Docket #: S25-309

Generation of Biological Profiles Using Electronic Medical Record Foundation Model

Large-scale, granular biological datasets such as proteomic, metabolomic, and transcriptomic profiles are powering the current shift toward precision medicine. These data enable biomarker discovery for early disease detection, prediction of treatment response, identification of therapeutic targets, and the development of accurate predictive models. However, widespread adoption has been limited by the high cost, long turnaround times, and invasive nature of traditional wet-lab methods.

The Aghaeepour Lab at Stanford has developed a machine learning model that addresses these challenges by generating high-fidelity biological profiles directly from existing electronic health records, offering a fast, cost-efficient, and non-invasive alternative to conventional laboratory approaches. The model demonstrated robust capability in capturing established and novel biological signals across diverse diseases, closely paralleling real proteomic data, including the universal disease of aging through organ-specific aging markers. Additionally, the inventors applied the model to a rapid cohort study to identify proteomic signatures predictive of therapeutic responses in patients with rheumatoid arthritis, demonstrating improved predictive performance over clinical benchmarks generalized to an external validation cohort. The invented synthetic proteomics framework represents a scalable, cost-effective approach to significantly enhance the integration of large-scale clinical data with high-dimensional omics to broadly inform translational and mechanistic biomedical research.

Applications

- -Synthetic omic profiling
- -Life Sciences Artificial Intelligence

- -Drug development and discovery platform

Advantages

- -First of its kind in digital, large-scale omics-profile generation
- -Minimal infrastructure and up-front cost
- -Low unit cost compared to standard wet-lab assays
- -Applicable to standard electronic health record repositories
- -Near-instant turnaround due to digital procedure
- -Scalability to population studies

Innovators

- David Seong
- Nima Aghaeepour

Licensing Contact

Michael Spaid

Technology Licensing Associate 2

Email