CURRICULUM VITAE

- **Personal Details**

Name	Nadav Rappoport
E-mail	nadavrap@bgu.ac.il
Address	Department of Software and Information Systems Engineering Ben-Gurion University of the Negev Beer Sheva, 8410501, ISRAEL
Work phone number	+972-74-779-5156
ORCID ID	0000-0002-7218-2558

- **Education**

B.Sc.	The Hebrew University of Jerusalem, Computer Science and Computational Biology	2006–2009
M.Sc.	The Hebrew University of Jerusalem, Computer Science and Computational Biology	2010–2011
Ph.D.	The Hebrew University of Jerusalem, Computer Science Advisor: Prof. Michal Linial Title of dissertation: *"Gaining functional and quantitative knowledge on genes and molecular machines using a computational approach"*	2010–2011

- **Employment History**

Senior Lecturer (Assistant Professor) Department of Software and Information Systems Engineering Ben-Gurion University of the Negev, ISRAEL	2019–*today*
Adjunct Assistant Professor Internal Medicine The University of Utah, USA	2024–*today*
Lecturer Yandex School of Data Science, Israel	2019–2020
Post-doctoral fellow Bakar Computational Health Science Institute University of California San Francisco, USA	2015–2019
Lecturer The Hebrew University of Jerusalem, Israel	2012–2015

Lecturer 2014
Azrieli College of Engineering Jerusalem, Israel

Teaching Assistant 2009–2015
School of Computer Science and Engineering. The Hebrew University of Jerusalem, Israel

- **Professional Activities**

(a) Positions in academic administration

Head of Medical Informatics M.Sc. track 2022–now
Department of Software and Information Systems Engineering
Ben-Gurion University of the Negev, Israel

Fourth year final projects' coordinator 2021–2023
Department of Software and Information Systems Engineering
Ben-Gurion University of the Negev, Israel

Graduate Committee member 2021–now
Department of Software and Information Systems Engineering
Ben-Gurion University of the Negev, Israel

Committee member of the Master studies in cyber security program 2021–now
Department of Software and Information Systems Engineering
Ben-Gurion University of the Negev, Israel

Teaching committee member of the B.Sc. program 2021–now
Department of Software and Information Systems Engineering
Ben-Gurion University of the Negev, Israel

Developing and presenting new M.Sc. program in Medical Informatics 2020–2024
Department of Software and Information Systems Engineering
Ben-Gurion University of the Negev, Israel

(b) Professional functions outside universities/institutions (inter-university, national, international)

Consultant 2019–now
Israeli Ministry of Health

Program committee member 2025
AMIA

Head of Special Session, IDSAI2025 2025

Organizing committee of IDSAI2025 2024–2025
The 4th International Israel Data Science Initiative Conference

Program committee member 2023
STC EFMI

Organizing committee of IBS2024 22st Israeli Bioinformatics Symposium	2023–2024
Organizing committee of IDSI2024 (The Third International Israel Data Science Initiative Conference)	2023–2024
Session chair "Data-science for improving health care systems", IDSI	2023
Program committee member AMIA	2022-2024
Organizer: AI and Big Data in Medical and Public Health Sciences Conference	2022
Program committee member, MIE	2022–2024
Program committee member AIME	2021–2024
Program Board Member Science Fund of the Republic of Serbia: Member of the Program Board for evaluation of project proposals within the Special COVID-19 Research Program of the Science Fund, 2020	2020
Program committee member AFP, ISMB/ECCB	2013

(c) Significant professional consulting

Consultant Private consultants to medical parties, including Clalit Health Services and the Israeli Ministry of Health	2019–today

(d) Editor or member of editorial board of scientific or professional journal

Section Editor PLoS Digital Health journal	2024–today
Associate Editor PLoS Digital Health journal	2021–today
Guest Editor PLoS Digital Health journal	2021
Sepcial Issue Editor Data Science in Health Care Journal of Big Data and Cognitive Computing	2021–2023

(e) Ad-hoc reviewer for journals

Reviewer 2021, 2023
IJERPH

Reviewer 2021
JAIR

Reviewer 2021
Journal of Biomedical Informatics

Reviewer 2021–2023
PLoS Digital Health

Reviewer 2020
Journal of Epidemiology and Public Health Reviews

Reviewer 2020–2022
Vaccines

Reviewer 2020
PLoS One

Reviewer 2020
Scientific Reports

Reviewer 2019–2021
Scientific Data

Reviewer 2017, 2020–2021
PLoS One

Reviewer 2017
Journal of Epidemiology and Public Health Reviews

Reviewer 2016
PLoS Computational Biology

Reviewer 2022
Frontiers

Reviewer 2022
Healthcare (MDPI)

Reviewer 2023
JMIR

Reviewer 2024
Heliyon

(f) Ad-hoc reviewer for conferences

Reviewer 2013, 2016, 2018, 2019, 2022
RECOMB

Reviewer 2021–2024
AIME

Reviewer AMIA	2018–2020, 2022–2024
Reviewer ISMB/ECCB	2019,2022
Reviewer ISMB	2022
Reviewer MIE	2022–2024
Reviewer ICML	2024
Reviewer STC	2023, 2024

(g) Ad-hoc reviewer for competitive grants

Reviewer Ministry of Science and Technology of Israel (MOST) research proposals	2023–2024

(h) Membership in professional/scientific societies

International Society for Computational Biology (ISCB)	2010–today
American Medical Informatics Association (AMIA)	2016–today
Israeli Association for Medical Informatics (ILAMI)	2020–today
The American Society of Human Genetics (ASHG)	2022–today

- **Educational activities**

(a) Courses taught

Machine Learning	Graduate	Ben-Gurion University	2020–today
Introduction to Clinical Data Science	Graduate	Ben-Gurion University	2020–today
Coordinating final undergrad projects	Undergraduate	Ben-Gurion University	2022–2023
Machine Learning MOOC	Undergraduate	Campous.gov.il	2022–2023
Machine Learning	Undergraduate	Ben-Gurion University	2022–today

Decision Analysis	Undergraduate	Ben-Gurion University	2021
C programming for CS	Undergraduate	The Hebrew University	2017–2019
C++ programming for CS	Undergraduate	The Hebrew University	2017–2019
C++ programming for non-CS	Undergraduate	The Hebrew University	2018
MS Access	Undergraduate	The Hebrew University	2018–2019

(b) Research students

M.Sc. Students

Tomer Hermelin	
Thesis title: Mortality prediction with alert intervals	Graduated 2022
Bar Ezra	
Thesis title: Multi-dimensional laboratory test results	Graduated 2022
Gal Shlomo	
Thesis title: Early detection for patients with high risk for non-alcoholic fatty liver disease	Graduated 2022
Yarden Hochberg	
Thesis title: Complex models for polygenic risk scores	
Recipient of Interdisciplinary fellowship (co-advised with Yonathan Rosenblat)	Graduated 2022
Erez Dor	
Thesis title: Gene-Environment effect for improving polygenic risk scores	Graduated 2022
Sharon Napadeski	
Thesis title: Time-series data analysis for outcome prediction during delivery	Graduated 2023
Lin Peretz	
Thesis title: Health consequences of people with a deviation of chronological age and lab-based physiological age	Graduated 2023
Liel Serfaty	
Thesis title: Effect of Mandatory Reporting on Surgeon and Hospital Surgery Performances	Graduated 2024
Matan Leventer	

Thesis title: Short- and long-term consequences of COVID-19 infection in young children	Graduated 2024
Ofir Ben Shoham	
Thesis title: Finetuning LLMs for health needs	Graduated 2024
Oriel Perets	
Thesis title: Improving Model's Performance Over Subpopulation Using Synthetic Data Generation in Clinical Datasets	Graduated 2023
Tomer Lavi	
Thesis title: Improving classification using intermediate clustering (co-advised with Bracha Shapira)	Graduated 2024
Shira Sadot	
Thesis title: Data-driven optimizing IVF protocols	Expected 2025
Natan Lubman	
Thesis title: Genomic associations of complex signal traits	Expected 2024
Charlotte Benadiba	
Thesis title: Survival-analysis-based genomic risk scores	Expected 2025
Yosi Meltser	
Thesis title: How does kidney stones forms - a metabolic perspective	Expected 2025
Elad Orpaz	
Thesis title: Alert system for epilepsy seizures using EEG (co-advised with Oren Shriky)	Expected 2026
Idan Cohen Zada	
Thesis title: Breast restoration complication prediction	
Recipient of Interdisciplinary fellowship (co-advised with Eitan Bachmat)	Expected 2026

Ph.D. Students

Kartheeswaran Thangathurai (co-advised with Eithan Rubin & Stass Engel)	Graduated 2024
Hadasa Kaufman (co-advised with Michal Linial)	Expected 2026
Oriel Perets	Expected 2028
Ofir Ben Shoham	Expected 2028
Tomer Lavi (co-advised with Bracha Shapira)	Expected 2028
Natan Lubman	Expected 2029

Supervision of visiting foreign students

Gerta Salilari, M.Sc. candidate

Thesis title: Mortality prediction with rejection

Visiting from Dept. of Information Engineering, Nagoya University 04/2021 – 08/2022

Undergraduate students - final project instructor

Year	Project title	Names	Comments
2021	MS prediction	Guy Zamostiano, Itai Katz & Shira Wertheim	Best project award
2021	Surgeon and hospital experiences affecting surgery outcomes	Daniel Gorsia, Hilla Peterfreund & Yaara Rumney	
2022	Standardizing Hadassah women ward to OMOP common data model	Almog Sarafian, Ayalong Azlan, Lahan Barmocha & Tom Duany	
2023	Modeling IVF treatment paths	Eyal Maor, Mazal Mahari, Nadav Arbel & Yuval Gorelik	
2023	Multi-center study without data sharing using OMOP CDM	Chen Vilinsky, Guy Kiperstik & Or Eitan	
2025	Genetic association of EEG Signals	Orel Jaraian, Selin Neomi Ivshin & Stav Farajun	

- **Awards, Citations, Honors, Fellowships**

(a) Honors, Citation Awards

Prize for Academic Excellence Hebrew University of Jerusalem	2008
Sudarsky Fellowship	2010
ECCB12 Conference Fellowship	2012
Noah Lichtenstein Prize	2013
Department selected monthly article (SMART) prize The Hebrew University of Jerusalem	2013
Charles Clore Foundation Fellowship for Ph.D. students	2014-2015
ISCB Travel Fellowship	2014
RECOMB Travel Fellowship	2015
Best poster McGil Anesthesia Bromage Research Day award for ”Revisiting Common Endometriosis Risk Factors with the UK Biobank”	2021

Best poster award: Brikman S., Liel S., Beiber A., Schlezinger N, Abu Hazira R. & Rappoport N., European Alliance of Associations for Rheumatology	2024

Students' Awards

Yarden Hochenberg (M.Sc. student) - Interdisciplinary fellowship	2020
Tomer Hermelin (M.Sc. student) - Excellence in Research	2022
Bar Ezra (M.Sc. student) – Engineering's Dean Excellence in Research	2023
Yarden Hochenberg (M.Sc. student) – Engineering's Dean Excellence in Research	2023
Ofir Ben-Shoham (M.Sc. student) – Engineering's Dean Excellence in Research	2023
Erez Dor (M.Sc. student) – Engineering's Dean Excellence in Research	2023
Liel Serfati (M.Sc. student) – Engineering's Dean Excellence in Research	2023
Ofir Ben Shoham (PhD student) – Engineering's Dean Excellence in Research	2024
Ofir Ben Shoham (PhD student) – Negev Zin Excellence in Research fellowship	2024

- **Scientific Publications**

  - H-index: ISI[1] 12, GS[2] 15.
  - Total number of citations: ISI 1,002 (997 without self-citations), GS 1,894.
  - The $*$ sign marks papers published after the beginning of the current review period. A paper marked with + is a conference paper that has not yet been published in a journal. The superscripts on co-author names indicate career stage where S=Student, PD=Post Doc., C=Collaborator, and PI=Principal Investigator. When available, citations are taken from Google Scholar (GS) and SCOPUS (ISI).

(c) Refereed chapters in collective volumes, conference proceedings

1. **N. Rappoport**[S] and M. Linial[PI], "Annotation of short proteins in the expanding universe of protein sequences from insects and crustaceans," in *Automated Function Prediction SIG 2012*, Oxford University Press, 2012 (citations ISI 0/GS 0; Not ranked)

2. D. Ofer[S], **N. Rappoport**[S], and M. Linial[PI], "The little known universe of short proteins in insects: A machine learning approach," *Short Views on Insect Genomics and Proteomics: Insect Genomics, Vol. 1*, pp. 177–202, 2015. DOI: 10.1007/978-3-319-24235-4_8 (citations ISI 3/GS 4)

   Contributions: leading, conceptualization, methodology, performing computational analysis, writing - original draft.

$*$ 3. B. H. Ezra[S], S. Havaldar[S], B. Glicksberg[C], and **N. Rappoport**[PI], "Multi-dimensional laboratory test score as a proxy for health," in *Challenges of Trustable AI and Added-Value on Health*, IOS Press, 2022, pp. 219–223. DOI: 10.3233/SHTI220441 (citations ISI 0/GS 0)

   Not ranked, acceptance rate=54%.
   Contributions: conceptualization, leading, methodology, supervision, writing - review & editing.

$*$ 4. L. Peretz[S] and **N. Rappoport**[PI], "Deviation of physiological from chronological age is associated with health," in *Challenges of Trustable AI and Added-Value on Health*, IOS Press, 2022, pp. 224–228. DOI: 10.3233/SHTI220442 (citations ISI 0/GS 7)

   Not ranked, acceptance rate=54%.

$*$ 5. T. Hermelin[S], P. Singer[C], and **N. Rappoport**[PI], "Improving prediction models' propriety in intensive-care unit, by enforcing an advance notice period," in *International Conference on Artificial Intelligence in Medicine*, Springer International Publishing Cham, 2022, pp. 167–177. DOI: 10.1007/978-3-031-09342-5_16 (citations ISI 0/GS 0)

   Rank=B, acceptance rate=15%.
   Contributions: conceptualization, methodology, supervision, writing - review & editing.

$*$ 6. G. Salillari[S] and **N. Rappoport**[PI], "Comparison of classification with reject option approaches on mimic-iv dataset," in *International Conference on Artificial Intelligence in Medicine*, Springer International Publishing Cham, 2022, pp. 210–219. DOI: 10.1007/978-3-031-09342-5_20 (citations ISI 0/GS 1)

[1] https://www.webofscience.com/wos/author/record/AAA-3769-2019
[2] https://scholar.google.com/citations?user=zL-QE0cAAAAJ&hl=en

Rank=B, acceptance rate=15%.
Contributions: conceptualization, methodology, supervision, writing - review & editing.

∗ 7. O. Perets[S] and **N. Rappoport**[PI], "DSF-GAN: Downstream feedback generative adversarial network," in *The Second Tiny Papers Track at ICLR 2024*, 2024. DOI: 10.48550/arXiv.2403.18267. [Online]. Available: https://openreview.net/forum?id=Vfp8jhwcCc (citations ISI 0/GS 0; **rank A***)

Acceptance rate=13%.
Contributions: conceptualization, leading, methodology, supervision, writing - review & editing.

(d) Refereed articles and refereed letters in scientific journals, running numbers

1. **N. Rappoport**[S], M. Fromer[S], R. Schweiger[S], and M. Linial[PI], "Pandora: Analysis of protein and peptide sets through the hierarchical integration of annotations," *Nucleic acids research*, vol. 38, no. suppl_2, W84–W89, 2010. DOI: 10.1093/nar/gkq320 (citations ISI 10/GS 10; IF=16.971, JR=8/295 Q1 in Biochemistry & Molecular Biology)

   top 5%
   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

2. **N. Rappoport**[S], S. Karsenty[S], A. Stern[S], N. Linial[PI], and M. Linial[PI], "Protonet 6.0: Organizing 10 million protein sequences in a compact hierarchical family tree," *Nucleic Acids Research*, vol. 40, no. D1, pp. D313–D320, 2012. DOI: 10.1093/nar/gkr1027 (citations ISI 36/GS 55; IF=16.971, JR=8/295 Q1 in Biochemistry & Molecular Biology)

   top 5%
   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

3. **N. Rappoport**[S] and M. Linial[PI], "Viral proteins acquired from a host converge to simplified domain architectures," *PLoS computational biology*, vol. 8, no. 2, e1002364, 2012. DOI: 10.1371/journal.pcbi.1002364 (citations ISI 25/GS 42; IF=5.379, JR=8/58 Q1 in Mathematical & Computational Biology)

   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

4. **N. Rappoport**[S] and M. Linial[PI], "Functional inference by protonet family tree: The uncharacterized proteome of daphnia pulex," in *BMC bioinformatics*, BioMed Central, vol. 14, 2013, pp. 1–15. DOI: 10.1186/1471-2105-14-S3-S11 (citations ISI 1/GS 1; IF=3.629, JR=16/58 Q2 in Mathematical & Computational Biology)

   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

5. **N. Rappoport**[S], N. Linial[PI], and M. Linial, "Protonet: Charting the expanding universe of protein sequences," *Nature Biotechnology*, vol. 31, no. 4, pp. 290–292, 2013. DOI: 10.1038/nbt.2553 (citations ISI 12/GS 12; IF= 54.908, JR=2/159 Q1 in Biotechnology & Applied Microbiology)

   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

6. S. Karsenty[PI], **N. Rappoport**[S], D. Ofer[S], A. Zair[S], and M. Linial[PI], "Neuropid: A classifier of neuropeptide precursors," *Nucleic acids research*, vol. 42, no. W1, W182–W186, 2014. DOI: 10.1093/nar/gku363 (citations ISI 9/GS 14; IF=16.971, JR=8/295 Q1 in Biochemistry & Molecular Biology)

   top 5%
   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

7. **N. Rappoport**[S], A. Stern[S], N. Linial[PI], and M. Linial[PI], "Entropy-driven partitioning of the hierarchical protein space," *Bioinformatics*, vol. 30, no. 17, pp. i624–i630, 2014. DOI: 10.1093/bioinformatics/btu478 (citations ISI 2/GS 2; IF=8.47, JR=3/58 Q1 in Mathematical & Computational Biology)

   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

8. **N. Rappoport**[S] and M. Linial[PI], "Protobug: Functional families from the complete proteomes of insects," *Database*, vol. 2015, bau122, 2015. DOI: 10.1093/database/bau122 (citations ISI 1/GS 1; IF=4.159, JR=14/58 Q1 in Mathematical & Computational Biology)

   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

9. **N. Rappoport**[S] and M. Linial[PI], "Trends in genome dynamics among major orders of insects revealed through variations in protein families," *BMC genomics*, vol. 16, pp. 1–15, 2015. DOI: 10.1186/s12864-015-1771-2 (citations ISI 5/GS 5; IF=4.478, JR=71/176 Q2 in Genetics & Heredity)

   Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

10. Y. Jiang[S], T. R. Oron[S], W. T. Clark[S], A. R. Bankapur[S], D. D'Andrea[S], R. Lepore[S], C. S. Funk[S], I. Kahanda[S], K. M. Verspoor[S], A. . Ben-Hur[S], **N. Rappoport**[PI], *et al.*, "An expanded evaluation of protein function prediction methods shows an improvement in accuracy," *Genome biology*, vol. 17, pp. 1–19, 2016. DOI: 10.1186/s13059-016-1037-6 (citations ISI 254/GS 439; IF=17.433, JR=6/176 Q1 in Genetics & Heredity)

    Top 5%
    Contributions: conceptualization, leading, methodology, performing computational analysis, validation, writing - review & editing.

11. M. Linial[PI], **N. Rappoport**[PD], and D. Ofer[S], "Overlooked short toxin-like proteins: A shortcut to drug design," *Toxins*, vol. 9, no. 11, p. 350, 2017. DOI: 10.3390/toxins9110350 (citations ISI 7/GS 13; IF=4.8, JR=21/93 Q1 in Toxicology)

    Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

12. **N. Rappoport**[PD], J. Toung[S], D. Hadley[S], R. J. Wong[C], K. Fujioka[C], J. Reuter[C], C. W. Abbott[C], S. Oh[C], D. Hu[PI], C. Eng, *et al.*, "A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth," *Scientific reports*, vol. 8, no. 1, p. 226, 2018. DOI: 10.1038/s41598-017-18246-5 (citations ISI 33/GS 51; IF= 5.134, JR=17/72 Q1 in Multidisciplinary Sciences)

Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

13. **N. Rappoport**[PD], H. Paik[PD], B. Oskotsky[C], R. Tor[C], E. Ziv[C], N. Zaitlen[PI], and A. J. Butte[PI], "Comparing ethnicity-specific reference intervals for clinical laboratory tests from ehr data," *The journal of applied laboratory medicine*, vol. 3, no. 3, pp. 366–377, 2018. DOI: 10.1373/jalm.2018.026492 (citations ISI 23/GS 33; IF=2, JR=18/32 Q3 in MEDICAL LABORATORY TECHNOLOGY)

    Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

14. N. Crawford[C], J. W. Oehlert[C], G. M. Shaw[C], D. K. Stevenson[C], **N. Rappaport**[PD], M. Sirota[C], S. A. Tishkoff[C], N. Sondheimer[C], *et al.*, "Divergent patterns of mitochondrial and nuclear ancestry are associated with the risk for preterm birth," *The Journal of pediatrics*, vol. 194, pp. 40–46, 2018. DOI: 10.1016/j.jpeds.2017.10.052 (citations ISI 16/GS 0; IF= 5.042, JR=11/129 Q1 in Pediatrics)

    Contributions: conceptualization, validation, methodology, performing computational analysis, writing - original draft.

∗ 15. B. Glicksberg[PD], J. Lee[C], **N. Rappoport**[PD], F. Liu[C], V. Pedoia[C], A. Butte[PI], and S. Majumdar[C], "Elucidating genetic associations that differentiate pain progression cluster groups in knee osteoarthritis patients," *Osteoarthritis and Cartilage*, vol. 27, S294–S295, 2019. DOI: 10.1016/j.joca.2019.02.684 (citations ISI 0/GS 1; IF= 7.063, JR=2/82 Q1 in Orthopedics)

    Top 5%
    Contributions: conceptualization, methodology, writing - original draft.

∗ 16. B. S. Glicksberg[PD], B. Oskotsky[C], P. M. Thangaraj[C], N. Giangreco[C], M. A. Badgeley[C], K. W. Johnson[C], D. Datta[C], V. A. Rudrapatna[C], **N. Rappoport**[PD], M. M. Shervey[PI], *et al.*, "Patientexplorer: An extensible application for dynamic visualization of patient clinical history from electronic health records in the omop common data model," *Bioinformatics*, vol. 35, no. 21, pp. 4515–4518, 2019. DOI: 10.1093/bioinformatics/btz409 (citations ISI 26/GS 47; IF=8.47, JR=3/58 Q1 in Mathematical & Computational Biology)

    Contributions: conceptualization, methodology, writing - review & editing.

∗ 17. H. Paik[PD], M. J. Kan[C], **N. Rappoport**[PD], D. Hadley[PD], M. Sirota[C], B. Chen[PD], U. Manber[C], S. B. Cho[C], and A. J. Butte[PI], "Tracing diagnosis trajectories over millions of patients reveal an unexpected risk in schizophrenia," *Scientific data*, vol. 6, no. 1, p. 201, 2019. DOI: 10.1038/s41597-019-0220-5 (citations ISI 8/GS 12; IF= 9.051, JR=11/72 Q1 in Multidisciplinary Sciences)

    Contributions: conceptualization, methodology, writing - review & editing.

∗ 18. D. Klinger[S], I. Blass[S], **N. Rappoport**[PI], and M. Linial[C], "Significantly improved covid-19 outcomes in countries with higher bcg vaccination coverage: A multivariable analysis," *Vaccines*, vol. 8, no. 3, p. 378, 2020. DOI: 10.3390/vaccines8030378 (citations ISI 50/GS 101; IF= 5.513, JR=63/140 Q2 in Medicine, Research & Experimental)

    Contributions: conceptualization, leading, methodology, performing computational analysis, supervision, writing - original draft.

∗ 19. N. Ghantous[S], M. Heshin-Bekenstein[C], K. Dequattro[C], Y. Lakovsky[C], A. M. Hendel[C], **N. Rappoport**[PI], Y. B. Aviel[S], I. Tirosh[S], L. Harel[S], P. F. Weiss[C], *et al.*, "Do geography and ethnicity play a role in juvenile spondyloarthritis? a multi-center binational retrospective study," *Pediatric Rheumatology*, vol. 19, pp. 1–8, 2021. DOI: 10.1186/s12969-020-00489-8 (citations ISI 2/GS 3; IF= 3.497, JR=34/129 Q2 in Pediatrics)

Contributions: conceptualization, leading, methodology, performing computational analysis, supervision, writing - original draft.

∗ 20. B. Sheppard[S], **N. Rappoport**[PD], P.-R. Loh[C], S. J. Sanders[S], N. Zaitlen[PI], and A. Dahl[PD], "A model and test for coordinated polygenic epistasis in complex traits," *Proceedings of the National Academy of Sciences*, vol. 118, no. 15, e1922305118, 2021. DOI: 10.1073/pnas.192230511 (citations ISI 10/GS 20; IF= 12.291, JR=8/72 Q1 in Multidisciplinary Sciences)

(*PNAS*)
Contributions: conceptualization, leading, methodology, performing computational analysis, writing - original draft.

∗ 21. D. Klinger[S], B. L. Hill[S], N. Barda[S], E. Halperin[C], O. N. Gofrit[C], C. L. Greenblatt[C], **N. Rappoport**[PI], M. Linial[C], and H. Bercovier[C], "Bladder cancer immunotherapy by bcg is associated with a significantly reduced risk of alzheimer's disease and parkinson's disease," *Vaccines*, vol. 9, no. 5, p. 491, 2021. DOI: 10.3390/vaccines9050491 (citations ISI 41/GS 59; IF= 5.513, JR=63/140 Q2 in Medicine, Research & Experimental)

Contributions: conceptualization, methodology, performing computational analysis, supervision, writing - review & editing.

∗ 22. N. Brandes[S], D. Ofer[S], Y. Peleg[C], **N. Rappoport**[PI], and M. Linial[C], "Proteinbert: A universal deep-learning model of protein sequence and function," *Bioinformatics*, 2022. DOI: 10.1093/bioinformatics/btac020 (citations ISI 239/GS 582; IF= 6.931, JR=4/68 Q1 in MATHEMATICAL & COMPUTATIONAL BIOLOGY)

Contributions: conceptualization, leading, methodology, performing computational analysis, supervision, writing - original draft, writing - review & editing.

∗ 23. R. K. Perez[S], M. G. Gordon[C], M. Subramaniam[S], M. C. Kim[S], G. C. Hartoularos[S], S. Targ[C], Y. Sun[C], A. Ogorodnikov[C], R. Bueno[C], A. Lu[C], M. Thompson[C], **N. Rappoport**[PI], A. Dahl[C], C. M. Lanata[C], M. Matloubian, L. Maliskova[C], S. S. Kwek[C], T. Li[C], M. Slyper[C], J. Waldman[C], D. Dionne[C], O. Rozenblatt-Rosen[C], L. Fong[C], M. Dall'Era[C], B. Balliu[C], A. Regev[C], J. Yazdany[C], L. A. Criswell[C], N. Zaitlen[PI], and C. J. Ye[PI], "Single-cell rna-seq reveals cell type–specific molecular and genetic associations to lupus," *Science*, vol. 376, no. 6589, eabf1970, 2022, ISSN: 1095-9203 (Electronic) 0036-8075 (Print) 0036-8075 (Linking). DOI: 10.1126/science.abf1970. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/35389781 (citations ISI 174/GS 0; IF= 63.832, JR= 2/135 Q1 in MULTIDISCIPLINARY SCIENCES)

Top 5% (*Science*)
Contributions: methodology, performing computational analysis, writing - review & editing.

∗ 24. **N. Rappoport**[PI], G. Aviel[S], D. M. Shahian[C], A. Korach[S], S. Carmi[C], J. F. Keaney Jr[C], and O. M. Shapira[C], "Resternotomy coronary artery bypass 1999-2018: Insights from the society of thoracic surgeons adult cardiac surgery database," *The Annals of Thoracic Surgery*, vol. 115, no. 1, pp. 62–71, 2023. DOI: 10.1016/j.athoracsur.2022.02.087 (citations ISI 6/GS 7; IF=5.113, JR=28/213 Q1 in Surgery)

Contributions: conceptualization, leading, methodology, performing computational analysis, supervision, writing - original draft.

∗ 25. I. Blass[S], T. Sahar[S], A. Shraibman[S], D. Ofer[S], **N. Rappoport**[PI], and M. Linial[C], "Revisiting the risk factors for endometriosis: A machine learning approach," *Journal of Personalized Medicine*, vol. 12, no. 7, p. 1114, 2022. DOI: https://doi.org/10.3390/jpm12071114 (citations ISI 12/GS 19; IF=3.508, JR= 57/329 Q1 in MEDICINE, GENERAL & INTERNAL)

Contributions: conceptualization, methodology, supervision, writing - review & editing.

∗ 26. B. Abbou[C], O. Tal[C], G. Frenkel[S], R. Rubin[C], and **N. Rappoport**[PI], "Optimizing operation room utilization—a prediction model," *Big Data and Cognitive Computing*, vol. 6, no. 3, p. 76, 2022. DOI: 10.3390/bdcc6030076 (citations ISI 6/GS 17; IF=3.7, JR=87/190 Q1 in COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE)

Contributions: conceptualization, leading, methodology, supervision, writing - original draft, writing - review & editing.

∗ 27. H. Paik[PI], J. Lee[S], C. S. Jeong[S], J. S. Park[C], J. H. Lee[C], **N. Rappoport**[PI], Y. Kim[C], H. Y. Sohn[C], C. Jo[C], J. Kim[C], and S. B. Cho[C], "Identification of a pleiotropic effect of adipoq on cardiac dysfunction and alzheimer's disease based on genetic evidence and health care records," *Translational Psychiatry*, vol. 12, no. 1, p. 389, 2022. DOI: 10.1038/s41398-022-02144-0 (citations ISI 3/GS 3; IF=7.989, JR=25/258 Q1 in PSYCHIATRY)

Contributions: conceptualization, methodology, performing computational analysis, writing - review & editing.

∗ 28. **N. Rappoport**[PI], D. M. Shahian[C], N. Galai[C], G. Aviel[S], J. F. Keaney Jr[C], and O. M. Shapira[PI], "Volume-outcome relationship of resternotomy coronary artery bypass grafting," *The Annals of Thoracic Surgery*, vol. 116, no. 2, pp. 287–295, 2023. DOI: 10.1016/j.athoracsur.2022.09.049 (citations ISI 3/GS 5; IF=5.113, JR=14/99 Q1 in RESPIRATORY SYSTEM)

Contributions: conceptualization, leading, methodology, performing computational analysis, supervision, writing - original draft.

∗ 29. A. Ostropolets[S], Y. Albogami[C], M. Conover[C], J. M. Banda[C], W. A. Baumgartner[C], C. Blacketer[C], P. Desai[C], S. L. DuVall[C], S. Fortin[C], J. P. Gilbert[C], A. Golozar[C], J. Ide[C], A. S. Kanter[C], D. M. Kern[C], C. Kim[C], L. Y. H. Lai[C], C. Li[C], F. Liu[C], K. E. Lynch[C], E. Minty[C], M. I. Neves[C], D. Q. Ng[C], T. Obene[C], V. Pera[C], N. Pratt[C], G. Rao, **N. Rappoport**[PI], I. Reinecke[C], P. Saroufim[C], A. Shoaibi[C], K. Simon[C], M. A. Suchard[C], J. N. Swerdel[C], E. A. Voss[C], J. Weaver[C], L. Zhang[C], G. Hripcsak[C], and P. B. Ryan[PI], "Reproducible variability: Assessing investigator discordance across 9 research teams attempting to reproduce the same observational study," *Journal of the American Medical Informatics Association*, vol. 30, no. 5, pp. 859–868, 2023. DOI: 10.1093/jamia/ocad009 (citations ISI 4/GS 10; IF=7.942, JR=18/246 Q1 in Computer Science, Information Systems)

Contributions: conceptualization, methodology, performing computational analysis, validation, writing - review & editing.

∗ 30. O. B. Shoham[S] and **N. Rappoport**[PI], "Federated learning of medical concepts embedding using behrt," *JAMIA open*, vol. 7, no. 4, ooae110, 2024. DOI: 10.1093/jamiaopen/ooae110 (citations ISI 0/GS 10; IF=2.5, JR=69/174 Q2 in HEALTH CARE SCIENCES & SERVICES)

∗ 31. N. A. Walton[S], R. Nagarajan[S], C. Wang[S], M. Sincan[S], R. R. Freimuth[S], D. B. Everman[C], D. C. Walton[C], S. P. McGrath[C], D. J. Lemas[C], P. V. Benos[C], A. V. Alekseyenko[C], Q. Song[C], E. Gamsiz

Uzun[C], C. O. Taylor[C], A. Uzun, T. N. Person[C], **N. Rappoport**[PI], Z. Zhao[C], and M. S. Williams[PI], "Enabling the clinical application of artificial intelligence in genomics: A perspective of the amia genomics and translational bioinformatics workgroup," *Journal of the American Medical Informatics Association*, vol. 31, no. 2, pp. 536–541, 2024. DOI: 10.1093/jamia/ocad211 (citations ISI 4/GS 6; IF=7.942, JR=18/246 Q1 in Computer Science, Information Systems)

Contributions: conceptualization, methodology, writing - original draft, writing - review & editing.

∗ 32. M. Elkan[S], N. Kofman[S], S. Minha[C], **N. Rappoport**[PI], R. Zaidenstein[C], and R. Koren[PI], "Does the "obesity paradox" have an expiration date? a retrospective cohort study," *Journal of Clinical Medicine*, vol. 12, no. 21, p. 6765, 2023. DOI: https://doi.org/10.3390/jcm12216765 (citations ISI 0/GS 3; IF=3.9, JR=53/327 Q1 in Medicine, General & Internal)

Contributions: conceptualization, leading, methodology, supervision, writing - review & editing.

∗ 33. D. Azriel[C], Y. Rinott[C], O. Tal[C], B. Abbou[C], and **N. Rappoport**[PI], "Surgery duration prediction using multitask feature selection," *IEEE Journal of Biomedical and Health Informatics*, 2024. DOI: 10.1109/JBHI.2024.3374783 (citations ISI 0/GS 0; IF=7.7, JR=21/251 Q1 in Computer Science, Information Systems)

Contributions: conceptualization, methodology, supervision, writing - review & editing.

∗ 34. S. Brikman[S], L. Serfaty[S], R. Abuhasira[C], N. Schlesinger[C], A. Bieber[C], and **N. Rappoport**[PI], "A machine learning-based prediction model for gout in hyperuricemics: A nationwide cohort study," *Rheumatology*, keae273, 2024. DOI: 10.1093/rheumatology/keae273 (citations ISI 1/GS 1; IF=4.7, JR=9/57 Q1 in RHEUMATOLOGY)

∗ 35. O. Ben Shoham[S] and **N. Rappoport**[PI], "Cpllm: Clinical prediction with large language models," *PLOS Digit Health*, vol. 3, no. 12, e0000680, 2024, ISSN: 2767-3170 (Electronic) 2767-3170 (Linking). DOI: 10.1371/journal.pdig.0000680. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/39642102 (citations ISI 0/GS 20; Not ranked yet)

∗ 36. O. Ben Shoham[S] and **N. Rappoport**[PI], "Medconceptsqa: Open source medical concepts qa benchmark," *Computers in Biology and Medicine*, vol. 182, 2024. DOI: 10.1016/j.compbiomed.2024.109089. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482524011740 (citations ISI 0/GS 2; IF=7, JR=7/109 Q1 in Biology)

(e) Unrefereed professional articles and publications

(f) Articles under review

∗ 1. O. Perets[S] and N. Rappoport[C], "Subpopulation-specific synthetic ehr for better mortality prediction," *arXiv preprint arXiv:2305.16363*, 2023. DOI: 10.48550/arXiv.2305.16363 (citations ISI 0/GS 1;)

∗ 2. O. Perets[S], E. Stagno[C], E. B. Yehuda[C], M. McNichol[C], L. A. Celi[C], **N. Rappoport**[PI], and M. Dorotic[C], "Inherent bias in electronic health records: A scoping review of sources of bias," *medRxiv*, 2024. DOI: 10.1101/2024.04.09.24305594. [Online]. Available: https://www.medrxiv.org/content/10.1101/2024.04.09.24305594v1 (citations ISI 0/GS 1;)

Submitted to the Transactions on Intelligent Systems and Technology

∗ 3. H. Kaufman[S], Y. Hochenberg[S], M. Linial[C], and **N. Rappoport**[PI], "Drop-deep: Dimensionality reduction for polygenic risk score using deep learning approach," *medRxiv*, pp. 2024–05, 2024. DOI: doi.org/10.1101/2024.05.01.24306609. [Online]. Available: https://www.medrxiv.org/content/10.1101/2024.05.01.24306609v1 (citations ISI 0/GS 0;)

Submitted to RECOMB

∗ 4. M. Pineda-Moncusí[S], A. Rekkas[C], Á. M. Pérez[C], *et al.*, "Trends of use of drugs with suggested shortages and their alternatives across 52 real world data sources and 18 countries in europe and north america," *medRxiv*, pp. 2024–08, 2024. DOI: 10.1101/2024.08.28.24312695 (citations ISI 0/GS 0;)

Submitted to JAMA

- **Lectures and Presentations at Meetings and Invited Seminars**

(a) Invited plenary lectures at conferences/meetings - (running numbers)

ECCB 2010 2010
Gent, Belgium
Poster: Functional Inference for Experimental Proteomics: The PANDORA (Annotations graph) and ProtoNet (Family Tree) Resources

ISMB / ECCB 2011 2011
Vienna, Austria
Poster: Viral proteins acquired from a host converge to simplified domain architectures

AFP / CAFA Satellite, ISMB / ECCB 2011 2011
Vienna, Austria
Poster: The Daphnia Uncharacterized Proteome of Short Sequences Is Revealed via the ProtoNet Classification Tree

InCoB/ISCB-Asia 2011
MalaysiaRappoport
N. & Linial M., 2011, "ProtoNet: Automatic Organization of the 10 million protein sequences into an evolutionary driven family tree"

ECCB 2012 2012
Basel, Switzerland
Highlight presentation: Viral-host coevolution: Playing 'seek and hide'

ECCB 2012 2012
Basel, Switzerland
Poster: 18 Insects whole proteomes tree construction identify proteins families special- ized in different Insects branches

ISMB 2012 2012
Long Beach, CA USA
Highlight presentation: Viral-host coevolution: Playing 'seek and hide'

ISMB 2012 2012
Long Beach, CA USA
Poster: A tree construction of all 18 insects complete proteomes reveals extensive expansion and shrinking of proteins families

IBS 2012 2012
Jerusalem, Israel
Poster: Functional Inference by ProtoNet Family Tree: The Uncharacterized proteome of Daphnia pulex

ISMB/ECCB 2013 2013
Berlin, Germany
Poster: Representation of Protein Complexes as Multilayer Graphs

RECOMB 2013 2013
Beijing, China
Poster: Representation of Protein Complexes as Multilayer Graphs

ISMB 2014 2014
Boston, MA US
Poster: ProtoBug: Automatic Classification of Insects? Complete Proteomes

ECCB 2014 2014
Strasbourg, France
Lecture: Entropy driven partitioning of the hierarchical protein space Poster: Comparative Analysis of Insects' Complete Proteomes

ISMB/ECCB 2015 2015
Dublin, Ireland
Poster: Entropy-driven partitioning of the hierarchical protein space

RECOMB 2015 2015
Warsaw, Poland
Poster: Entropy-driven partitioning of the hierarchical protein space

RECOMB 2016 2016
Los Angeles, California
Poster: Predict demographic features from lab tests

Transatlantic Data Science Workshop 2016
NIH, Bethesda, MD

Pediatric academic societies meeting 2017 2017
San Francisco, California
Poster: Genome-wide Association Study Identifies Ancestry Specific Variants Associated with Spontaneous Preterm Birth

Health Disparities Research Symposium, 2017 2017
San Francisco, California
Lecture: Creating ethnicity-specific reference intervals for diagnostic tests from I data

Biomedical big data science conference, 2018 2018
Israel
Poster: Creating ethnicity-specific reference intervals for lab tests from I data

Frontiers in Computational Healthcare 2019
Israel
Lecture: A system for prediction of bleeding volumes during surgery from medical records

Heshin-Bekenstein, M., Ghantous, N., Tirosh, I., Butbul, Y., Harel, L., Lakovsky, Y., Weiss, P., Rappoport, N., Dequattro, K., von Scheven, E. and Gensler, L.S., 2019. 2019
Comparison of clinical and demographic features of juvenile spondyloarthritis between israeli and us children.

Israeli Human Population Genetics Meeting 2019
Israel
Rappoport N. Zaitlen N, & Butte A., Leveraging electronic health records to define ethnicity-specific reference ranges for common lab tests"

ILANIT2020 2020
Eilat, ISRAEL
Rappoport N., Zaitlen N., Hofer I., "Using electronic health records to estimate risk of bleeding during surgery"

OHDSI2020 2020
Rappoport N. "Changing Impact Of Reoperative Coronary Artery Bypass Grafting On Clinical Outcomes"

ECTRIMS 2021
N. Haham, L. Brill, N. Rappoport, O. Zveik, A. Rechtman, A. Vaknin-Dembinsky
Using circulatory biomarkers to predict disease progression in multiple sclerosis (MS)"

ILAMI2021 2021
Abbou B & Rappoport N.
"Surgery duration prediction using big data for increasing OR utility"

STS2022 2022
Shapira O & Rappoport N.
"Volume-Outcome Relationship of Coronary Artery Bypass Grafting Performed after a Prior Sternotomy – Insights from the Society of Thoracic Surgeons National Database"

Medical Informatics Europe 2022 (MIE22) 2022
Ezra, Bar H., et al. "Multi-Dimensional Laboratory Test Score as a Proxy for Health."

Medical Informatics Europe 2022 (MIE22) 2022
Peretz, Lin, and Nadav Rappoport
"Deviation of Physiological from Chronological Age Is Associated with Health."

International Conference on Artificial Intelligence in Medicine 2022
Hermelin, Tomer, Pierre Singer, and Nadav Rappoport
Improving Prediction Models' Propriety in Intensive-Care Unit, by Enforcing an Advance Notice Period."

International Conference on Artificial Intelligence in Medicine 2022
Salillari, Gerta, and Nadav Rappoport. "Comparison of Classification with Reject Option Approaches on MIMIC-IV Dataset."

ISMB 2022 2022
Brandes N., Ofer D., Peleg Y., Rappoport N., Linial M, ProteinBERT: a universal deep-learning model of protein sequence and function

IASP world congress on pain 2022
Blass, I., Sahar, T., Shraibman, A., Ofer, D., Rappoport, N., & Linial, M. "Revisiting Common Endometriosis Risk Factors with the UK Biobank"

ISMB 2022
Blass, I., Sahar, T., Shraibman, A., Ofer, D., Rappoport, N., & Linial, M. "Predictive Model for Endometriosis with Clinical, Lifestyle and Genetic Information"

AI and Big Data in Medical and Public Health Sciences Conference 2022
Israel
Hochenberg Y. & Rappoport N. "Machine Learning approach for genetic-based risk scores"

AI and Big Data in Medical and Public Health Sciences Conference 2022
Israel
Dor E. & Rappoport N. "Model development for estimating Gene-Environment interactions"

AI and Big Data in Medical and Public Health Sciences Conference 2022
Israel
Thangathurai K., Engel S., Rappoport N. & Rubin E. "A general framework for incorporating agnostic prior knowledge for an improved feature selection approach"

The 14th, NIHP meeting, Israel, March 29th, 2022 2022
Rappoport N., "Surgery duration prediction"

IEEE-EMBS 2022
Perets O. & Rappoport N., "Improving Model's Performance Over Subpopulation Using Synthetic Data Generation in Clinical Datasets"

IDSI 2023
Ein-Gedi, Israel
Salillari G. & Rappoport N. "Classification with Reject Option Approaches on MIMIC-IV Dataset"

IDSI 2023
Ein-Gedi, Israel
Abbou B., & Rappoprt N., "Optimizing Operation Room Utilization—A Multi-Centers Prediction"

The 5th Winter School in Computer Science and Engineering AI for Better Medicine 2023
Jerusalem, Israel
Salillari G. & Rappaport N, "Classification with Reject Option Approaches on MIMIC-IV Dataset"

ILANIT 2023
Eilat, Israel
Peretz L., Bar E. & Rappoport N. "What can we tell from routine laboratory test results?"

25th European Congress of Endocrinology 2023
Matan Elkan, Natalia Kofman, Nadav Rappoport, Ronit Zaidenstein, Ronit Koren, "Does the 'obesity paradox' have an expiration date? A bigdata retrospective cohort"

ISPOR Europe 2023
C. Read, R. Claire, J. Elvidge, J. Dietz, T. Debray, M.A. Mayer, J.M. Ramirez Anguita, D. Dawoud, N. Rappoport, "Decision Making and the Questions Raised When Real World Evidence From a Federated Data Network Conflicts With RCT Data: A Use Case on COVID-19 Treatments", Value in Health, Volume 26, Issue 12, Supplement

Medicine in the AI Era 2023
HIT, May 15th, 2023
Nadav Rappoport, "Exploitation of Language models to decode protein structure"

3rd Mini-conference on AI and the Israeli Gov 2023
Jerusalem, May 29th, 2023
Kineret Rosenstein & Nadav Rappoport, "Kinneret Project - Data Aggregation for Research and AI for Industry and Academia, Hospital Division"

The 15th meeting of the Israel National Institute For Health Policy Research 2023
Tel-Aviv
Rappoport N., "Hierarchy-based model prediction of surgery duration", May 31st

2nd annual Data for Health (D4H) conference September 11th, 2023 2023
Rappoport N. "What can we tell more from routine laboratory test results?"

Digital Health 2023
Lago, Israel, March 29th, 2023, Israel
Rappoport N. "The Future of the Digital Health"

European Society of Human Genetics conference (ESHG 2023) 2023
Glasgow, United Kingdom June 2023
"Using PDR as prior knowledge to discover new associations to Parkinson's disease"

ISMB 2023
Poster: Kaufman H., Rappoport N., Hochenber Y., and Linial M., "Improving Polygenic Risk Score Prediction by Phenotype-Agnostic Dimensionality Reduction"

ASHG, 2023 2023
Washington, USA
Dor E. & Rappoport N., ": A new genetic risk score model for incorporating Gene-Environment interactions"

ASHG, 2023 2023
Washington, USA
Rappoport N. et al. "Selecting Covariates for Genome-Wide Association Studies"

SMFM 2024 Pregnancy Meeting 2024
https://cdmcd.co/7brgxX
Lee Reicher, Emmanuel Attali, Sharon Napadenski, Nadav Rappoport, Yariv Yogev "Personalized Machine Learning Model for Predicting Time to Delivery in Active Labor"

SMFM 2024 Pregnancy Meeting 2024
https://cdmcd.co/6vZzQ4
Lee Reicher, Sharon Napadenski, Emmanuel Attali, Nadav Rappoport, Yariv Yogev "Personalized Machine Learning Model for Predicting Successful Vaginal Delivery during the Active Phase of Labor"

ILAMI 2024
Tel-Aviv, Israel
Serfati L., & Rappoport N.

European Alliance of Associations for Rheumatology (EULAR) 2024
Brikman S., Liel S., Beiber A., Schlezinger N, Abu Hazira R. & Rappoport N.

EULAR 2024 European Congress of Rheumatology 2024
Best Poster Award
Brikman S, Serfaty L, Abuhasira R, et al OP0003 A MACHINE LEARNING-BASED PREDICTION MODEL FOR GOUT IN HYPERURICEMICS – A NATIONWIDE COHORT STUDY Annals of the Rheumatic Diseases 2024;83:7.

IBS2024 2024
Israel
Ivited lecture: Dimensionality Reduction for Polygenic-Risk-Score Using Deep Learning Approach

IDSI2024 2024
Israel
Perets et al. ”DSF-GAN: Downstream Feedback Generative Adversarial Network”

IDSI2024 2024
Israel
Lubman et al. Genome-Wide-Association Study of a Signal-Based Trait Using Latent Representation

IDSI2024 2024
Israel
Poster: Thangathurai Kartheeswaran ”A Framework for Features Selection in Fat Datasets Using External Knowledge Integration”

IDSI2024 2024
Israel
Poster & Flash talk: Ben Shoham O. et al

AIME2024 2024
Poster: ”Subpopulation-Specific Synthetic EHR Can Increase Mortality Prediction Performance”

AIME2024 2024
Poster: ”A Framework for Features Selection in Fat Datasets Using External Knowledge Integration”

AIME2024 2024
Poster: ”Federated Learning of Medical Concepts Embedding using BEHRT”

PHSS Workshop-XIII 2024
Lecture: Subpopulation-Specific Synthetic EHR Can Increase Mortality Prediction Performance

AI and Precision Medicine Scientific Workshop 2024
Lecture: CPLLM: Clinical Prediction with Large Language Models

AI for Reliable and Equitable Real World Evidence Generation in Medicine” Workshop 2024
Federated Learning of Medical Concepts Embedding using BEHRT

Altera Digital Health (dbMotion) 2024
Israel
Ivited seminar

ACR24 2024
Invited talk

ACR24 2024
Poster: “The Incidence of Gout Among Individuals with Hyperuricemia over Time – an Insight from a Nationwide Cohort Study”

ACR24 2024
Poster: "Exploring Multi Factorial Model for the Prediction of Gout in Patients with Hyperuricemia"

(b) Seminar presentations at universities and institutions

Tel-Aviv University, Israel 2015
"Automatic annotation of proteomes protein families using ProtoNet platform"

UCSF, USA 2015
"Proteins' functional annotation prediction using big data"

The Hebrew University of Jerusalem 2015
"Automatic annotation of proteomes protein families using ProtoNet platform"

Mt. Sinai, NY, USA 2015
"Functional inference by ProtoNet family tree: the uncharacterized proteome of daphnia pulex"

UCSF, USA 2016
"Harnessing Big Data for Laboratory test reference ranges"

Stanford, USA 2016
"Genomic Wide Association Study of Preterm Birth"

UCSF, USA 2017
"Comparing ethnicity-specific reference intervals for clinical laboratory tests from EHR data"

Hadassah, Israel 2017
"A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth"

Stanford, USA 2017
"A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth"

Sema4, MS, USA 2018
"Harnessing Big Data for Laboratory test reference ranges"

UCLA, USA 2018
"Identifying patients with high risk for Hemorrhage"

Sema4, MS, USA 2019
"Harnessing Big Data for Laboratory test reference ranges"

Ariel University, Israel 2019
"Big Biomedical data for healthcare improvement and Precise Medicine"

Bar-Ilan University, Israel 2019
"Big Biomedical data for healthcare improvement and Precise Medicine"

KI Institute, Israel 2019
"Big Biomedical data for healthcare improvement and Precise Medicine"

KSM, Israel 2019
"Big Biomedical data for healthcare improvement and Precise Medicine"

IBM Research, Israel 2019
"Big Biomedical data for healthcare improvement and Precise Medicine"

Israeli Ministry of Health 2020
"Prediction analytics for advancing health care"

Fertility AI Academy 2022
"Machine Learning and Fertility"

Shamir Medical Center 2022
"Surgery duration prediction using multicenter health records"

SISE, BGU 2023
"What can we tell from routine laboratory test results?"

Haemek Medical Center 2023
"Hyperuricemia and Gout Clustering and prediction models"

Hillel Yafe Research Seminar 2023
Retrospective data studies in Kineret, June 19th, 2023

Utah Medical Center, Rheumatology Seminar 2024
An Eye of AI: What can we learn from routine laboratory tests?

OHDSI IL, KI Institute 2024
Data studies in Kineret Platform

(c) Other talks and interviews and activities

2017-18 Catalyst Awards program consulting, University of California, San Francisco

2020 Radio interviewed by Dudu Erez, BCG-Covid-19

2020 IDF – Hackathon – Mentor

2020 Mentor in UK BioBank-Hackathon, The Hebrew University of Jerusalem

2020 Mentor in Medical Hackathon, Medical Corps, Israel Defense Force

2021 Podcast interviewed, "Ma Shekoreh Mahar" with Rami Shani

2021 Mentor in UK BioBank-Hackathon, The Hebrew University of Jerusalem

2022 Radio Interview by Sharon Kantor, "Three Who Know" 10/2022

2023 MeDS Meetup "Swimming in the Nation Medical Data Lake - Kineret Big Data Platform", 15/5/2023

- **Patents**
- **Grants**

Highlights

- Acquired funding for approximately $1,000,000$ NIS since joining BGU.

Competitive Research Grants ($1,000,000$ NIS)

Competitive Agencies as recognized by Planning and Budgeting Committee (PBC) – VATAT

∗ ISF IPMP — The Israeli Science Foundation (ISF) 2023–2024
Subject: What makes a stone? Unraveling the molecular mechanism driving kidney stone formation
Grantees: **Nadav Rappoport**, Gur D., Duvdevani M., & Zilberman D.)
Total amount: 2600K NIS

∗ Cheif Scietist of Israeli Ministry of Health 2022–2023
Subject: Deviation of Biological Age From Chronological Age as Health Biomarker
Grantees: **Nadav Rappoport**
Total amount: 25K NIS

∗ Israel Ministry of Science and Technology (MOST) 2024-2026
Subject: The impact of SGLT2 inhibitors on acid-base balance, pH, and ketones in urine and their relationship to kidney function, cardiovascular events and mortality - A retrospective study
Grantees: **Nadav Rappoport**, & E. Krutzel
Total amount: 250K NIS

Other Competitive Agencies

∗ Horizon 2022–2023
Subject: OMOP Standardizing OBGYN data
Grantees: **Nadav Rappoport**, Sompolinksy Y., Liphsitz M.
Total amount: 70K Euro

External Grants ($155,000$ USD)

∗ Milgrom foundation 2020–2021
Subject: Predicting Medical Care Burden from Early Life Records
Grantees: **Nadav Rappoport**
Total amount: 10K USD

* EHDEN 2019–2021
  Subject: Standardizing Galilee Medical Center data for retrospective studies
  Grantees: **Nadav Rappoport**, Israeli MOH, & Galilee Medical Center
  Total amount: 115K USD

* Macabbi 2022–2023
  Subject: COVID-19 effects in newborns
  Grantees: **Nadav Rappoport** & Michael Davidovich
  Total amount: 30K USD

Internal (BGU) Grants (100,000 NIS)

* Data Science center 2020–2021
  Subject: Estimating Biological Age from routing laboratory test results
  Grantees: **Nadav Rappoport**
  Total amount: 10K USD

* Data Science center 2022–2023
  Subject: Estimating Genetic interaction in large scale dataset
  Grantees: **Nadav Rappoport** & Mark Last
  Total amount: 41K USD

* Google 2023–2024
  Subject: GCP for DS
  Grantees: **Nadav Rappoport**
  Total amount: 5K USD

* Data Science center 2023–2024
  Subject: Using prior knowledge to reduce the search space for finding novel associations for Neurodegenerative disorders
  Grantees: **Nadav Rappoport** & Eitan Rubin
  Total amount: 41K USD

* SimREC 2023–2024
  Subject: Empathy estimation from nurses videos
  Grantees: **Nadav Rappoport**, Miky Fier, & Odeya Cohen)
  Total amount: 9K NIS

- **Present Academic Activities**

(a) Research in progress

Subject: Long-term consequences of COVID-19 in young children 2024
Participants: Matan Leventer (M.Sc. student)

Subject: Optimizing IVF protocols 2025
Participants: Shira Sadot (M.Sc. candidate)

Subject: LLMs models in the clinic 2024
Participants: Ofir Ben-Shoham (M.Sc. student)

Subject: Genetic risk of complex cardiac traits **Participants:** Natan Lubman (M.Sc. student)	2024
Subject: Dimensionality reduction-based Polygenic Risk Scores **Participants:** Hadasa Kaufman (Ph.D. student)	2024
Subject: GAN models in clinical prediction models **Participants:** Oriel Perets (Ph.D. candidate)	2024
Subject: Classification with rejection **Participants:** Tomer Lavi (M.Sc. student)	2024
Subject: Survival model-based Polygenic Risk scores **Participants:** Charlotte Benadiba (M.Sc. student)	2024
Subject: Spatial omics data of kidney stones **Participants:** Yosef Meltser (M.Sc. candidate)	2024
Subject: Patterns recoginition in EEGs **Participants:** Elad Orpaz (M.Sc. student)	2026

- **Additional Information**

(a) Personal Statement

- **Synopsis of Research**

My main research interests are in the fields of Machine Learning, Deep Learning, Big Data, Health Care, Clinical Data, EHR (electronic health record), Biomolecular data, Population Genetics, Prediction Models, and Precision Medicine. I'm interested in advancing health and healthcare in data-driven approaches. (Citations refer to the publications listed in the CV).

Recent advances in big-data analyses, are increasing the availability of large-scale clinical datasets, and the rapid cost reduction in molecular-genetic assays has transformed medical genomics research. Discoveries in these rich data sets are rate-limited by approaches to integrate across modalities. Modern computational methods will improve the healthcare system and deliver more precise medicine. I want to take advantage of recent achievements in big-data collection, analysis, algorithms, computation resources, and storage to improve the health system. Harnessing the huge amount of clinical and bio-molecular data for advancing health forward is not trivial. There specific challenges and filed I'm interested in.

(a) Laboratory test results as proxy and predictors for health status

Clinical laboratory tests contribute to medical diagnoses and interventional decisions. They also provide insight into physiological states that are not directly observable. They are widely used, as many laboratory tests are inexpensive, easy to extract, and have standard protocols. However, differences in population demographics, geography, and laboratory instruments can alter the distribution of test results. Therefore, each clinical laboratory is expected to define reference intervals for each laboratory test. However, I wanted to show that having a single reference interval per laboratory is not ideal. So first, I wanted to define a more precise reference interval, in a data-driven way. We showed that this can be achieved using pre-collected samples. Moreover, I showed that population-specific reference intervals are more predictive of future outcomes and, therefore, more useful [j14]. Moreover, in order to get a patient-specific reference interval, we need to find the best resolution of available data,

which is adequate for the patient. For example, two 60-year-old female patients may have different reference intervals based on their specific clinical conditions and taken medications. In this context, I'm interested in understanding how we can use routine laboratory test results to help physicians. In current practice, a physician is focused on marginal deviation from the reference intervals. We are developing a method that considers abnormalities in laboratory tests using mutual information between the different tests. We are looking for an approach for alerting the physician for abnormal laboratory test results which consider the high dimensional distributions [c2]. Aging is now understood as a gradual and progressive deterioration of integrity across multiple organ systems. Aging is associated with an increased rate of diseases, comorbidities, and mortality. However, there is a large variability in health among individuals of the same age group. This is assumed to be related to Biological Age (BA). BA also called physiological age, is a measure of how well or poorly a person's body functions relative to his actual calendar age (CA), also called chronological age. The BA of a person can be higher or lower than their CA, since aging is not only limited to time, but is, in fact, a complex process with multiple causes. Previous studies have shown that there are some laboratory tests that are correlated with age. We are developing a method for estimating the biological age using routine laboratory test results. Then, we have a way to find people whose CA deviates from the expected BA. Depending on the direction of the deviation, we see different associations with specific clinical conditions [c3]. Moreover, not only deviation of BA from CA is associated with health but also the rate of aging as we showed [s1].

(b) Precise genetic-based risk

Complex traits, in contrast to Mendelian diseases, are traits that are affected by multiple genetic regions. I'm interested in finding genetic regions affecting such complex phenotypes. For example, we looked for, and found genetic regions affecting the risk for preterm birth [j12][j13], that differentiate pain progression cluster groups in knee osteoarthritis patients[j15] and genetic associations to lupus [j23]. Statistical models and twin studies can estimate the heritability of a phenotype, meaning the amount of variation in the population coming from inheritance and not from the environment. However, genetic studies and genetic-based models still lack the ability to explain the full heritability variation. This phenomenon is called the "missing heritability" problem. I'm interested in closing the gap of missing heritability gap. One approach is by using models that consider epistatic interactions. In a previous paper, we showed that there are phenotypes that have an epistatic effect [j19]. We build polygenic risk scores (PRS) specifically for the even and odd chromosomes and then test their interaction in a linear regression on phenotype. We call this the even–odd (EO) test, and we prove analytically that EO reliably estimates CE under polygenicity. We now study other Machine Learning (ML)-based models to close the gap of missing heritability [s2][s3][s7][s8][p1][j26].

(c) Novel Machine Learning and Deep Learning approaches in health care

Prediction models in healthcare may have specific characteristics, whereas general ML models may not be applicable or should be modified. For example, not all misclassifications are the same. Classifying a tumor as benign where it's actually malignant may have worse consequences than classifying a malignant tumor as benign. One such model is a classification with rejection. We developed a model that, besides learning the prediction of the labels, also learns to return confidence in the prediction power of the label. Such confidence is learned disentangled from the labels. A possible application of such a model is in the intensive care unit discharge system. Where patients with high confidence can be discharged from or kept in the unit, while a patient with low confidence in the predictor should be kept or sent for further examinations [c5]. Another clinical-specific constraint that drives a novel machine learning model is an alert system where an alter interval is required.

Currently, alert systems exist. For example, an alert system for patients at risk for mortality or for sepsis. The limitation of this approach is that they evenly weight all samples. We showed that events occurring close to the prediction point are easier to classify and, therefore, push evaluation metrics to over-estimation. We developed an evaluation metric that reduces the weights as a function of the distance of the event from the prediction point. And we develop a deep learning model with the constrains of a required alert interval before alerting. Such an alert system will enable the care provider to intervene in case of an alert in order to prevent a life-threatening event [c4]. Another challenge specific to the medical field is the small sample size, which requires special solutions. We demonstrated such a solution and applied it to Multiple Sclerosis deterioration prediction. For example, it can be done by generating synthetic samples [c6][s10]. Another challenge in clinical data is that the data is often time-series. A patients get different treatments, and diagnosis and observations along his life, and person vary in the length of sequences. We introduced a novel Deep Learning architecture that make the model highly efficient and flexible to very large sequence lengths. It was first applied for proteins annotation predictions, where each protein is represented by a sequence. The architecture of ProteinBERT consists of both local and global representations, allowing end-to-end processing of these types of inputs and outputs[s2]. Clinical practice, guidelines, and possibilities, changes over time. We are interested in how we can use 20 years of clinical data and generate recommendations for treatment that take into account trends. We demonstrated such an approach by analyzing data of about 1.5M Coronary Artery Bypass Grafting (CABG) surgeries performed in the last 20 years. We developed a model, adjusting for many clinical conditions, and showed that the risk of redo CABG decreased to the point where it has the same risk for first-time CABG, and therefore the default procedures like heart catheterization should be reconsidered [j27][j28].

(d) Applications of Machine Learning in Healthcare

I'm also interested in general applications of ML in health care. Using statistical models, we showed that there is an association between countries with higher rates of BCG vaccines and lower rates of COVID-19 cases and deaths [j18]. We also showed a negative correlation between BCG treatment for bladder cancer and a significantly reduced risk of Alzheimer's disease and Parkinson's disease [j20]. We tested retrospectively differences in juvenile spondyloarthritis in different countries [j21]. And we revealed risk factors for Endometriosis [j24]. Improving healthcare process using ML methods [j25][j32]. We also developed a visualization tool for time-series clinical data in a tool called PatientExploreR [j17]. Such a tool can be useful for researchers, but also for physicians. And we developed algorithms that can be used in many fields, and specifically where clinical, biomedical, and biomolecular data is involved [c6][j22][s4][s6].

(e) Distributed learning

Clinical and medical data regulation prevent data sharing, which limit the ability to study diseases and clinical conditions with low prevalence. And it limits the opportunities to validate results externally. One approach to overcome this limitation is by standardizing the data in every site, and then only share the model, scripts or queries, and its results. I started working with such a common data model called OMOP CDM. I led the conversion of clinical data from Israel's governmental hospitals. An application of such conversion is a study in which we developed a prediction model for surgery durations, based on pre-surgery clinical data [j25][j26][j29][j31]. Another approach is to use distributed systems [s9].

(f) Large Language Models in Healthcare

Large Language Models (LLMs) like ChatGPT hold immense potential to revolutionize healthcare by enhancing clinical decision-making, patient management, and healthcare system efficiency. Recent studies underscore their versatility and the challenges that remain. For instance, the MedConceptsQA benchmark demonstrates that while current clinical LLMs struggle with medical concept reasoning, advanced models like GPT-4 show significant promise, achieving notable performance gains in zero-shot and few-shot learning[67]. Similarly, federated learning approaches, such as those using BEHRT, have shown that privacy-preserving methodologies can achieve near-centralized accuracy for tasks like next-visit diagnosis prediction, paving the way for secure, decentralized AI solutions[49]. Other innovations, such as DSF-GAN for generating high-utility synthetic data and CPLLM for fine-tuning LLMs in clinical prediction tasks, highlight how novel architectures and adaptive methods can elevate both data utility and predictive accuracy, surpassing traditional models like RETAIN and Med-BERT[10].

Building on these advancements, future research will focus on estimating the potential and the limitations of LLMs in the healthcare systems. Moreover, we will focus on integrating LLMs seamlessly into healthcare workflows, ensuring their utility in real-time clinical environments. This involves developing benchmarks that encompass broader and more complex medical scenarios, enhancing the interpretability of LLM-driven predictions to foster clinician trust, and refining federated learning frameworks to ensure scalability across diverse healthcare systems. Additionally, efforts will aim to optimize LLMs for patient-specific predictions, leveraging historical data to provide tailored recommendations for disease management and hospital readmission risks. By combining these strategies with robust evaluation methods and privacy-preserving technologies, LLMs can drive a paradigm shift in healthcare delivery, offering safer, more efficient, and patient-centered solutions.

- **Teaching Statement**

As a senior lecturer deeply committed to fostering meaningful learning experiences, I approach my teaching with a holistic perspective that integrates both the theoretical foundations and practical applications of machine learning and data science. My teaching philosophy revolves around nurturing critical thinking, problem-solving skills, and a deep understanding of the subject matter, preparing students to excel in their academic pursuits and professional endeavors.

a) Goals for Student Learning: In each of my courses, my primary goal is to empower students with the knowledge, skills, and confidence to tackle complex challenges in their respective fields. For the graduate course in machine learning, I aim to equip students with a comprehensive understanding of advanced machine learning algorithms and techniques, enabling them to apply these methods effectively in real-world scenarios. In the undergraduate machine learning course, my objective is to provide students with a solid foundation in machine learning principles and methodologies, preparing them for further studies or careers in data science and related fields. In the "Clinical Data Science" elective course, my goal is to bridge the gap between data science and healthcare, empowering students to leverage data-driven approaches for improving healthcare outcomes and advancing medical research.

b) Methods and Strategies to Achieve Goals: To achieve these goals, I employ a variety of teaching methods and strategies tailored to the specific needs of each course. These include lectures supplemented with interactive discussions, hands-on coding exercises, case studies, and real-world projects. I emphasize active learning and problem-based approaches, encouraging students to explore concepts through experimentation and collaboration. Additionally, I incorporate the latest advancements and case studies from relevant industries to provide students with practical insights and real-world relevance. For example, bringing to class examples from the general news recently

published.

c) Assessing and Evaluating Student Engagement and Comprehension: I utilize a multifaceted approach to assess student engagement and comprehension throughout the semester. This includes regular quizzes, assignments, coding projects, and class participation. I also provide constructive feedback to students, guiding them toward deeper understanding and improvement.

d) Future Directions and Improvements: Looking ahead, I am committed to continuously enhancing my teaching methodologies and course content to adapt to evolving educational needs and technological advancements. I plan to incorporate more interdisciplinary perspectives and hands-on experiences into my courses, fostering connections between theory and practice. Additionally, I aim to further integrate project-based learning and research opportunities into the curriculum, enabling students to apply their knowledge in real-world contexts and contribute meaningfully to the field.

In conclusion, my teaching philosophy revolves around fostering a dynamic and inclusive learning environment where students are challenged, inspired, and empowered to reach their full potential. I am dedicated to cultivating lifelong learners who are equipped to make meaningful contributions to society and drive innovation in their respective fields.