Health Types in Chronic Disease Trajectories: Socioeconomic Disparities over the Life Course

Ailun Shui[a]

[a] *University of Groningen*

Keywords: Socioeconomic status, Chronic diseases, Health types, Health inequality

1. Introduction

It is widely acknowledged that socioeconomic health disparities persist across the life course. In early adulthood, health disparities begin to manifest as differences in biomarkers that reflect underlying physiological dysregulation, often occurring before the onset of clinical disease (Shui et al., 2025). As individuals age, these under-the-skin risks translate into disparities in diagnoses and subsequent health outcomes, ultimately contributing to substantial mortality gaps in later life (Cutler et al., 2006; Van Kippersluis et al., 2010; Danesh et al., 2024). At any given age, individuals from socioeconomically disadvantaged backgrounds tend to experience poorer health outcomes. The pronounced socioeconomic disparities observed in later life reflect the cumulative effects of long-term health processes unfolding across the life course. Although the socioeconomic health gradient at specific time points is well documented, less is known about how socioeconomic status (SES) shapes health trajectories over time. This gap raises an important question: Are there latent types of chronic disease trajectories, and how do socioeconomic conditions shape these trajectories in ways that ultimately contribute to the substantial disparities observed in later life?

Health evolves in dynamic and path-dependent ways. The earlier and longer individuals remain in poor health, the more difficult recovery becomes and the greater the associated costs (De Nardi et al., 2025). Recent research shows that individuals' health tends to follow distinct trajectories, which strongly predict subsequent health outcomes and are associated with substantial economic consequences (De Nardi et al., 2025; Borella et al., 2025). When health disparities, particularly in midlife and later life, are treated as outcomes, these types can be understood as reflecting the cumulative processes that generate such disparities. Accordingly, examining the role of SES in shaping health trajectories not only helps identify the origins of health disparities but also sheds light on why and how these disparities widen over time.

In this paper, I extend existing research by identifying latent types of health trajectories and the socioeconomic gradient across and within these types. Specifically, I investigate whether latent health types exist in the chronic disease trajectories and how SES is related

to the types. The evolution of chronic diseases provides a particularly informative context for studying the dynamic nature of socioeconomic health disparities. Chronic conditions are persistent, cumulative, and often co-occurring, making them key drivers of health deterioration at older ages and major contributors to health-care costs (Bauer et al., 2014; Boersma et al., 2020; Hacker, 2024). In the Netherlands, the National Institute for Public Health and the Environment (RIVM) projects that by 2030, nearly 7 million people, around 40% of the population, will be affected by chronic conditions. Moreover, there is a pronounced SES gradient in the prevalence of chronic diseases (Danesh et al., 2024; Pallesen et al., 2024). Therefore, understanding the processes that shape large SES disparities in chronic disease is of critical importance, as it offers valuable insights for preventive screening.

2. Data and Methodology

The Dutch Lifelines Cohort Study and Biobank is a population-based, prospective study conducted in the Northern Netherlands, with over 167,000 participants at baseline. I use data from Lifelines to follow more than 22,000 individuals over six waves, covering an average span over 10 years. The dataset includes self-reported chronic disease conditions, health outcomes, socioeconomic characteristics, and health-related behaviors. These longitudinal records allow me to track changes in having chronic diseases over time and to analyze long-term health dynamics.

The analysis consists of two main parts. The first part focuses on measuring health trajectories, identifying latent health types, and analyzing how health evolves across these types. I construct a composite chronic disease index (CDI), defined in Equation 1, to capture chronic health conditions. The CDI measures the number of chronic diseases an individual currently has or has had at each age. I define a chronic disease trajectory as a five-element vector of CDI values across five waves, as formalized in Equation 2.

$$CDI_{i.a} = \sum_{k=1}^{19} I_{i,a,k} \tag{1}$$

Here, k indexes chronic diseases, and I is a binary indicator that equals one if an individual has or has had the disease at a given survey wave a.

$$H_i = [h_{i,1a}, h_{i,1a}, h_{i,1a}, h_{i,1a}, h_{i,1a}] \tag{2}$$

Here, H_i denotes a five-dimensional vector representing individual i's health status. Each element corresponds to the CDI for individual i at different survey waves (waves 1a, 1b, 1c, 2a, and 3a).

Based on these trajectories, I classify individuals into latent health types based on their health trajectories. I follow the concept of health types proposed by (De Nardi et al., 2025)

and extend the approach of (Borella et al., 2025), using a machine learning k-means clustering technique to group individuals into three types. Each individual is assigned to a type based on their five-wave CDI trajectory. I then validate the health types by evaluating the predictive power of these types using chronic disease data from the wave 3b.

$$H_i = [h_{i,1a}, h_{i,1a}, h_{i,1a}, h_{i,1a}, h_{i,1a}] \tag{3}$$

The second part of the analysis investigates the relationship between educational attainment and health types. I employ a multinomial logit model to assess the probability of belonging to each health type by education, conditional on demographic characteristics and health behaviors. Finally, I examine health dynamics by education, conditional on health type, in order to explore socioeconomic disparities within each type.

3. Results

Individuals are classified into three health types based on their CDI trajectories across five waves, and the types are subsequently ordered according to their mean CDI levels. Type 1 ("low and stable burden") comprises individuals whose CDI values remain close to zero throughout the observation period, indicating little to no chronic disease burden. Type 2 ("moderate and increasing burden") includes individuals with a moderate average CDI that rises gradually over time; although their baseline burden is relatively low, it increases with age. Type 3 ("high and progressively worsening burden") consists of individuals with the highest baseline CDI, which continues to increase during the follow-up period, reflecting the most severe and persistently worsening chronic disease burden.

Distinct trajectories across health types

Figure 1 depicts the health dynamics of CDI across health types. Distinct trajectories are observed. The good health type (Type 1) exhibits an almost negligible number of chronic diseases throughout the observation period. In contrast, the poor health type (Type 3) shows a substantially higher average number of chronic conditions, with a continuous increase following the baseline measurement. The moderate health type (Type 2) falls between these two extremes, displaying a moderate chronic disease burden that increases steadily over time, although at a slower rate than that observed for Type 3.

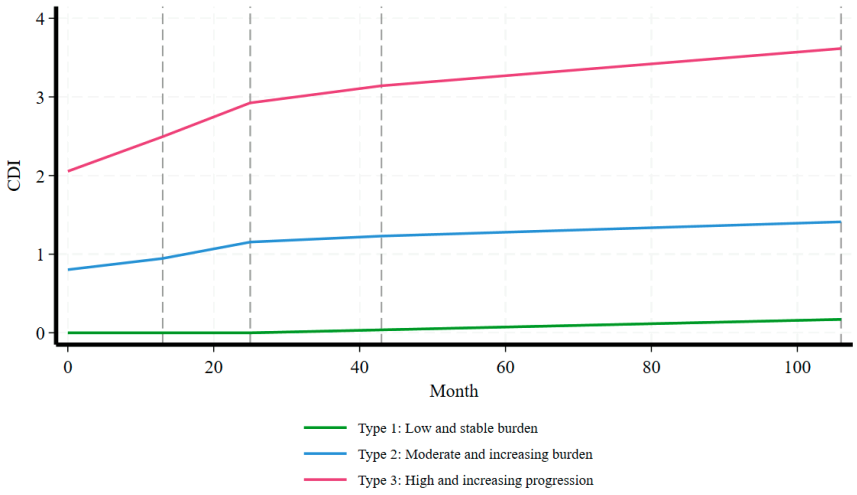


Figure 1: The Health Dynamics of CDI by Health Types

Notes: This figure illustrates the dynamic trajectories of the chronic disease index (CDI) by health type. The x-axis represents months since the baseline survey. Vertical dashed lines denote the average survey timing for waves 1b, 1c, 2b, and 3a since the first wave. The CDI is adjusted to account for irregular observation intervals across individuals.

Transitions and health evolution across health types over the life cycle

To further examine how health types evolve over the life cycle, specifically how the distribution of observations across health types changes with age, Figure 2 presents a stacked area graph illustrating these dynamics using the pooled balance panel. As age increases, the proportion of observations classified as Type 2 and Type 3 rises correspondingly. Figure 3 complements Figure 2 by illustrating health dynamics across the life cycle by health type using the pooled balanced panel. The results reinforce the previously observed patterns.

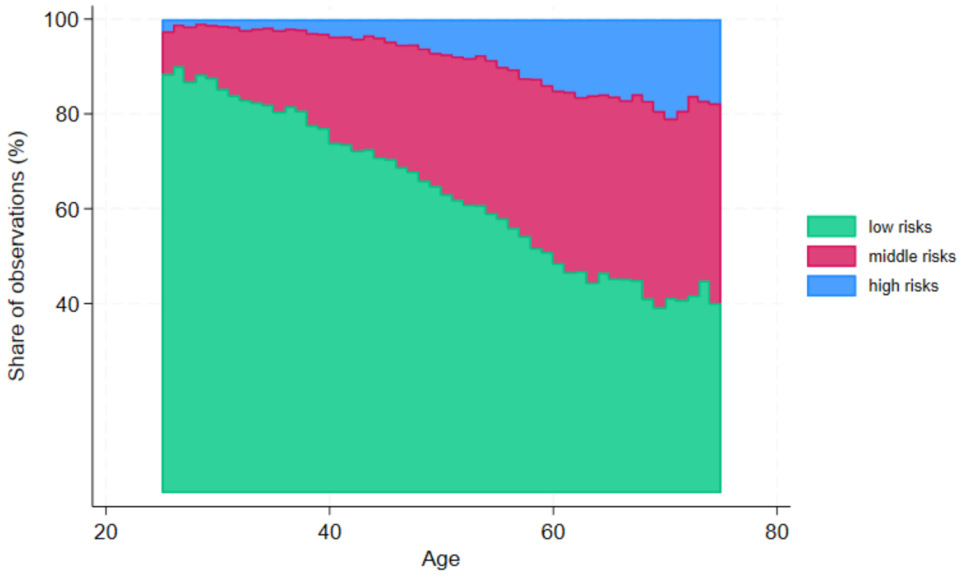


Figure 2: Evolution of Health Types over the Life Cycle

Notes: This figure presents the distribution of observations across health types based on pooled data from a balanced panel. The sample includes individuals aged 25 to 75.

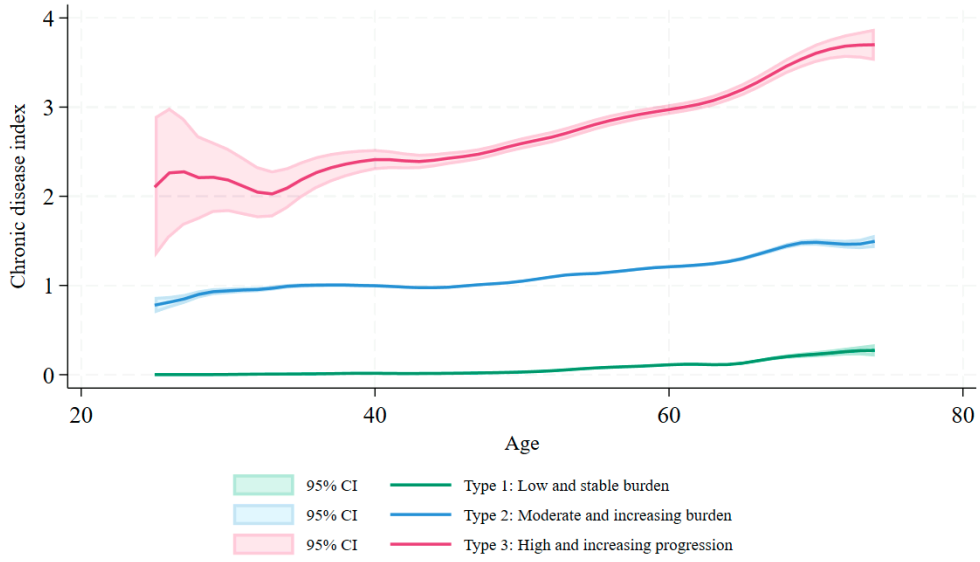


Figure 3: The Health Dynamics of CDI by Health Types

Notes: This figure depicts the health evolution by health types using the based on pooled data from a balanced panel. The sample includes individuals aged 25 to 75.

The socioeconomic gradient across and within health types

To examine the extent to which baseline SES predicts individuals' health type, I estimate a multinomial logit model to assess the probability of belonging to each health type across different SES groups. Table 1 reports the average marginal effects of education. As shown in Column 1, individuals in the high-education group are, on average, 2.95 percentage points more likely to belong to the low and stable burden type (Type 1). In contrast, they are 1.57 and 1.37 percentage points less likely to belong to Type 2 and Type 3, respectively. After further adjusting for employment status, partnership status, and health-related behaviors, a clear gradient in the probability of belonging to different health types persists. Individuals with higher educational attainment consistently exhibit a relative advantage.

Next, I investigate socioeconomic disparities in health dynamics conditional on health type. As shown in Figure 4, the health gradient is most pronounced in the worst health type (Type 3). In contrast, there is little evidence of a gradient in Type 1. This pattern suggests that SES may act as a buffer for individuals who are more vulnerable in the poorest health type. Overall, these results indicate that educational disparities in health reflect both differences in the distribution of individuals across health types and variations in health dynamics within the same type.

Table 1: Average Marginal Effects of Education on Health Types

	(1) Model 1	(2) Model 2	(3) Model 3
Low and stable burden (Type 1)	0.0295***	0.0291***	0.0217***
	(0.0063)	(0.0063)	(0.0064)
Moderate and increasing burden (Type 2)	-0.0157**	-0.0157**	-0.0111*
	(0.0062)	(0.0062)	(0.0063)
High and increasing burden (Type 3)	-0.0137***	-0.0134***	-0.0105***
	(0.0028)	(0.0028)	(0.0028)
Controls: Basic	Yes	Yes	Yes
Controls: Employment + Partnership		Yes	Yes
Controls: Health-related behaviors			Yes
Number of individuals	22,318	22,318	22,318

Notes: This table reports average marginal effects of high education (relative to low education) from multinomial logit models. Basic controls include age, age squared, gender, province, and cohort. Specification (2) additionally controls for current employment status and partnership status. Specification (3) further includes health-related behaviors, including smoking, drinking (frequency and intensity), and physical activity. Standard errors are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

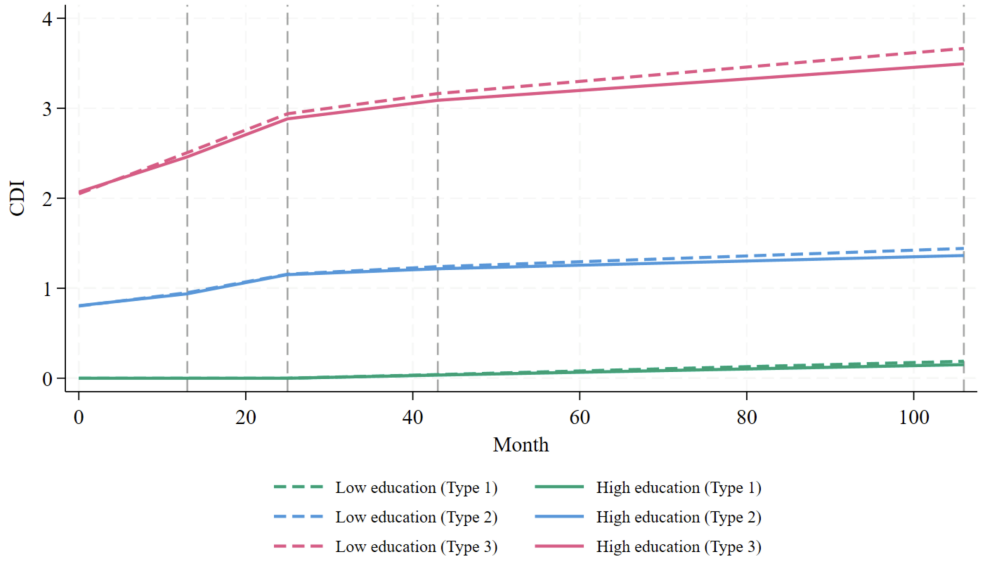


Figure 4: The Health Dynamics of CDI by Health Types and Education

Notes: This figure depicts the health dynamics by health types and education.The x-axis represents months since the baseline survey. Vertical dashed lines denote the average survey timing for waves 1b, 1c, 2b, and 3a since the first wave. The CDI is adjusted to account for irregular observation intervals across individuals.

4. Conclusion

This study identifies three distinct latent health types in the trajectories of chronic diseases over the life course. While individuals in the good health type maintain a near-zero chronic disease index (CDI) with only minimal increases over time, those in the bad health type experience a substantially higher disease burden and a faster deterioration in health. Although transitions between health types may occur, health trajectories are generally stable, with future health outcomes strongly conditioned on prior health status.

The results also reveal pronounced socioeconomic disparities both across and within health types: socioeconomically advantaged individuals are more likely to belong to healthier trajectories, while socioeconomic gradients in health dynamics are particularly pronounced within the worst health type. These findings suggest that socioeconomic status shapes both the distribution of individuals across health trajectories and the degree of divergence within them, contributing to the widening of health disparities over the life course.

From a policy perspective, the persistence of health trajectories highlights the importance of early and targeted interventions in chronic disease prevention, particularly for socioeconomically disadvantaged populations who are more likely to follow unfavorable health paths. Strengthening preventive care and reducing socioeconomic barriers to health may therefore play a crucial role in mitigating the long-term accumulation of chronic disease burden.

5. Acknowledgements

The author thanks Gerard J. van den Berg, Jochen O. Mierau, and Laura Viluma for their valuable comments and suggestions. The author also thanks the participants of the CPHEB and Lifelines seminars for their valuable feedback. The author express their gratitude to the Lifelines Cohort Study and Biobank, the research centers providing data to Lifelines, and all the participants of the study. The Lifelines initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG), Groningen University, and the Provinces in the North of the Netherlands (Drenthe, Friesland, Groningen. Ailun Shui acknowledges the support of the CSC scholarship program (Grant No. 202108500029), and is grateful for the resources provided by the International Max Planck Research School for Population, Health, and Data Science (IMPRS-PHDS). All funding was used for Ailun Shui's daily expenses and academic visits during his PhD and was not directly allocated to the research of this paper.

References

A. Shui, G. J. van den Berg, J. O. Mierau, L. Viluma, Lifetime trajectories and drivers of socioeconomic health disparities: Evidence from longitudinal biomarkers in the netherlands, Journal of Human Resources (2025).

D. Cutler, A. Deaton, A. Lleras-Muney, The determinants of mortality, Journal of economic perspectives 20 (2006) 97–120.

H. Van Kippersluis, O. O'Donnell, E. Van Doorslaer, T. Van Ourti, Socioeconomic differences in health over the life cycle in an egalitarian country, Social science & medicine 70 (2010) 428–438.

K. Danesh, J. T. Kolstad, W. D. Parker, J. Spinnewijn, The chronic disease index: Analyzing health inequalities over the lifecycle, Technical Report, National Bureau of Economic Research, 2024.

M. De Nardi, S. Pashchenko, P. Porapakkarm, The lifetime costs of bad health, Review of Economic Studies 92 (2025) 1987–2026.

M. Borella, F. Bullano, M. De Nardi, B. Krueger, E. Manresa, Health inequality and health types, The Econometrics Journal 28 (2025) 341–384.

U. E. Bauer, P. A. Briss, R. A. Goodman, B. A. Bowman, Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the usa, The Lancet 384 (2014) 45–52.

P. Boersma, L. I. Black, B. W. Ward, Prevalence of multiple chronic conditions among us adults, 2018, Preventing chronic disease 17 (2020) E106.

K. Hacker, The burden of chronic disease, Mayo Clinic Proceedings: Innovations, Quality & Outcomes 8 (2024) 112–119.

A. V. J. Pallesen, J. O. Mierau, F. K. Christensen, L. H. Mortensen, Educational and income inequalities across diseases in denmark: a register-based cohort study, The Lancet Public Health 9 (2024) e916–e924.