

Adapting Monotonic Subtype and Stage Inference (SuStaIn) to Model Patterns of Disease Progression for Psychiatric Disorders

Ethan Yan¹, Katie M Lavigne^{2,3}

¹Department of Psychology, McGill University, Montreal, QC, Canada

²Douglas Research Centre, Montreal, QC, Canada

³Department of Psychiatry, McGill University. Montreal, QC, Canada

Keywords: Disease progression modelling, Event-based models, Psychiatric disorders, Subtype and stage inference. Computational modelling

Corresponding author email: siliang.yan@mail.mcgill.ca

Published: 03/25/2025

<https://doi.org/10.26443/msurj.v21i2.494>

Significant progress has been made in disease progression modelling, with crucial implications for understanding and managing chronic, progressive diseases, such as neurodegenerative diseases. One prominent framework is Subtype and Stage Inference (SuStaIn), an event-based machine learning model that characterizes disease progression and can reveal spatiotemporal and phenotypic heterogeneity in chronic diseases using only cross-sectional data. SuStaIn has demonstrated success across neurodegenerative, psychiatric, and pulmonary diseases. However, a key limitation of SuStaIn is its assumption of monotonic disease progression, which may be biologically implausible and clinically restrictive, particularly for diseases that commonly involve remission and recovery (e.g., psychiatric disorders). *To examine the impact of this assumption, this proof-of-concept study systematically evaluated SuStaIn's performance on non-monotonic data.* We generated both monotonic and bidirectional ground-truth disease progression datasets. The bidirectional datasets were further manipulated across sample sizes, numbers of biomarkers, and subtype proportions. Model performance was evaluated using Kendall's tau to compare the inferred event sequences with the ground-truth progression. Results show that SuStaIn performs significantly worse on datasets with bidirectional underlying disease progression than on baseline monotonic datasets. While SuStaIn reliably captures the number of subtypes and their proportions, it fails to accurately model progression within subtypes that contain switch biomarkers. Moreover, increasing the sample size and the number of biomarkers does not substantially improve performance under bidirectional conditions, in contrast to monotonic progressions. These findings provide empirical evidence that SuStaIn's monotonic assumption limits its ability to model bidirectional disease processes, as in many psychiatric disorders. More broadly, this work deepens the understanding of SuStaIn as an event-based model and informs future methodological extensions to relax the monotonic progression constraint.

© The Authors. This abstract is published under a CC-BY license:

<https://creativecommons.org/licenses/by/4.0/>