AI-Driven Insights into Early-Onset Multimorbidity Clusters and Long-Term Health Trajectories: A Path Towards Improved Healthcare Planning

Mozhdeh Shiranirad, Emilia Holland, Roberta Chiovoloni, Michael Boniface, Nisreen A Alwan, Simon DS Fraser, and Rebecca B Hoyle
with the MELD-B Consortium

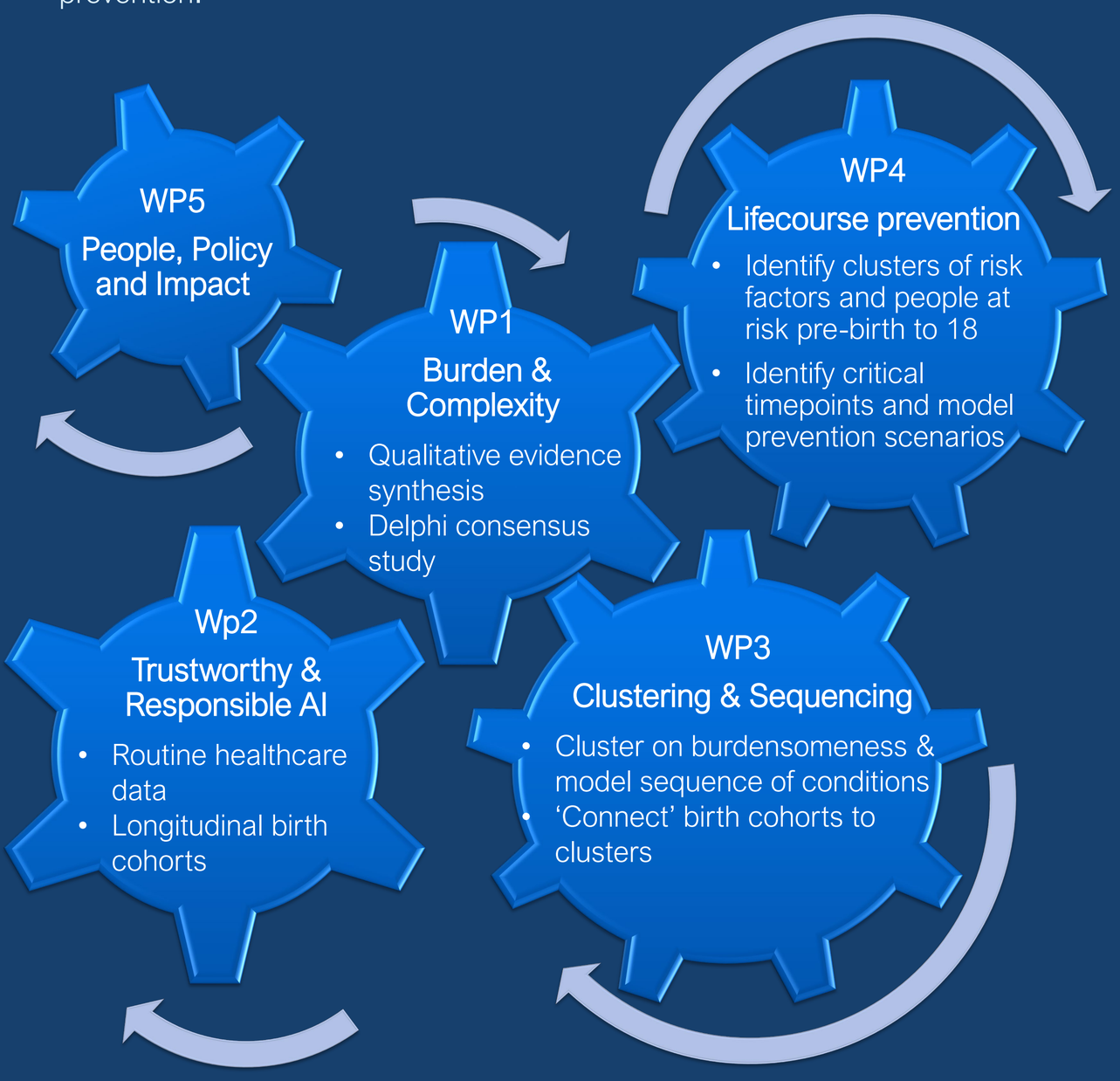
FUNDED BY NIHR | National Institute for Health and Care Research

Multiple Long-term Condition Multimorbidity

One in four individuals grapples with managing two or more long-term health conditions (LTCs) resulting in lower quality of life and increased mortality risk. Deprived areas experience the onset of these conditions 10-15 years earlier than affluent ones. With an aging population, the prevalence of people having multiple LTCs (MLTCs) is expected to rise significantly, impacting individuals, families, and society, leading to increased human and service costs particularly associated with mental health conditions linked to long-term physical health issues[1] .

meld-b

The MELD-B (Multidisciplinary Ecosystem to study Lifecourse Determinants and Prevention of Early-onset Burdensome Multimorbidity) research project aims to understand when MLTC multimorbidity (MLTC-M) becomes **'burdensome'** and identify potential opportunities for intervention. We use comprehensive analysis of birth cohorts and electronic health records, harnessing the power of cutting-edge artificial intelligence (AI) techniques to pinpoint pivotal life-course periods for potential early-onset MLTC-M prevention.



Primary Burden indicators

A **qualitative evidence synthesis** comprehensively explores the lived experience of individuals with multiple long-term conditions, focusing on understanding the associated burdens. This informs a modified **Delphi consensus study** to gather feedback from clinicians, academics, patients, and carers on identified indicators. The goal is to validate these indicators for incorporation into clustering analysis.

Accumulation & complexity	Finance	
Health service & administration	Emotions	Medication
Investigation & monitoring	Learning & adapting	Symptoms

The burden indicators from the qualitative evidence synthesis are categorized into 8 themes representing the burden of living with MLTCs.

Dataset

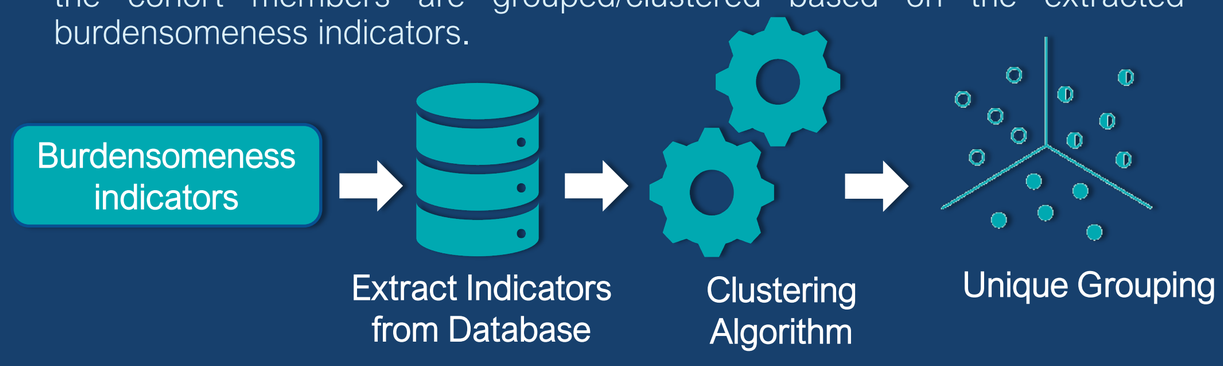
Four exploratory indicators of 'burdensomeness' were identified from a cohort of 4,896,520 individuals in the SAIL[2] (Secure Anonymized Information Linkage) Databank between January 2000 and December 2021: age at onset, interval between accrual of LTCs, presence of Mental Health Conditions component of LTC (PMHC), and frequency of hospital appointments.

Cohort Demographic Characteristics

Sex		Ethnicity				
Female	Male	White	Mixed	Asian	Black	Other
49.6%	50.4%	92.4%	1.3%	4.0%	1.0%	1.3%

Clustering in Burdensomeness Space

Utilizing unsupervised machine learning clustering methods (k-prototype)[3] , the cohort members are grouped/clustered based on the extracted burdensomeness indicators.



Clustering Results

Using these four exploratory burdensomeness indicators and the k-prototypes algorithm[3], we found five clusters in 'burdensomeness space', shown below by colour and list of key characteristic features.

- Age at onset > 68
- High accrual interval between conditions (>6 years)
- No PMHC

- Early to medium age at onset
- PMHC

- Age at onset of first condition after 56
- No PMHC

- Early age of onset
- High frequency of hospital appointments
- Very low (<1 year)/ very high (>6 years) accrual interval between conditions
- No PMHC

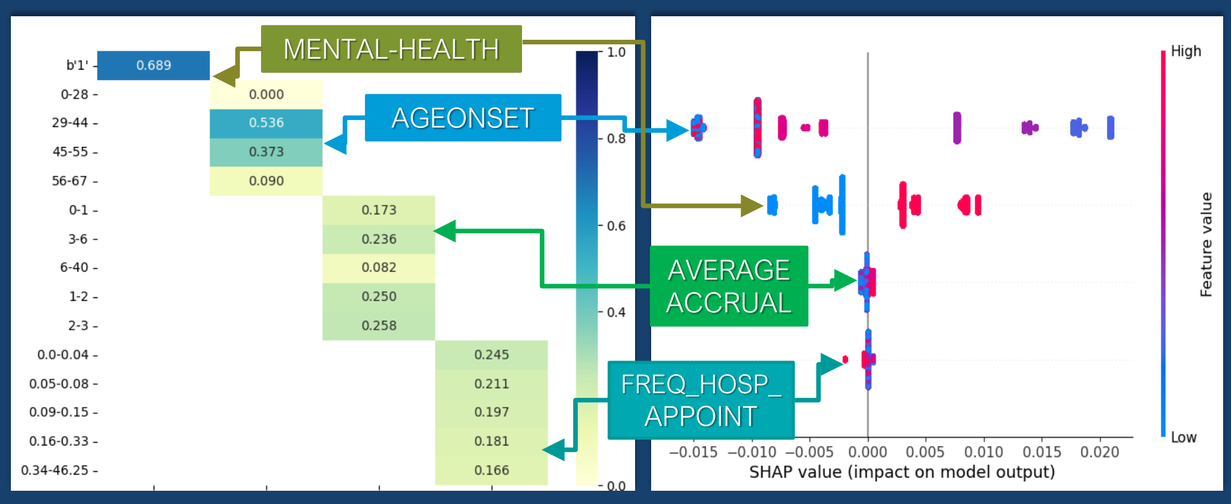
- Age at onset of first condition very young
- PMHC

Initial Burdensomeness Indicators

The feature characteristics of each cluster are explained via two methods:

1. Shapley Additive explanation [4] (SHAP), which quantifies the magnitude and direction (positive or negative) of an indicator's effect on cluster membership,
2. Frequency analysis, which shows the prevalence of different categories for a feature.

The frequency and SHAP plots for the purple cluster are shown on left and right respectively.



Collaborating Institutions

References

[1]Fraser, S. D. S., et al.: J. Multimorb. Comorb. 13, (2023)
[2]https://saildatabank.com
[3]Huang, Z.: Data Min. Knowl. Discov. 2, 283 (1998),
[4]Lundberg, S. M., and Lee, S.-I.: Adv. Neural Inf. Process. Syst. 30, (2017)