

Sabina Stefan Oller

sabinastefanoller@gmail.com | Boston, MA | [Google Scholar](#) | [Personal Website](#) | [GitHub](#)

AI scientist and computational biologist with a PhD in Biomedical Engineering and four years of postdoctoral research in oncology. I develop multimodal and self-supervised machine-learning methods that integrate functional genomics, multi-omics, and clinical data to model treatment response, predict cancer vulnerabilities, identify biomarkers, and stratify patients. My work focuses on translating signals across preclinical models and patient tumors and extracting interpretable, testable biological hypotheses; it has resulted in multiple first-author publications, including a recently accepted paper in *Nature Genetics*.

Technical Expertise

AI / ML: transformers and attention; multimodal representation learning; self-supervised and foundation-model approaches; individual treatment-effect modeling; generative modeling; model interpretability; LLM-based agentic workflows

Computational biology: translational oncology; functional genomics and CRISPR screening; bulk and single-cell transcriptomics; multi-omics integration; cancer dependency and drug-response modeling; biomarker discovery; patient stratification

Programming: Python, PyTorch, R, TensorFlow, scikit-learn, Hugging Face, NumPy/pandas, Linux, Git, Weights & Biases, GPU/HPC computing; modular, reproducible analysis and model-development pipelines

Research experience

Postdoctoral Research Fellow | Dana-Farber Cancer Institute / Harvard Medical School
Advisors: Dr. Volker Hovestadt & Dr. Kimberly Stegmaier

2022-present

Translational oncology and treatment-response modeling

- Developed MANTIS, a multimodal AI framework integrating DNA methylation, copy-number, and clinical data from 2,786 pediatric brain tumor patients to generate individualized survival predictions and estimate treatment effects.
- Demonstrated improved prognostic performance over established clinical risk indicators and identified a subgroup of average-risk medulloblastoma patients who may be candidates for treatment de-intensification; first-author study accepted at *Nature Genetics*.
- Model interpretation identified the prognostic importance of a distal enhancer of the oncogene *MYC*, generating a biological hypothesis that was subsequently functionally validated.
- Integrated and validated heterogeneous molecular and clinical data across multiple independent patient cohorts, supporting robust predictions across diverse datasets and disease subgroups.

Functional genomics and preclinical-to-patient transfer

- Developed a unified multimodal framework for genetic dependency and drug-response prediction that integrates genome-scale CRISPR and pharmacological screens with molecular profiling and patient tumor data, learns from partially observed and unpaired datasets, and remains robust to missing modalities and assay-specific readouts.
- Developed self-supervised representations to improve transfer from preclinical models to patient tumors, mitigating batch effects and improving alignment between cancer cell lines and patient samples with improved downstream predictions in patient data.
- The unified model outperforms standard tabular baselines for over 90% of prediction targets and, when applied to patient tumors, prioritizes therapeutic vulnerabilities that may be overlooked when preclinical data are analyzed in isolation.
- Develop LLM-based agentic workflows that orchestrate multi-step biological analyses, integrate evidence across data sources, and support hypothesis generation.

PhD Researcher in Biomedical Engineering | Brown University

2016-2022

- Developed deep learning methods for high-dimensional biomedical imaging, including image enhancement, segmentation, vascular quantification, and estimation of physiological parameters from optical coherence tomography.
- Built end-to-end computational pipelines combining simulation, deep learning, and quantitative validation to extract biologically meaningful measurements and track longitudinal changes from imaging data.
- Led and published first-author methodological research in *Nature Communications*, *Frontiers in Neuroscience*, and *Biomedical Optics Express*.
- Taught and mentored students across biomedical optics, electricity and magnetism, and applied mathematics.

Research Intern | International Prevention Research Institute, France

2016

- Developed machine learning-based algorithms to detect signs of diabetic retinopathy from retinal images.

Education

2016-2022	PhD in Biomedical Engineering , <i>Brown University</i> Advisor: Dr. Jonghwan Lee Thesis: Quantifying cortical tissue and vascular properties using optical coherence tomography Relevant coursework: ENGN 2520 - Pattern Recognition and Machine Learning, CSCI 1470 – Deep Learning
2015	B.S. (Honors) in Bioinformatics , awarded with distinction <i>University of Cape Town, South Africa</i> GPA: 4.0
2012-2014	B.S. in Physics and Applied Mathematics , awarded with distinction <i>University of Cape Town, South Africa</i> GPA: 4.0
2013	Visiting Student <i>University of California, Berkeley</i>

Publications

Accepted Manuscripts

Stefan, S., Amann, A. S., Smith, K. S., Boniolo, F., Pfitzer, E. C., Li, Q., Robinson, G. W., Harter, P. N., Lotter, W., Northcott, P. A., & Hovestadt, V. AI-based Precision Prognostics and Therapy Personalization for Childhood Brain Tumors. Accepted at *Nature Genetics*.

Manuscripts in Preparation

Stefan, S., Stegmaier, K., Hovestadt, V. Multi-task representation learning prioritizes genetic and pharmacological vulnerabilities in patient tumors.

Peer-Reviewed Publications

2025	Assi, J., Stefan, S. , Tang, R., Küreli, G., Kura, S., Long, E., Jiang, J., Boas, D., & Lee, J. Machine learning-based method to detect capillary stalling using optical coherence tomography. <i>Biomedical Optics Express</i> , 16(5), 1807-1818.
2024	Daenekas, B., Pérez, E., Boniolo, F., Stefan, S. , Benfatto, S., Sill, M., Sturm, D., Jones, D. T., Capper, D., Zapatka, M., & Hovestadt, V. Conumee 2.0: enhanced copy-number variation analysis from DNA methylation arrays for humans and mice. <i>Bioinformatics</i> , 40(2), btae029.
2023	Walek, K. W.*, Stefan S.* , Lee, J. H., Puttigampala, P., Kim, A. H., Park, S. W., Marchand, P. J., Lesage, F., Liu, T., Huang, Y. A., Boas, D. A., Moore, C., & Lee, J. Near-lifespan longitudinal tracking of microvascular dynamics and structure. <i>Nature Communications</i> , 14(1), 2982. *Equal contribution.
2023	Lee, J. H., Stefan, S. , Walek, K., Nie, J., Min, K., Yang, T. D., & Lee, J. Investigating the correlation between early vascular alterations and cognitive impairment in Alzheimer's disease in mice with SD-OCT. <i>Biomedical Optics Express</i> , 14(4), 1494-1508.
2023	Sianga-Mete, R., Hartnady, P., Mandikumba, W. C., Rutherford, K., Currin, C. B., Phelanyane, F., Stefan, S. , Pond, S. L. K., & Martin, D. P. Viral genome sequence datasets display pervasive evidence of strand-specific substitution biases that are best described using non-reversible nucleotide substitution models. <i>eLife</i> , 12, RP87361.

- 2022 **Stefan, S.**, Kim, A., Marchand, P. J., Lesage, F., & Lee, J. Deep learning and simulation for the estimation of red blood cell flux with optical coherence tomography. *Frontiers in Neuroscience*, 16, 835773.
- 2020 **Stefan, S.**, & Lee, J. Deep learning toolbox for automated enhancement, segmentation, and graphing of cortical optical coherence tomography microangiograms. *Biomedical Optics Express*, 11(12), 7325-7342.
- 2018 **Stefan, S.**, Jeong, K. S., Polucha, C., Tapinos, N., Toms, S. A., & Lee, J. Determination of confocal profile and curved focal plane for OCT mapping of the attenuation coefficient. *Biomedical Optics Express*, 9(10), 5084-5099.
- 2018 **Stefan, S.**, Schorr, B., Lopez-Rolon, A., Kolassa, I., Shock, J. P., Rosenfelder, M., Heck, S., & Bender, A. Consciousness indexing and outcome prediction with resting-state EEG in severe disorders of consciousness. *Brain Topography*, 31(5), 848-862.

Conference Presentations

- 2025 **Stefan, S.**, Amann, A. S., Boniolo, F., Lambo, S., Steinicke, T., Smith, K. S., Harter, P. W., Northcott, P. A., & Hovestadt, V. AI-based Precision Prognostics and Therapy Personalization for Pediatric Brain Tumors. *Pediatric Society for Neuro-Oncology*. Talk (abstract award).
- 2025 **Stefan, S.**, & Hovestadt, V. Adverse-event embeddings predict drug mechanisms and toxicity. *Machine Learning in Computational Biology (MLCB)*. Poster.
- 2024 **Stefan, S.**, Amann, A., Smith, K., Northcott, P., & Hovestadt, V. Precision prognostics in pediatric neuro-oncology using DNA methylation arrays and neural networks. *Machine Learning in Computational Biology (MLCB)*. Spotlight presentation.
- 2024 **Stefan, S.**, Amann, A., Smith, K., Northcott, P., & Hovestadt, V. Precision prognostics in pediatric neuro-oncology using DNA methylation arrays and neural networks. *PQG AI for Genomics and Health*. Poster.
- 2022 **Stefan S.***, Walek, K. W.*, Lee, J. H., Puttigampala, P., Kim, A. H., Park, S. W., Marchand, P. J., Lesage, F., Liu, T., Huang, Y. A., Boas, D. A., Moore, C., & Lee, J. Longitudinal tracking of vascular properties in Alzheimer's disease using optical coherence tomography. *SPIE Photonics West*. Talk. *Equal contribution.
- 2020 **Stefan, S.**, & Lee, J. Deep Learning Toolbox for Vectorizing Microangiograms. *Brown Unconference on Computational Intelligence and Applications*. Talk.
- 2019 **Stefan, S.**, Polucha, C., & Lee, J. Deep learning for OCT angiogram vectorization. Program No. 525.06. 2019 Neuroscience Meeting Planner. Society for Neuroscience, Chicago, IL. Poster.
- 2019 **Stefan, S.**, Polucha, C., & Lee, J. Deep Learning for OCT Angiogram Vectorization. *Northeast Symposium on Biomedical Optics*. Poster.
- 2018 **Stefan, S.**, Polucha, C., & Lee, J. Mapping the Attenuation Coefficient of Tumorous Cortical Tissue using OCT. *Northeast Symposium on Biomedical Optics*. Poster.
- 2017 **Stefan, S.**, Schorr, B., Lopez-Rolon, A., Kolassa, I. T., Shock, J., & Bender, A. Automated Outcome Prediction and Assessment with Quantitative EEG in Severe Disorders of Consciousness. *Annual Meeting of the German Society for Clinical Neurophysiology and Functional Imaging*. Poster.

Fellowships and Awards

- 2026 Invited Speaker as one of the "Rising Stars in Neuro-Oncology", Center for Excellence in Neuro-Oncology Symposium
- 2025 Abstract Award, Pediatric Society for Neuro-Oncology (top 2% of all submissions)
- 2019 First Prize, AI in Healthcare Hackathon, Harvard University.
Developed a concept using a slit lamp combined with a cellphone camera to conduct eye examinations in rural areas.
- 2019 Best Real-World Application, Hack@Brown Hackathon, Brown University.
Used an OpenBCI EEG headset to determine alpha-band power and guide users through meditation in a virtual reality experience.

- 2015 Hyman Liberman Scholarship.
Awarded to the top-performing B.S. student at the University of Cape Town.
- 2015 Scarce Skills Development Fellowship (National Research Foundation of South Africa).
- 2014 Undergraduate Research Fellowship (National Research Foundation of South Africa).