multi-modal machine learning




Luo Y, Wunderink RG, Lloyd-Jones D. Proactive vs Reactive Machine Learning in Health Care: Lessons From the COVID-19 Pandemic. *JAMA. 2022*.

# Let us work together and bring it to a whole new level

- Collaboration welcome
- yuan.luo@northwestern.edu
- @yuanhypnosluo
- We are hiring, multiple postdoc positions available
- https://labs.feinberg.northwestern.edu/luolab/
- Main funding support acknowledgement
  - U01TR003528
  - U54HL160273
  - R01LM013337
  - R01GM105688
  - R21LM012618
  - UL1TR001422
  - U01HG011169

NUCATS Clinical and Translational Sciences Institute

Northwestern

Northwestern Medicine

eMERGE Network electronic medical records & genomics