Evaluating the Alzheimer's disease data landscape

Colin Birkenbihl, Yasamin Salimi, Daniel Domingo-Fernández, Simon Lovestone, AddNeuroMed consortium, Holger Fröhlich, Martin Hofmann-Apitius, the Japanese Alzheimer’s Disease Neuroimaging Initiative and the Alzheimer’s Disease Neuroimaging Initiative

Introduction: Numerous studies have collected Alzheimer’s disease (AD) cohort data sets. To achieve reproducible, robust results in data-driven approaches, an evaluation of the present data landscape is vital.

Methods: Previous efforts relied exclusively on metadata and literature. Here, we evaluate the data landscape by directly investigating nine patient-level data sets generated in major clinical cohort studies.

Results: The investigated cohorts differ in key characteristics, such as demographics and distributions of AD biomarkers. Analyzing the ethnорacial diversity revealed a strong bias toward White/Caucasian individuals. We described and compared the measured data modalities. Finally, the available longitudinal data for important AD biomarkers was evaluated. All results are explorable through our web application ADataViewer (https://adata.scai.fraunhofer.de).

Discussion: Our evaluation exposed critical limitations in the AD data landscape that impede comparative approaches across multiple data sets. Comparison of our results to those gained by metadata-based approaches highlights that thorough investigation of real patient-level data is imperative to assess a data landscape.

Alzheimer’s & Dementia: Translational Research & Clinical Interventions

Published Online

16 December 2020

https://doi.org/10.1002/trc2.12102