

Co-occurring diseases and mortality in patients with chronic heart disease modelling their dynamically expanding disease portfolio: A nationwide register study

Nikolaj Normann Holm, Anne Frølich, Helena Dominguez, Kim Dalhoff, Helle Gybel Juul-Larsen, Ove Andersen, Anders Stockmarr

Submitted to: JMIR Cardio
on: February 26, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 44

Figures 45

Figure 1..... 46

Figure 2..... 47

Figure 3..... 49

Figure 4..... 50

Figure 5..... 51

Figure 6..... 52

Figure 7..... 53

Figure 8..... 54

Figure 9..... 56

Multimedia Appendixes 57

Multimedia Appendix 1..... 58

Multimedia Appendix 2..... 58

Multimedia Appendix 3..... 58

Co-occurring diseases and mortality in patients with chronic heart disease modelling their dynamically expanding disease portfolio: A nationwide register study

Nikolaj Normann Holm¹ BSc, MSc, PhD; Anne Frølich^{2,3}; Helena Dominguez⁴; Kim Dalhoff^{5,6}; Helle Gybel Juul-Larsen⁷; Ove Andersen^{6,7,8}; Anders Stockmarr¹

¹ Department of Applied Mathematics and Computer Science Technical University of Denmark Kgs. Lyngby DK

² Innovation and Research Centre for Multimorbidity Slagelse Hospital Slagelse DK

³ Department of Public Health University of Copenhagen Copenhagen DK

⁴ Department of Biomedical Sciences University of Copenhagen Copenhagen DK

⁵ Department of Clinical Pharmacology Bispebjerg and Frederiksberg University Hospital Copenhagen DK

⁶ Department of Clinical Medicine University of Copenhagen Copenhagen DK

⁷ Department of Clinical Research Copenhagen University Hospital Amager and Hvidovre Hvidovre DK

⁸ Emergency Department Copenhagen University Hospital Amager and Hvidovre Hvidovre DK

Corresponding Author:

Nikolaj Normann Holm BSc, MSc, PhD

Department of Applied Mathematics and Computer Science

Technical University of Denmark

Richard Petersens Plads, Building 324

Kgs. Lyngby

DK

Abstract

Background: Medical advances in managing chronic heart disease (HD) patients permit co-occurrence of other chronic diseases due to increased longevity, causing them to become multimorbid. Prior research on the effect of co-occurring diseases on mortality among HD patients often considers disease counts or clusters at the time of HD diagnosis, overlooking the dynamic development of patients' disease portfolios over time where new chronic diseases are diagnosed prior to death. Further, these studies do not consider interactions among diseases and between diseases, biological and socioeconomic variables, which is essential for addressing health disparities among HD patients. Therefore, a mapping of the effect of combinations of these co-occurring diseases on mortality among HD patients that considers such interactions in a dynamic setting is warranted.

Objective: This study aimed to examine the effect of HD patients' co-occurring diseases on mortality, modelling their dynamically expanding disease portfolios while identifying interactions among the co-occurring diseases and socioeconomic and biological variables.

Methods: The study utilized electronic health record data from the national Danish registries and algorithmic diagnoses of 15 chronic diseases to obtain a study population of all 766,596 adult HD patients in Denmark from the 1st of January 1995 until the 31st of December 2015. The time from HD diagnosis until death was modelled by an extended Cox model with time-varying covariates. We identified interactions among co-occurring diseases and socioeconomic and biological variables in a data-driven manner. We mapped the mortality impact of 1) the most common disease portfolios, 2) the disease portfolios subject to the highest level of interactions, and 3) the most severe disease portfolios. Estimates from interaction-based models were compared to an additive model.

Results: Cancer had the highest impact on mortality (hazard ratio (HR) 6.72 males, 7.59 females). Excluding cancer revealed schizophrenia and dementia as those with the highest mortality impact (top five HRs in the range 11.72-13.37 males, 13.86-16.65 females, for combinations of four diseases). The additive model underestimated the effects up to a factor of 1.4 compared to the interaction model. Stroke, osteoporosis, COPD, dementia, and depression were identified as chronic diseases involved in the most complex interactions of the highest order.

Conclusions: The findings of this study emphasize the importance of identifying and modelling disease interactions for a

comprehensive understanding of mortality risk in heart disease patients.

(JMIR Preprints 26/02/2024:57749)

DOI: <https://doi.org/10.2196/preprints.57749>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in a JMIR publication, my accepted manuscript PDF will be visible to all users.

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Co-occurring diseases and mortality in patients with chronic heart disease modelling their dynamically expanding disease portfolio: A nationwide register study

Nikolaj Normann Holm ^a, Anne Frølich ^{b, c}, Helena Dominguez ^d, Kim Dalhoff ^{e, f}, Helle Gybel Juul-Larsen ^g, Ove Andersen ^{f, g, h}, Anders Stockmarr ^a

^a Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark

^b Innovation and Research Centre for Multimorbidity, Slagelse Hospital, Slagelse, Denmark

^c Department of Public Health, University of Copenhagen, Copenhagen, Denmark

^d Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark

^e Department of Clinical Pharmacology, Bispebjerg and Frederiksberg University Hospital, Copenhagen, Denmark

^f Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

^g Department of Clinical Research, Copenhagen University Hospital Amager and Hvidovre, Hvidovre, Denmark

^h Emergency Department, Copenhagen University Hospital Amager and Hvidovre, Hvidovre, Denmark

Abstract

Background

Medical advances in managing chronic heart disease (HD) patients permit co-occurrence of other chronic diseases due to increased longevity, causing them to become multimorbid. Prior research on the effect of co-occurring diseases on mortality among HD patients often considers disease counts or clusters at the time of HD diagnosis, overlooking the dynamic development of patients' disease portfolios over time where new chronic diseases are diagnosed prior to death. Further, these studies do not consider interactions among diseases and between diseases, biological and socioeconomic variables, which are essential for addressing health disparities among HD patients. Therefore, a mapping of the effect of combinations of these co-occurring diseases on mortality among HD patients that consider such interactions in a dynamic setting is warranted.

Objectives

This study aimed to examine the effect of HD patients' co-occurring diseases on mortality, modelling their dynamically expanding portfolios of chronic diseases while identifying interactions among the co-occurring diseases and socioeconomic and biological variables.

Methods

The study utilized electronic health record data from the national Danish registries and algorithmic diagnoses of 15 chronic diseases to obtain a study population of all 766,596 adult HD patients in Denmark from the 1st of January 1995 until the 31st of December 2015. The time from HD diagnosis until death was modelled by an extended Cox model with chronic diseases and interactions involving them as time-varying covariates. We identified interactions among co-occurring diseases and socioeconomic and biological variables in a data-driven manner using a hierarchical forwards-backwards selection procedure and stability selection. We mapped the mortality impact of 1) the most common disease portfolios, 2) the disease portfolios subject to the highest level of interactions, and 3) the most severe disease portfolios. Estimates from interaction-based models were compared to an additive model.

Results

Cancer had the highest impact on mortality (hazard ratio (HR) 6.72 males, 7.59 females). Excluding cancer revealed schizophrenia and dementia as those with the highest mortality impact (top five HRs in the range 11.72-13.37 males, 13.86-16.65 females, for combinations of four diseases). The additive model underestimated the effects up to a factor of 1.4 compared to the interaction model. Stroke, osteoporosis, chronic obstructive pulmonary disease, dementia, and depression were

identified as chronic diseases involved in the most complex interactions, which were of the fifth order.

Conclusions

The findings of this study emphasize the importance of identifying and modelling disease interactions to gain a comprehensive understanding of mortality risk in HD patients. The study illustrated that complex interactions are widespread among the heart disease patients' co-occurring chronic diseases. Failing to account for these interactions can lead to an oversimplified attribution of risk to individual chronic diseases, which may, in cases of multiple co-occurring conditions, result in an underestimation of mortality risk.

Keywords: Survival Analysis, Interaction effects, Chronic heart disease, Multimorbidity, Time-varying covariates

Introduction

Driven by the advancements in diagnostic tools and medical treatments, the mortality of patients with chronic heart disease (HD) has decreased considerably [1]. However, with a prolonged lifespan comes a risk of developing additional chronic diseases and complications to their HD [2], causing them to become multimorbid [3]. Multimorbidity is highly prevalent among HD patients [2,4,5], and the increasing disease burden may modify time to death [6].

Recent research has identified the most prevalent comorbidities in HD patients and how they affect mortality and other adverse health-related outcomes [5,7–9]. However, only a few studies have considered the effect of several diseases in the same person. Among these, there is a large variety in which diagnoses are considered and the statistical methods applied. The studies that considered multimorbidity either restricted their analyses to a subset of diagnosis combinations [7] or grouped diagnoses into multimorbidity clusters at baseline before analyzing the effects of the extracted clusters [5]. Despite modelling disease interactions, these kinds of analyses fail to capture the crucial dynamics in the HD disease trajectories, where additional diseases are cumulatively diagnosed before death [10], causing an augmented risk profile for the patient. As the chronology of disease onset has been associated with a change in mortality among common diagnoses [11], it is thus essential to consider this dynamic development when analyzing effects. Due to the high prevalence of multimorbidity among HD patients, the unique combination of chronic diseases a patient has at any given time — referred to as the 'disease portfolio' — is not static. Instead, it evolves over the observation period as new chronic diseases develop. This dynamic expansion reflects the progressive accumulation of chronic diseases in an individual following their HD diagnosis until death. As only a few studies consider these dynamics, there is a need for a thorough, large-scale study of the impact of disease interactions on mortality, modelling such a dynamic expansion of the patients' disease portfolio. Such an investigation would enable obtaining a deeper understanding of how the complexity of disease progression in patients with HD affects mortality over time.

The significance of understanding the effects of the emergence of multimorbidity over an individual's lifespan has previously been highlighted [3,12,13]. However, rather than treating multimorbidity as a singular risk factor, we take a more nuanced approach by dissecting the effects of multimorbidity based on the diseases appearing in the disease portfolio, recognizing that each combination of chronic diseases can affect mortality differently. Furthermore, as many chronic diseases have similar biological and socioeconomic risk factors, knowledge of the interplay between the impact of these is essential and can be utilized for possible preventive interventions and the development of guidelines for relevant coexisting diseases [14,15]. For instance, consider a disease portfolio comprising HD and osteoporosis. The impact on the mortality hazard rate may vary between men and women.

Expanding on this example, the effect of socioeconomic position may differ depending on both sex and the presence of osteoporosis in the portfolio. These variations in effects represents interactions in modelling terms. As such, identifying and emphasizing interactions among chronic diseases and demographic factors can shed new light on the impact of pathophysiological pathways on mortality.

The present large-scale study is based on data from the total adult Danish population recorded in nationwide primary and secondary healthcare registries, including medical diagnoses, medications, educational attainment level and healthcare utilization. We utilized an extended Cox model with time-varying covariates to model time until death for HD-diagnosed individuals, considering their dynamically expanding disease portfolio. In our model, the hazard ratio (HR) of a disease portfolio is constant. In contrast, the HR of an individual changes dynamically when the individual gets a new portfolio by obtaining a new chronic disease (see Fig. 1).

We performed a model and data driven selection of interaction effects. Subsequently, we studied the impact on time to death according to a) the most frequently occurring disease portfolios, b) the most complex disease portfolios in terms of order of interactions, and c) the disease portfolios with the highest hazards relative to solely HD.

We recognize the inherent complexity in interpreting interaction effects, especially in the case of higher-order interactions involving multiple factors. However, to emphasize the importance of modelling interaction effects, we also present a comparative analysis of effect estimates for disease portfolios, contrasting our interaction model with a simpler model where interactions are excluded. The differences observed in these comparisons serve to underscore the crucial role of modelling interactions in medical research.

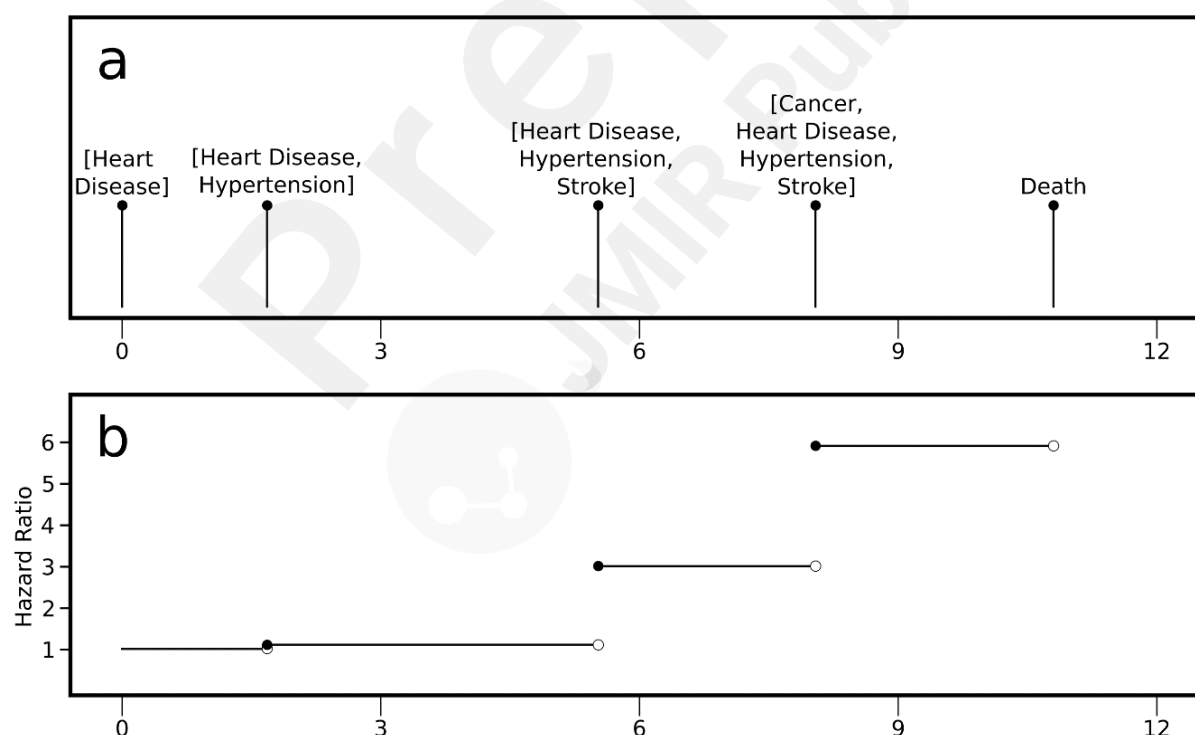


Fig. 1 Example of how the statistical model works. (a) Illustration of an event sequence where a hypothetical HD patient gets the diagnosis of HD at time 0 and subsequently hypertension, stroke and cancer diagnoses at different times (measured in years following HD) before death. (b) The corresponding longitudinal development of the hazard ratio of the subject, relative to a theoretical non-multimorbid patient who solely has HD.

Throughout this paper, we use a bracket notation to represent the disease portfolio of a specific HD

patient. For example, a patient with HD, diabetes and hypertension is denoted by the portfolio [diabetes, hypertension]. If the HD patient also has high cholesterol, their disease portfolio is [diabetes, high cholesterol, hypertension]. As all individuals in the study population have HD, we use the term disease portfolio without mentioning the coexisting HD diagnosis in the notation. We use the terms dyads, triads, tetrads, and pentads to describe disease portfolios of size two, three, four and five, with size being the number of chronic diseases in the portfolio including the HD.

Methods

Data Foundation

All children born in Denmark, or any new resident are, by law, required to get a unique personal identification number, which is stored in the Danish Civil Registration System (CRS) [16]. The personal identification number can link information from any additional Danish register at an individual level subject to GDPR restrictions [17]. Information about chronic diagnoses were based on diagnostic algorithms initially developed by the Research Center for Prevention and Health at Glostrup University Hospital [18]. They cover 15 diagnoses based on their clinical relevance which have previously been used in national reports of chronic diagnoses in Denmark [19,20]. Moreover, previous work with the diagnoses has shown prevalence results comparable to other European studies [21]. The algorithmic diagnoses are based on data recorded in four registries: the Danish National Patient Register (NPR) [22], the Danish Psychiatric Central Research Register (PCRR) [23], the Danish National Prescription Registry (DNPR) [24] and the Danish National Health Service Registry (NHSR) [25]. Therefore, a particular diagnosis can be given at a particular time (with temporal granularity of days) based on criteria for hospitalization diagnoses, medication, or repeated use of specific health services. As such, a single diagnosis corresponds to one disease and represents multiple ATC and/or ICD-10 codes with similar treatments and organization of health care. The diagnosis timestamps considered in this study are, thus, diagnostic timestamps and should not be regarded as timestamps for disease onsets. In addition to the registries used for diagnostic timestamps, we utilized the Danish Population Education Register [26] and the Danish Register of Causes of Death [27] for information on educational attainment and death.

Study Design and Population

Using our data foundation of the Danish registries, we obtained a study population of HD-diagnosed individuals covering the entire Danish adult population (aged ≥ 18 years) at some point during the observation period from the 1st of January 1995 to the 31st of December 2015, which had previously been created [28]. These people were followed, and data associated with visits to outpatient clinics, hospital stays, primary sector health services and prescriptions were collected for each person throughout the observation period. To define the study population, we applied algorithmic diagnoses (detailed in Supplemental Table 1) to identify individuals diagnosed with HD while determining diagnostic timestamps for 14 additional, selected chronic diseases [21]. Our inclusion criterion was thus broad, encompassing all Danish adults (aged ≥ 18 years) who received an algorithmic diagnosis of HD during the study period. No additional exclusion criteria were applied. Our outcome was time until death of any cause after the HD diagnosis.

Statistical Analysis

The prevalence of each of the chronic diseases was calculated at the time of HD diagnosis across all patients in the population. Similarly, we calculated prevalences of the diseases throughout the entire observation period by considering if the condition at all had occurred among the HD patients.

The data were analyzed within a survival analysis framework, with years following HD diagnosis as the time variable and an event defined as all-cause mortality. As such, we denoted the HD diagnosis time $t=0$ and aligned our timescale accordingly, meaning time $t=0$ correspond to potentially different age times and calendar times for distinct individuals. In addition, individuals lost to follow-up due to emigration or, due to reaching the end of the observation period, were censored at these times.

The time-varying information on individual diagnoses, information on sex (male or female), age, educational attainment level (none, short, medium, long, missing and missing pre-1920) and calendar time were included as explanatory variables in the analysis (see Tables 1 and 2 in [28]). We utilized an extended Cox model to estimate the effect of these explanatory variables on mortality, allowing for the inclusion of time-varying covariates. We classified our variables into primary and intrinsic categories [29]. Primary variables, such as the time-varying diagnosis indicators, cover variables of paramount interest. Intrinsic variables define the study individuals, i.e., the variables sex, age, educational attainment level and calendar time. Interactions both between and within each group of variables were considered. The numerical variables were mean-centered prior to analysis.

As the development of additional diagnoses is a continuous process, the primary variables were allowed to change over time. These variables were piece-wise constant in time, being zero when the diagnosis was not present and one when obtained and onward in time. Since the registries continuously cover clinical events for all individuals over the observation period, these diagnosis variables update at individual-specific time points dictated by the (sequence of) event(s) that trigger the algorithmic diagnosis (Supplemental Table 1). An example of a potential sequence of diagnoses is showcased in Fig. 1.

In the extended Cox proportional hazards model [30], the hazard h_i for the i 'th individual at time t is given by:

$$h_i(t, X_i(t)) = h_0(t) \exp\left(\sum_j \beta_j X_{ij}(t)\right) \quad (1)$$

Where $h_0(t)$ is the unspecified baseline hazard function for a male with no education without any diagnoses except the HD. $X_{ij}(t)$ denotes the variable j for individual i (with $X_i(t)$ denoting a vector of all variables) at time t , with $i=1, \dots, n$. The β_j are the effect parameters. Due to $h_0(t)$ being unspecified, these parameters are linked to the relative mortality hazard rate of a variable, as opposed to the absolute risk. Equation 1 assumes that variables have proportionate effects on the hazard function over time. We assessed this assumption for each variable by examining Schoenfeld residuals [31]. In addition, as the effect parameters β_j do not depend on time, the hazard rate associated with a particular combination of explanatory variables is assumed the same across all timepoints.

Selection of Variables and Interactions

It is essential to account for diagnosis interactions as such parameters serve to model the entire effect of disease portfolios associated with mortality. Possible omitted interaction effects from a model where a significant interaction exists can result in a misrepresentation of the relationship between the variables and the time until death. It may also lead to bias in parameter estimation [32,33].

A common way to perform variable selection is a backwards selection approach starting from a full model considering all possible interactions, reducing it to a model which best explains the observed data. However, such an approach was not computationally feasible as we are in a big-data setting with numerous observations and countless potential variable interactions. Instead, we considered two variations of a forward-backwards selection procedure to discover disease interactions. As a sensitivity analysis, we also performed variable selection using the stability selection methodology [34] with the regularization-based LASSO [35] approach.

In addition to the models including interaction effects, a model solely consisting of the primary and intrinsic variables' main effects (and squared and cubic terms) was estimated for reference.

We considered k -way interactions iteratively for $k=2,\dots,M$, with M being a predetermined upper limit. The selection procedure starts from an initial model including all main effects and works in the

- Generate n_c candidate variable additions obtained from the current model by adding a single k -way interaction to an already existing $(k-1)$ -way interaction, also adding necessary lower hierarchical terms.
- Repeat until no candidate models below the cutoff:
 - Estimate each of the candidate models obtained from adding any of the n_c variable additions not already added to the current model and compare with the current model by a likelihood-ratio test.
 - Select the candidate model with the lowest p-value below the cutoff $\frac{\alpha}{n_c}$ as

following way for each k :

The selection algorithm either runs until M -way interactions are included or until no k -way interactions are selected in the k 'th iteration. In the forward step of the selection algorithm, a Bonferroni-adjusted cutoff is utilized to minimize the risk of false discoveries, as each variable addition is potentially tested for inclusion n_c times. We note that all considered models are hierarchical, meaning that if a model contains a five-way interaction among five variables, it also contains all possible four, three and two-way interactions among those variables.

Due to the allowance of *any* k -way interaction between and among the primary and intrinsic variables, a possibly large number of candidate models were included for each k . Because of this, the selection forward step was relaxed, such that the candidate model p-values were ordered from lowest to highest after the first estimation for each k . In the following estimations, variable additions were checked in this order, immediately adding any interactions below the cutoff while discarding insignificant terms. Before backwards selection, any discarded terms were included again through forward selection. To introduce conservatism, all variable selection was performed with $\alpha=0.001$. The resulting model with all selected interactions was labelled as the ALL model.

In addition to the ALL model, the variable selection procedure was run without relaxation of the forward step but only considering interactions among the primary variables. We labelled this the disease interactions only (DIO) model. Furthermore, we used a variation of the stability selection framework [34], a method for improving variable selection in high-dimensional, sparse environments. This method selects variables repeatedly chosen on subsampled data by a structure learning method, such as the LASSO algorithm for the Cox model [36]. We used a selection threshold of 0.9, following the recommendation in [34]. Each subsample included ten randomly selected variables, considering all their possible interactions up to an order of five. This caused us to

consider 3400 subsamples in total. We then fit an unregularized Cox model using the stably selected terms and performed backward selection to reduce the model using all available data. The resulting model was labelled the stable (S) model. As a sensitivity criterion, we compare detected interactions among the chronic diseases across the ALL, DIO and S models. The additive model only including main effects was labelled the only main effects (OME) model.

Selecting Disease Portfolios

Due to the many possibilities when considering combinations of the 14 co-occurring diseases, some of our presented results are based on selected disease portfolios. These selections were made based on three criteria: Most common disease portfolios, disease portfolios subject to the highest order of disease interactions and disease portfolios with the highest mortality impact. The main results presented in this paper are based on the ALL model. To illustrate the importance of modelling interaction effects, the effect of specific disease portfolios in the ALL model was compared to additive effects from the OME model on the log-hazard scale.

Scenarios

As the considered diagnosis variables were subject to higher-order interactions, effects are not apparent just from the estimated parameters because the effect of a single diagnosis varied across different levels of other diagnoses and intrinsic variables. To supplement the effect of selected disease portfolios, the absolute mortality risk over time was estimated for multiple scenarios utilizing the estimated ALL model. We did this to illustrate the modification of an HD-diagnosed individual's risk profile over time. Each scenario represents the risk of a hypothetical individual whose disease portfolio expands at predetermined time points following HD diagnosis. The times at which the disease portfolio expanded in the hypothetical scenarios were determined in a data-driven fashion by gamma regressions, where the time points (at which the first, second, or third expansion of the disease portfolio following HD occurred) were regressed on the diagnoses in the sequence considered in the scenario. The scenarios were constructed for patients who received their HD at mean age and calendar time levels.

Ethical Considerations

In this study, we used data from the national Danish registries which are protected by the Danish Data Protection Act, meaning that they can only be accessed after application and subsequent approval. The study did not require additional approval from the Danish Research Ethics Committees or any informed consent as it solely involved use of national registry data, exempt under the Scientific Ethical Committees Act. Danish registry data are de-identified to protect the privacy of individuals.

Results

Characteristics of the Study Population

A total of 766,596 HD-diagnosed individuals were included (53.1% males). The mean age at the time of the HD diagnosis was 67.5 years for males and 73.0 years for females (further baseline characteristics available in [28], Table 2). At the end of the observation period, 58.0% were dead (54.6% males, 61.7% females). Overall, the prevalence of multimorbidity in the complete trajectories of each HD patient was 96.9%. This was an increase compared to the multimorbidity prevalence at time $t=0$ (86.3%). The prevalence of each of the 14 co-occurring diseases is presented in Fig. 2. Overall, hypertension, high cholesterol and allergies were among the most prevalent diseases in the

HD population, with a lifetime prevalence of 81.2, 44.9 and 28.9%, respectively (Fig. 2, Supplemental Table 2). Differences in prevalence by sex were large for some chronic diseases, particularly for osteoporosis and depression commonly occurring in females.

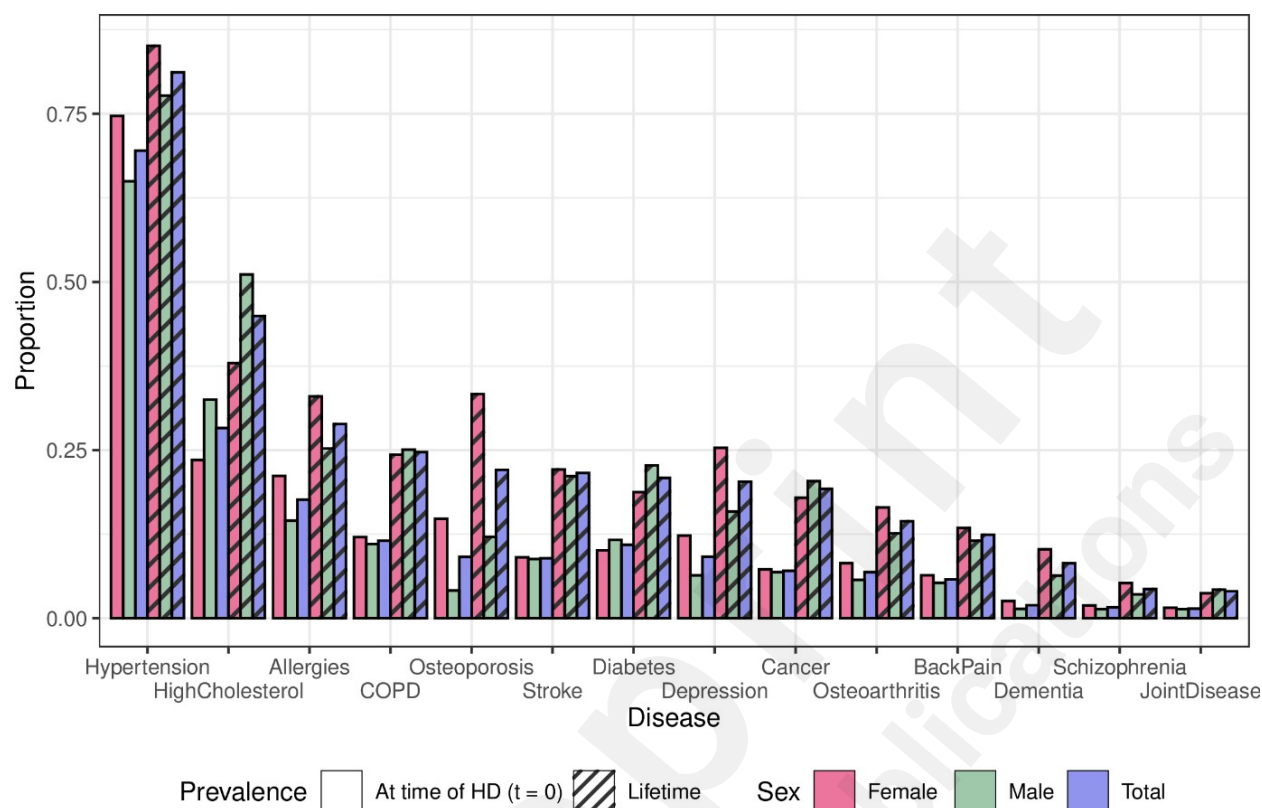


Fig. 2 Diagnosis prevalence according to sex. Prevalence is reported at the time of HD diagnosis and for the entire span of the observed disease trajectories (Lifetime).

Interactions

Following the inclusion of five-way interactions, the ALL model selection procedure terminated due to no six-way interactions being selected. All the primary and intrinsic variables were present in the final model. Fig. 3 illustrates statistically significant interaction relationships between chronic diseases detected in the ALL model. Connections between diseases in the ribbon chart illustrate the two chronic diseases appearing in an interaction, with the colour depicting the complexity of the interaction (darker color represents a higher-order interaction). The figure reveals all diseases interacting, with some diseases involved in more complex disease interactions than other chronic diseases. In total, 288 interactions were present in the final model. The interaction relationships between the considered diseases were highly diverse but dominated by cancer, which had statistically significant interactions with all other diseases. The diseases depression, stroke, chronic obstructive pulmonary disease (COPD), dementia and osteoporosis were involved in the most complex interactions, as they were the sole diseases involved in five-way interactions. Some of the most prevalent diseases, allergies and hypertension, were not part of these complex relationships.

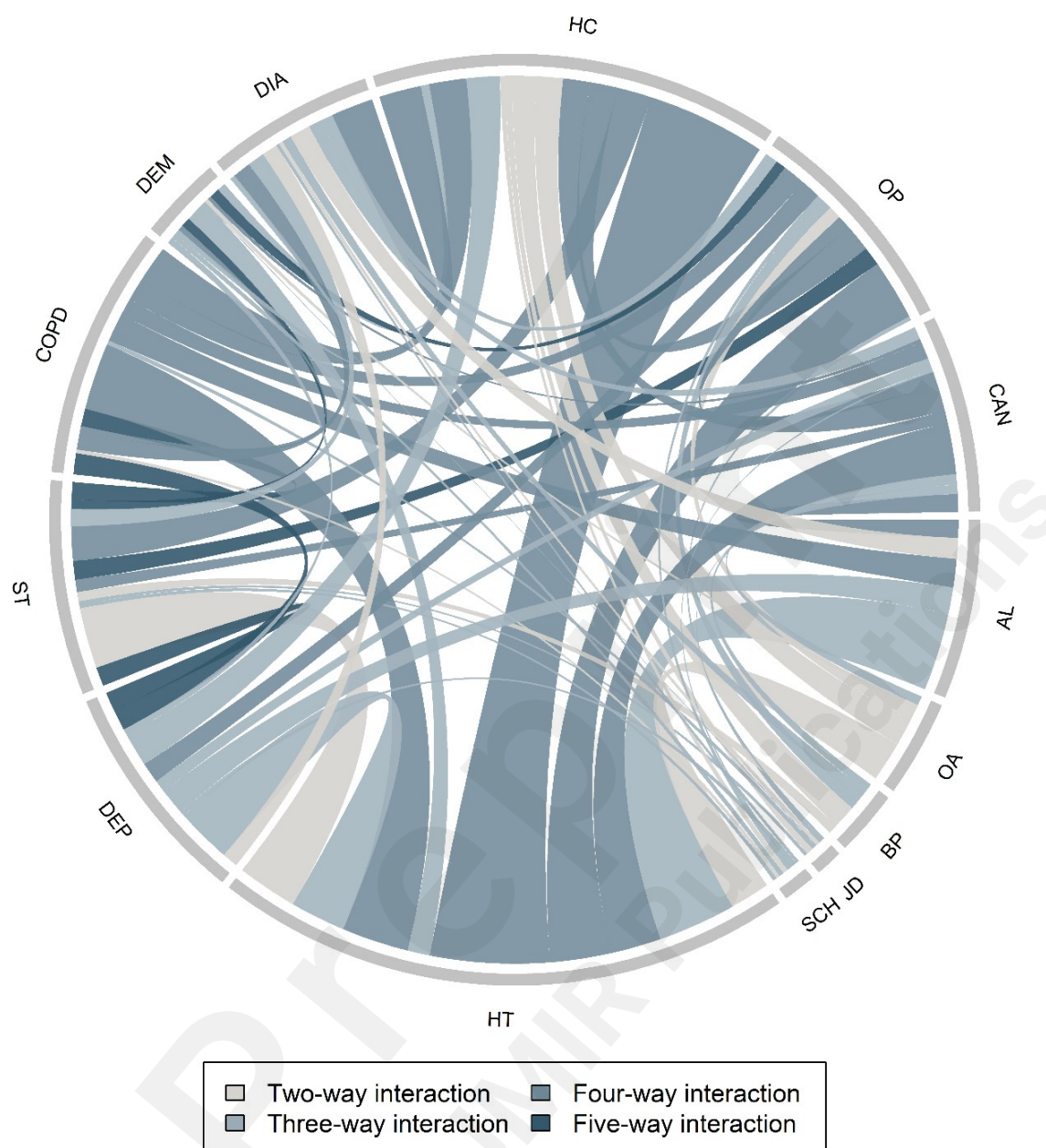


Fig. 3 Graphical representation of disease-disease interactions in the all interactions (ALL) model. A ribbon connects chronic diseases that have any significant interaction ($p \leq 0.001$) between them. The connection's width corresponds to the number of HD-diagnosed individuals obtaining both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between two diseases. The ribbon chart is ordered by number of connections between diseases, starting from AL with five connections all the way to CAN which interacts with all of the additional diseases.

AL = allergies, BP = back pain, JD = joint disease, OA = osteoarthritis, HT = hypertension, SCH = schizophrenia, DEP = depression, HC = high cholesterol, DIA = diabetes, ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, CAN = cancer.

The chronic disease allergies was part of five interaction relationships with other diseases, involving two four-way, two three-way and a single two-way interaction. Hypertension interacted with nine other diseases, involving four four-way, three three-way and two two-way interactions. Notably, dementia and depression appeared in higher-order interactions (two five-way interactions, respectively) despite having fewer co-occurrences in the population. Similar patterns were observed for the DIO and S models (Supplemental Fig. 1, Supplemental Fig. 2). In both of these models, COPD, dementia, stroke, and depression were involved in interactions of the highest order. The DIO

model included up to five-way interactions, also featuring complex interactions involving the chronic diseases diabetes and cancer (Supplemental Fig. 1). For the S model, only up to four-way interactions were detected (Supplemental Fig. 2). In general, most of the interactions between diseases identified in the ALL model were also present in the DIO and S model (Supplemental Table 5).

Effects

Difference in Effect Estimates by Disease Portfolio Size

To evaluate how the effects of disease portfolios on time until death differ between models with and without interactions, we calculated the effect differences between the OME model and the ALL model on the log hazard rate scale, denoted as $\hat{\theta}_{ALL} - \hat{\theta}_{OME}$. We focused on disease portfolios ranging from two to eight diseases, as these account for 99% of all disease portfolio observations of size two or greater. The effects in the ALL model for each disease portfolio were computed at mean age and calendar time levels, aggregating over combinations of both sexes and all educational attainment levels. To compute an overall estimate of the effect differences between the models for each sex, we calculated a weighted mean of the differences for each portfolio size. The weights were determined by the prevalence of individual disease portfolios across the different education levels for each sex. In Fig. 4, the aggregated differences are displayed on the hazard scale, indicating the multiplier required to convert the HR from the OME model into the HR from the ALL model. The figure illustrates substantial variations in disease portfolio effects when interactions are excluded compared to when they are included. The HR multiplier increases gradually for disease portfolios of increasing size, flattening around 1.4 at disease portfolios of size six. In general, for disease portfolios of size two, the HRs are on average slightly overestimated when interactions are not modelled. However, for disease portfolios of size four and greater, the HRs are on average underestimated for both sexes. The underestimation also applies to females with disease portfolios of size three. In general, the HR multiplier was slightly greater for females, compared to males, across all disease portfolios.

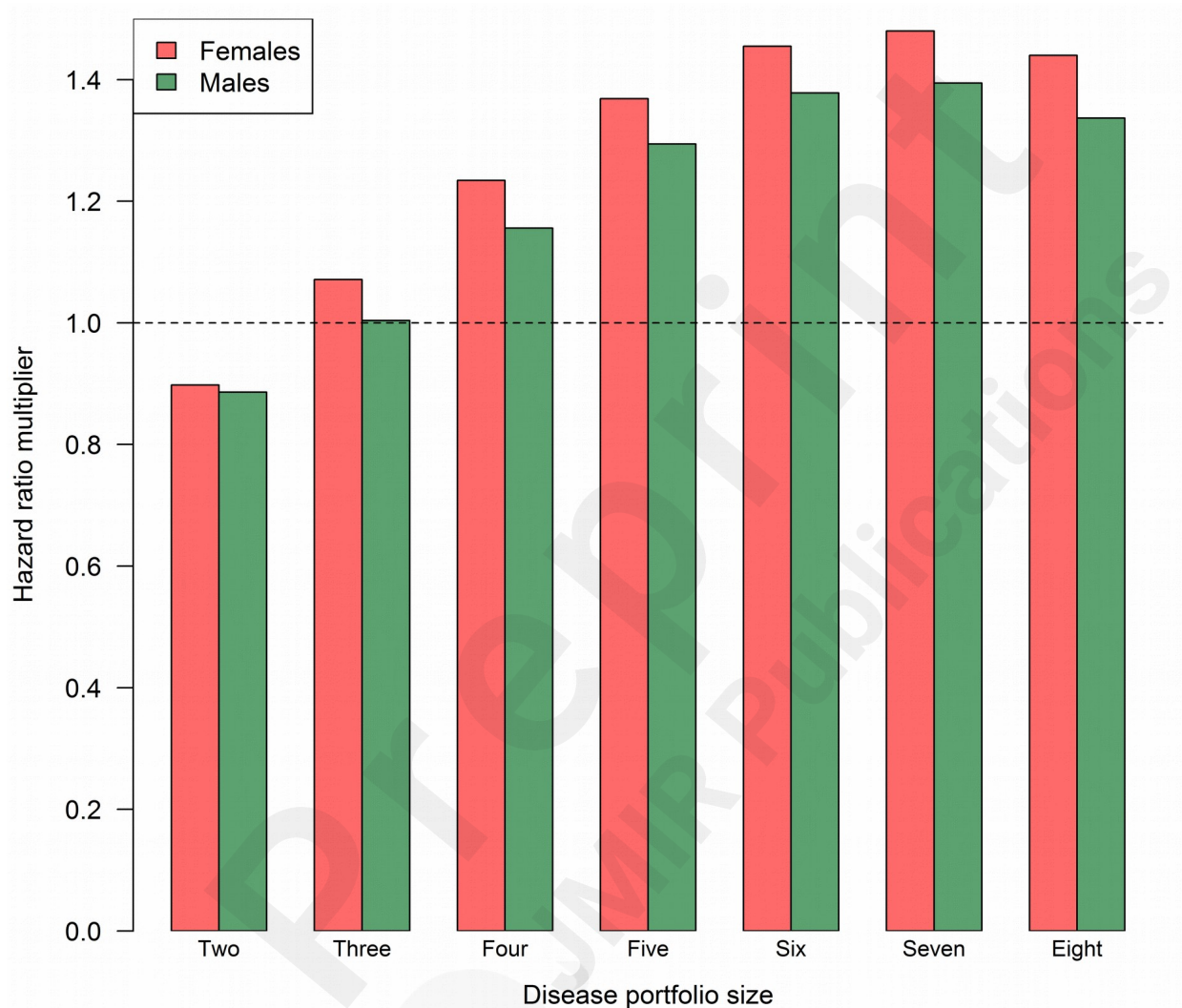


Fig. 4 Difference in effect estimates for disease portfolios of increasing sizes for females and males. Each bar represents a weighted average of the differences in effects between the additive OME model and the ALL model on the hazard scale $\exp(\hat{\theta}_{ALL} - \hat{\theta}_{OME})$. The bars thus indicate the average multiplier required to convert the HR from the OME model into the HR from the ALL model. The weights were determined based on the occurrence of each specific disease portfolio across the different educational attainment levels for each sex.

Most Frequent Disease Portfolios

The effects of the ten most frequent disease portfolio dyads, triads, tetrads and pentads are presented on the log hazard scale at increasing educational attainment levels for males in Fig. 5 and for females in Fig. 6, based on the ALL model. The associated HR estimates are presented in Supplemental Table 3 and Supplemental Table 4. Disease portfolios including high cholesterol and allergies were of particular concern, as many of them had a negative effect, corresponding to a decreased mortality

hazard rate relative to an HD-diagnosed individual who is not multimorbid. By comparing effects from the disease portfolios from the ALL model to effects from the OME model, generally, the direction of the effect (positive or negative) agreed between the models for both males and females. However, the magnitude of the effects were greater in the ALL model compared to the OME model for almost all disease portfolios, educational attainment levels and sexes. This indicates an underestimation of the risk associated with a disease portfolio for the positive effects, and an overestimation for the negative effects. For some disease portfolios, an inverse social gradient was visible in the educational dimension, where the higher the educational attainment level, the greater the effect of the disease portfolio (see e.g. the portfolio [diabetes, hypertension] in Fig. 5). Sex-related disparities in disease portfolio effects were also evident. For disease portfolios containing depression and osteoporosis, the effects of the portfolios were greater for males compared to females, whereas for COPD, cancer, stroke and diabetes, the effects were greater among females.

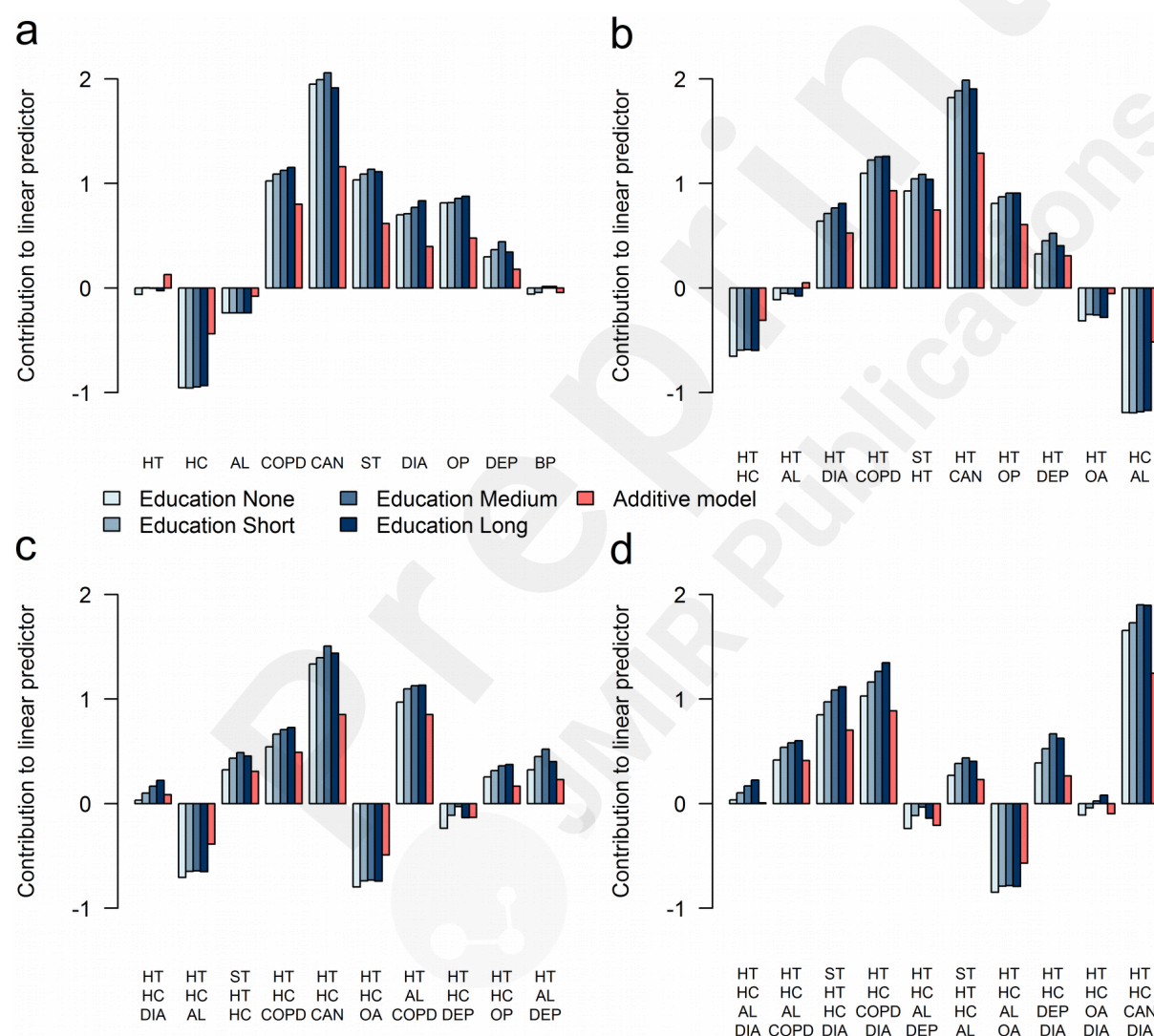
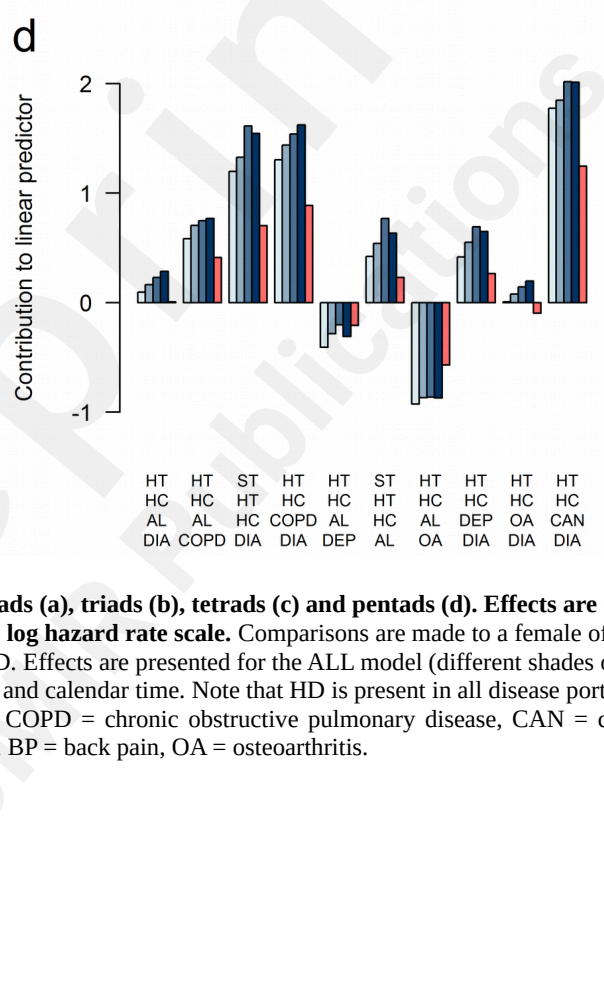


Fig. 5 Effects of the ten most frequent disease portfolio dyads (a), triads (b), tetrads (c) and pentads (d). Effects are displayed for males of varying educational attainment levels at the log hazard rate scale. Comparisons are made to a male of the corresponding educational attainment level who solely has HD. Effects are presented for the ALL model (different shades of blue) and the OME model (red). All comparisons are made at mean age and calendar time. Note that HD is present in all disease portfolios. HT = hypertension, HC = high cholesterol, AL = allergies, COPD = chronic obstructive pulmonary disease, CAN = cancer, ST = stroke, DIA = diabetes, OP = osteoporosis, DEP = depression, BP = back pain, OA = osteoarthritis.



Most complex disease portfolios

Fig. 7 shows the effects of disease portfolios containing combinations of stroke, osteoporosis, COPD, dementia and depression for males with differing educational attainment levels. These chronic diseases were all part of five-way interactions, making the effects associated with their portfolios the most complex. For dyads, triads, tetrads and pentads, the OME model generally yielded lower effects than the ALL model. This implies an underestimation of mortality risk in males for these portfolios when interactions are not modelled. The underestimation was greatest for disease portfolios involving dementia or stroke. Similar results were observed for females, but also included a large underestimation of mortality hazard rates for portfolios involving COPD (see Supplemental Fig. 3)

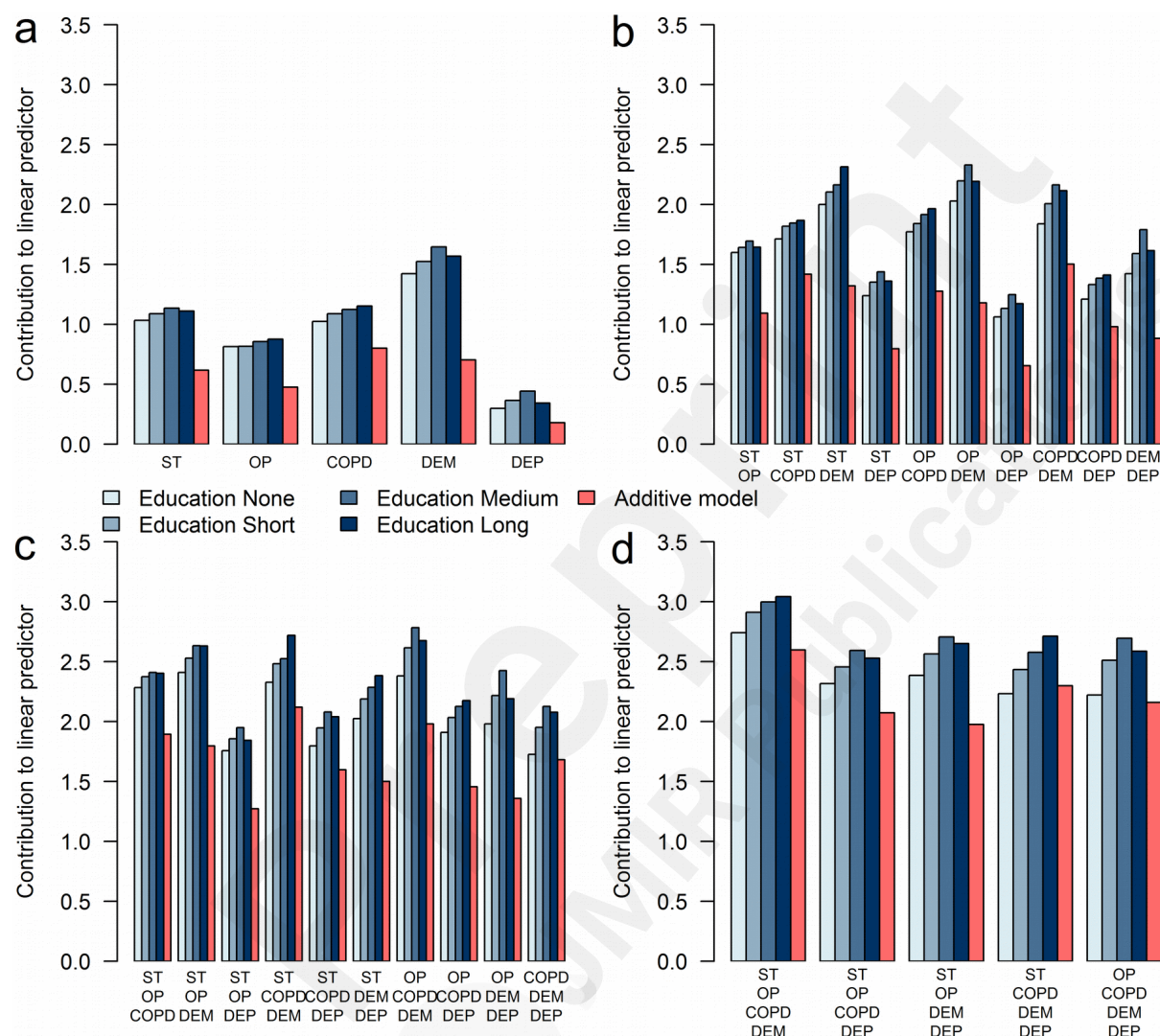


Fig. 7 Effects of disease portfolio dyads (a), triads (b), tetrads (c) and pentads (d) involving stroke, osteoporosis, COPD, dementia and depression. Effects are displayed for males of varying educational attainment levels at the log hazard rate scale. Comparisons are made to a male of the corresponding educational attainment level who solely has HD. Effects are presented for the ALL model (different shades of blue) and the OME model (red). All comparisons are made at mean age and calendar time. Note that HD is present in all disease portfolios.

ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, DEP = depression.

Disease Portfolios with Highest Mortality Impact

Table 1 presents the largest HRs for disease portfolio dyads, triads and tetrads among males and females. Generally, the HRs of the disease portfolios were greater in females, however the portfolio [schizophrenia] exhibited a greater HR in males. For dyads, the portfolios [cancer], [dementia], [schizophrenia], [stroke] and [COPD] rank within the top five for both sexes. Notably, [cancer] exhibits the largest hazard ratio (HR 6.72 males, 7.59 females). When considering triads and tetrads, cancer is similarly consistently featured in the top five portfolios for both sexes. This indicates that

cancer contributes to a greatly increased relative mortality risk whenever present. Among triads, the portfolio [cancer, schizophrenia] had the largest hazard ratio for males (HR 13.26) and the second largest for females (HR 13.38). The top-ranking portfolio for females was [cancer, COPD] (HR 15.39), while for males it was the second largest (HR 11.34). Notably, four out of the five tetrad portfolios with highest mortality impact include both cancer and COPD for males and females. Since cancer was consistently present in the triads and tetrads with highest mortality impact, we separately examined the triads and tetrads among portfolios without cancer. The results are presented in Table 2. Upon excluding cancer, we observed that portfolios including dementia and schizophrenia were prominent in most of the triads and tetrads with highest mortality impact. Among tetrads, the portfolios for males always involved osteoporosis paired with dementia or schizophrenia. In contrast, for females, the tetrads typically consisted of stroke in combination with dementia or schizophrenia.

Table 1 The five largest hazard ratios for dyads, triads and tetrad disease portfolios. The reference group comprises males or females with only HD. Hazard ratio estimates were aggregated on the log-hazard scale for males and females across all educational attainment levels, using weights corresponding to the number of individuals with each portfolio within that subpopulation. Portfolios with fewer than ten individuals were excluded.

CAN = cancer, COPD = chronic obstructive pulmonary disease, DEM = dementia, DIA = diabetes, OP = osteoporosis, SCH = schizophrenia, ST = stroke.

Males					Females				
Dyads									
Rank	Portfolio ^a	HR	CI (99.9%)	n ^b	Rank	Portfolio ^a	HR	CI (99.9%)	n ^b
1	[CAN]	6.72	(6.06-7.45)	6702	1	[CAN]	7.59	(6.83-8.43)	3559
2	[DEM]	3.99	(3.59-4.43)	1272	2	[DEM]	4.41	(3.98-4.89)	1180
3	[SCH]	3.04	(2.85-3.24)	888	3	[ST]	3.60	(3.27-3.97)	3386
4	[ST]	2.89	(2.66-3.14)	5722	4	[COPD]	3.57	(3.23-3.95)	4335
5	[COPD]	2.81	(2.55-3.10)	7884	5	[SCH]	2.74	(2.56-2.92)	663
Triads									
1	[CAN, SCH]	13.26	(11.50-15.29)	66	1	[CAN, COPD]	15.39	(13.51-17.53)	622
2	[CAN, COPD]	11.34	(9.89-12.99)	1356	2	[CAN, SCH]	13.38	(11.70-15.31)	58
3	[CAN, OP]	10.35	(9.01-11.90)	433	3	[CAN, DEM]	12.84	(10.60-15.56)	90
4	[CAN, DEM]	10.06	(8.38-12.07)	131	4	[CAN, ST]	12.65	(10.91-14.67)	296
5	[CAN, ST]	9.87	(8.59-11.35)	773	5	[CAN, DIA]	10.44	(9.24-11.80)	251
Tetrads									
1	[CAN, COPD, SCH]	19.21	(16.33-22.60)	28	1	[CAN, COPD, SCH]	24.10	(20.45-28.41)	14
2	[CAN, SCH, ST]	16.82	(14.14-20.01)	14	2	[CAN, COPD, DEM]	23.13	(17.89-29.91)	13
3	[CAN, COPD, OP]	16.40	(14.10-19.07)	157	3	[CAN, COPD, ST]	22.80	(18.84-27.59)	54
4	[CAN, COPD, ST]	15.92	(13.29-19.07)	168	4	[CAN, DEM, ST]	19.14	(15.03-24.37)	20
5	[CAN, COPD, DEM]	14.71	(11.59-18.67)	30	5	[CAN, COPD, OP]	17.57	(15.06-20.48)	168

^a All portfolios contain the HD diagnosis.

^b The number of unique males or females who has exactly this combination of diseases at any time during the observation period.

Table 2 The five largest hazard ratios for dyads, triads and tetrad disease portfolios excluding portfolios with cancer. The reference group comprises males or females with only HD. Hazard ratio estimates were aggregated on the log-hazard scale for males and females across all educational attainment levels, using weights corresponding to the number of individuals with each portfolio within that subpopulation. Portfolios with fewer than ten individuals were excluded.

COPD = chronic obstructive pulmonary disease, DEM = dementia, DIA = diabetes, SCH = schizophrenia, ST = stroke, OP = osteoporosis.

Males					Females				
Dyads									
Rank	Portfolio ^a	HR	CI (99.9%)	n ^b	Rank	Portfolio ^a	HR	CI (99.9%)	n ^b

1	[DEM]	3.99	(3.59-4.43)	1272	1	[DEM]	4.41	(3.98-4.89)	1180
2	[SCH]	3.04	(2.85-2.24)	888	2	[ST]	3.60	(3.27-3.97)	3386
3	[ST]	2.89	(2.66-3.14)	5722	3	[COPD]	3.57	(3.23-3.95)	4335
4	[COPD]	2.81	(2.55-3.10)	7884	4	[SCH]	2.74	(2.56-2.92)	663
5	[OP]	2.47	(2.26-2.69)	2341	5	[DIA]	2.31	(2.18-2.44)	2939
Triads									
1	[DEM, OP]	8.58	(7.49-9.84)	257	1	[ST, DEM]	9.77	(8.49-11.24)	268
2	[DEM, ST]	7.54	(6.58-8.65)	380	2	[COPD, DEM]	8.68	(7.36-10.24)	113
3	[COPD, SCH]	7.37	(6.58-8.24)	177	3	[COPD, SCH]	8.44	(7.52-9.47)	106
4	[DEM, SCH]	7.12	(6.34-8.00)	228	4	[ST, COPD]	8.42	(7.27-9.75)	324
5	[SCH, ST]	6.50	(5.80-7.28)	117	5	[OP, DEM]	7.96	(6.99-9.06)	649
Tetrads									
1	[DEM, OP, ST]	13.37	(11.32-15.78)	98	1	[COPD, DEM, ST]	16.65	(13.54-20.47)	32
2	[DEM, OP, SCH]	12.36	(10.46-14.61)	52	2	[DEM, DIA, ST]	15.04	(12.95-17.46)	35
3	[DEM, DIA, OP]	12.09	(10.21-14.31)	19	3	[COPD, SCH, ST]	14.79	(12.56-17.41)	12
4	[COPD, DEM, OP]	11.90	(10.00-14.16)	42	4	[DEM, OP, ST]	14.58	(12.36-17.20)	142
5	[COPD, OP, SCH]	11.72	(10.26-13.40)	26	5	[COPD, DEM, OP]	13.86	(11.51-16.70)	72

^a All portfolios contain the HD diagnosis.

^b The number of unique individuals who has exactly this combination at any time during the observation period.

Sex effect across socioeconomic subpopulations

The complex interactions at play indicate that the effect of sex on mortality varies by disease portfolio. This is illustrated in Fig. 8, which presents HRs comparing females to males across the 50 most prevalent disease portfolios at different education levels. Overall, the figure shows a decrease in female mortality risk compared to males, with most HRs falling below 1, ranging from 0.41 ([HT, AL, OP]) to 0.93 ([ST, HC, DIA] and [ST, HT, HC, DIA]). However, the magnitude of this decrease varies across comorbidity patterns. For example, portfolios that include osteoporosis consistently show HRs below 0.66, indicating a notable lower mortality risk for females with these portfolios, compared to males. Conversely, more complex disease portfolios that include stroke and diabetes—such as [stroke, hypertension, high cholesterol, diabetes] and [stroke, hypertension, diabetes]—have HRs closer to 1, suggesting only a slight reduction in female mortality hazard rate compared to males.

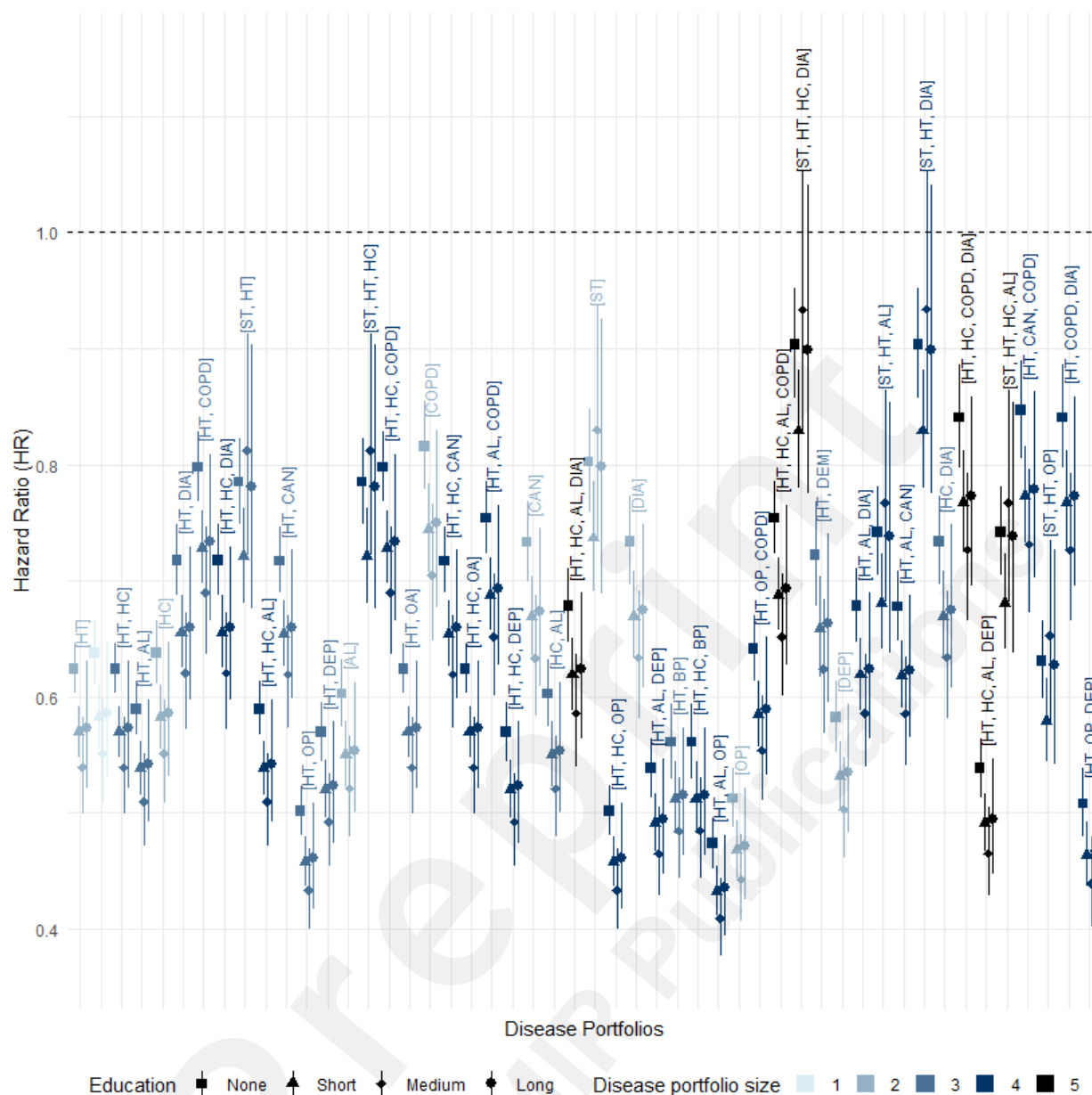


Fig. 8 Hazard ratios of female sex (vs male) by disease portfolio and educational attainment level. Estimates for the 50 most common disease portfolios are shown with 99.9% confidence intervals. The estimates are presented for each of the educational attainment levels: None, short, medium and long, indicated by different shapes and always in ascending order from none to long. The reference group comprises males with the same disease portfolio and educational attainment level. The disease portfolios are ordered by prevalence from left to right with [HT] the most frequent disease portfolio. Note that all portfolios contain the HD condition, so it is not labelled in the plot. Therefore, the disease portfolio without a label in the plot (the second from the left) correspond to the disease portfolio with only HD.

HT = hypertension, HC = high cholesterol, AL = allergies, DIA = diabetes, COPD = chronic obstructive pulmonary disease, ST = stroke, CAN = cancer, OP = osteoporosis, DEP = depression, OA = osteoarthritis, BP = back pain, DEM = dementia.

The impact of COPD

To illustrate that the effect of a single disease varies depending on the other diseases present in the portfolio, we estimated the effect of COPD in each observed disease portfolio in the population. The aggregated results are shown in for males and females of increasing disease portfolio size. The effect of COPD is greatest in triads (HR 2.81 males, 3.57 females), and generally higher in females compared to males. For increasing disease portfolio sizes, the aggregated effect of COPD decreases considerably with increasing disease portfolio sizes.

Table 3 Effect of COPD for increasing disease portfolio sizes. Each cell is the aggregated effect of chronic obstructive pulmonary disease (i.e. hazard ratio comparing the portfolio with and without COPD). The effects were aggregated on the log hazard scale with weights determined based on the occurrence of each specific disease portfolio across the different educational attainment levels for each sex.

		Disease portfolio size						
Sex		Two	Three	Four	Five	Six	Seven	Eight
Males HR		2.81	2.98	2.74	2.50	2.27	2.08	1.91
Females HR		3.57	3.77	3.43	3.08	2.75	2.45	2.19

Scenarios

We present four scenarios in Fig. 9 to illustrate the workings of the ALL model on the risk scale. In Fig. 9(a), we depict the first scenario, which consists of the trajectory of schizophrenia, followed by cancer, and then dementia. The figure illustrates an increase in the mortality rate with the additions of schizophrenia and cancer to the disease portfolio. However, when dementia is obtained, its involvement in interactions prevents a substantial increase in the mortality rate, compared to simply continuing undiagnosed. This is despite dementia being the disease with the second highest mortality impact when considered in isolation (HR 3.99 males, 4.41 females, Table 1). The interaction effects between the diseases in the portfolio and dementia create a situation where adding dementia does not further elevate the mortality hazard rate substantially.

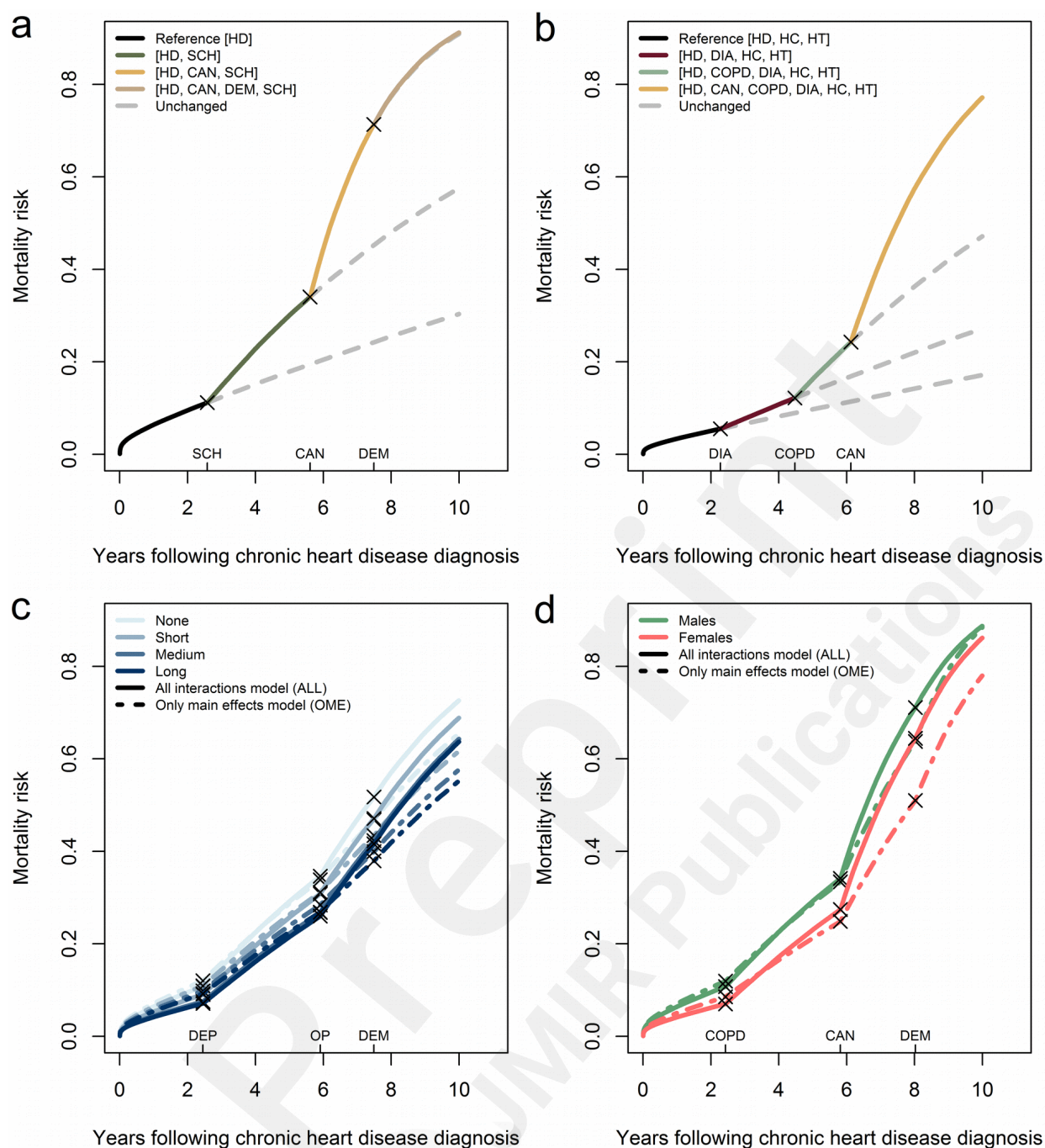


Fig. 9 Disease progression scenarios. Scenario representing the mortality risk over time of a hypothetical (a) male with no education, at mean age and calendar time who gets schizophrenia, cancer and dementia at 2.6, 5.6 and 7.5 years following HD. (b) male with no education who has hypertension, high cholesterol at time of HD and diabetes, COPD and cancer at 2.3, 4.8 and 6.1 years following HD. (c) male of varying educational attainment levels who gets depression, osteoporosis and dementia at 2.5, 5.9 and 7.5 years following HD under the ALL model (solid lines) and under the additive OME model (dot-dashed lines). (d) male (green color) and female (red color) with no education who gets COPD, cancer and dementia at 2.4, 5.8 and 8.0 years following HD under the ALL model (solid lines) and the additive OME model (dot-dashed lines). CAN = cancer, COPD = chronic obstructive pulmonary disease, DEM = dementia, DEP = depression, DIA = diabetes, OP = osteoporosis, SCH = schizophrenia.

Fig. 9(b) depicts a scenario that could resemble the disease trajectory of a male heavy smoker. In this scenario, the patient initially obtains HD while also having hypertension and high cholesterol. Over the following years, the patient receives a diabetes diagnosis, which further elevates the mortality risk. The risk accelerates even more with the addition of a COPD diagnosis and, finally, a cancer diagnosis. In Fig. 9(c) a scenario depicting the risk over time for a depression, osteoporosis and

dementia trajectory at different educational attainment levels for both the ALL and OME model is presented. A deviation between the ALL and OME model is most visible at the dementia disease, after which the risk in the ALL model accelerates compared to the OME model. Additionally, the scenario visualizes that despite the inverse social gradient of the disease portfolios on the log hazard scale (Fig. 7), lower educational attainment is still associated with a greater risk of death. Another scenario illustrating this relation is presented in Fig. 9(d) for a COPD, cancer and dementia trajectory. In the scenario we observe general increased mortality in males compared to females. However, due to the HRs of the disease portfolios being greater in females, compared to males (Table 1), the sex difference decreases over time.

Discussion

Patients with HD will often be diagnosed with other chronic diseases during their lifetime [2,5]. The effect of these co-occurring diseases on adverse outcomes is an important research focus, as it is a clinically emerging challenge. In the present study on the effect of disease portfolios on time until death, an extended Cox model allowing for time-varying covariates was applied to a large, longitudinal dataset encompassing all adult Danish HD patients in the period 1995-2015. We identified interactions through a model and data-driven variable selection procedure, revealing the severe diseases depression, stroke, COPD, dementia and osteoporosis as involved in the most complex interaction. In addition, we estimated a simpler additive model consisting solely of main effects, which on average underestimated the effect of severe disease portfolios by a factor of 1.4. We did this to elucidate the importance of considering interaction effects when modelling the mortality risk associated with multiple chronic diseases. To the best of our knowledge, our work is the most extensive study examining the effect of co-occurring diseases on mortality within HD patients.

We found that depression, stroke, COPD, dementia and osteoporosis were involved in interaction relationships of the highest order, indicating that when any of these diseases are added to the HD patient's disease portfolio, the risk contribution of them extensively depend on other diseases already in the portfolio, or the intrinsic variables describing the patient. These diseases were also identified under alternative variable-selection procedures. Our comparisons between the interaction model and the simpler additive model showed differences in the magnitude of the effects for several disease portfolios. Overall, if interactions are not modelled, the average effect of disease portfolios on time until death appears underestimated for disease portfolios with more than three diseases (up to a factor of 1.4, Fig. 4). For females, this average underestimation also applies to disease portfolios of size three. We observed an inverse socioeconomic gradient in the educational dimension for some of the most frequent and complex disease portfolios, where the greater the education attainment level, the greater the associated HR of the disease portfolio (Fig. 5-Fig. 7, Supplemental Fig. 3). We found that cancer was present in all cases in the disease portfolios with highest mortality impact (Table 1). When considering disease portfolios that did not include cancer, we observed that the psychiatric diseases schizophrenia and dementia frequently appeared in conjunction with osteoporosis for males and in conjunction with stroke for females (Table 2). Schizophrenia also often appeared with cancer among the disease portfolios with highest mortality impact. These results highlight effect modification when multiple diseases co-occur in the HD patient, and interventions should, therefore, carefully evaluate the entire disease portfolio of the HD patient.

Effects and Interactions

The high complexity of the estimated interaction model is clearly illustrated in Fig. 3. The figure shows the many dynamics between diseases at play in the HD population, where multiple chronic diseases are rampant. Depression, stroke, COPD, dementia and osteoporosis were the chronic diseases included in the most complex interactions, also allowing interactions between these and the

patient's intrinsic factors. When considering interactions between chronic diseases exclusively (the DIO model), we observed that cancer and diabetes were also involved in the most complex interactions (Supplemental Fig. 1). Interactions with the intrinsic variables sex and age might trivially explain some of these interactions involving cancer and diabetes, which could be why they are not identified among the most complex interactions in the ALL model. Nevertheless, most interactions between individual chronic diseases identified in the ALL model variable selection were similarly discovered in either the S or DIO variable selections (Supplemental Table 5), indicating robustness in the detected interactions.

The consequences of modelling effects of interactions are meticulously visualized on the risk scale in the scenario illustrated in Fig. 9(a), where the addition of dementia does not change the risk profile of the hypothetical patient much, as he already possesses the severe diseases schizophrenia and cancer along with HD. In fact, many of the effect modifications implied by the presence of interactions lead to an attenuation of the combined effect of the diseases compared to their effects in an additive model. Biologically, this is reasonable, as the considered patients are generally frail due to their HD, thereby causing the continued addition of chronic diseases to increase frailty before death eventually occur. Our results showing the effect of COPD decreasing for increasing disease portfolio sizes supports this finding ().

Our analysis showed that both the psychiatric diseases dementia and long-term depression were involved in the most complex interactions (**Fig. 3**). Although not part of five-way interactions, schizophrenia was involved in four-way interactions with several other diseases. These high-order interaction effects in disease portfolios with psychiatric diseases complicate the interpretation of their impact on mortality, as the effects of having these psychiatric diseases depend heavily on the other present chronic diseases in the disease portfolio, as well as intrinsic factors such as age, sex and socioeconomic position. From a biological point of view, this illustrates the interplay between somatic and psychiatric diseases concerning mortality [37,38]. Studies report increased prevalence and risk of psychiatric diagnoses for patients with cardiovascular diseases and their risk factors [39], and efforts should be made to improve these patients' psychological function. Additionally, several studies indicate an increased mortality risk in psychiatric patients when comorbidities are present [7,37,38]. Indeed, we also found that the psychiatric diseases schizophrenia and dementia are manifested in the disease portfolios with the highest hazard ratios (Table 1, Table 2). As a result, the present study gives substantial implications for the priority of identification of psychiatric manifestations of multimorbidity among HD patients, as mortality risk is heavily modified when these diagnoses are present, at least among the chronic diseases and the population considered in the present study.

Cancer was present in all portfolio dyads, triads and tetrads with the highest HR (Table 1). This finding is supported by previous studies reporting that the majority of deaths from cardiovascular disease occur in patients diagnosed with breast, bladder and prostate cancer [40]. However, the cancer diagnosis in our study encapsulates a larger spectrum of cancer conditions. Among the triads and tetrads with highest mortality impact, cancer was often present with schizophrenia. However, when considering portfolios excluding cancer, dyads with dementia had a higher mortality impact. Previous research shows higher cancer mortality rates in individuals with schizophrenia, often attributed to factors such as nonadherence to treatment, diagnostic overshadowing and limited collaboration between medicine and psychiatry [41]. For HD patients, our results highlight these combinations of diseases as having some of the most substantial mortality impacts.

We note that among the variables identified in higher-order interactions, Fig. 7 and Supplemental Fig. 3 show differences in effects when comparing estimates from models with and without interactions. These contrasts emphasize the importance of considering an HD patient's complete

disease portfolio when assessing risk. Our findings show that when interactions are not recognized, the model underestimates the effect of severe disease such as cancer, stroke, and COPD, while overestimating the effect of less severe diseases such as high cholesterol and allergies (Fig. 5). A previous study demonstrated the adverse impact of ignoring statistical interactions in epidemiologic studies, showing a potential bias in main effect parameters estimates [33], which could be a reason for these observed differences. As the underestimation of effects asserted itself even for disease portfolios of small size, it could possibly be attributed to the first few manifestations of multimorbidity (i.e., the first diseases obtained post-HD) being more important for survival than later. While the risk continuously increases with the addition of extra diagnoses, the individual disease effects do not combine additively. As a result, some patients might reach a high risk profile with just a few diagnoses, trivializing the extra effect of obtaining a new diagnosis, as illustrated by the scenario in Fig. 9(a). The situation illustrated in Fig. 9(a) with the mortality risk not changing with the addition of a (on its own) deadly chronic disease can only be modelled when interactions are allowed. We speculate that the simple additive model breaks down due to situations like these by underestimating the effect of individual severe diseases such as cancer, stroke and COPD while overestimating the effect of some of the more common, less severe diseases like high cholesterol, hypertension and allergies. While it was observed that, on average, the additive model underestimates the effect of disease portfolios (Fig. 4), it is essential to mention that the individual disease portfolio effect differences were aggregated across the HD population.

In this study, we observed an apparent negative effect of the high cholesterol diagnosis, indicating increased survival relative to an individual without the disease. This artefact can be attributed to the phenomenon that some HD-diagnosed individuals who are also diagnosed with high cholesterol are likely being treated with lipide modifying agents such as statins, which have many beneficial properties such as cholesterol reduction and anti-inflammatory activities [42,43]. Despite having an additional diagnosis, these HD-diagnosed individuals might represent a less frail part of the HD population, who might have a higher degree of health literacy thus being more aware of their conditions and getting attention by their general practitioners. Another possible explanation is our use of diagnosis time instead of the time of actual disease onset, which is unknown. High cholesterol is a condition where a considerable amount of time may pass before diagnosis [44], and among those HD patients who are undiagnosed, some may have the disease but are not undergoing treatment. It is also essential to consider other consequences of multimorbidity. Increased survival, relative to an individual without a particular disease, may appear beneficial at first glance. However, it is crucial to recognize that an additional chronic disease introduces new challenges, such as new medication management, consultations with general practitioners and specialists, and potential functional impairments. It is essential to remember that increased survival in these cases does not necessarily equate to improved quality of life.

We found a more pronounced effect in disease portfolios including osteoporosis in males, compared to females (Fig. 5, Fig. 6, Fig. 8, Table 1). Notably, despite the generally higher prevalence of osteoporosis in females compared to males, it is well documented that osteoporosis diagnosed males experience higher mortality rates than their female counterparts [45]. Our study reaffirms this observation within a nationwide HD population.

Our findings revealed an inverse socioeconomic gradient for some disease portfolios, where the isolated effect of disease portfolios generally increased as educational attainment levels rose (Fig. 5-Fig. 7, Supplemental Fig. 3). Thus, the higher educated the patient, the higher the mortality hazard rate of the disease portfolio, compared to a person of the same education level with solely HD. It is widely known that individuals with higher levels of education enjoy better overall health and lower mortality hazard rates than people with lower levels of education [46]. Consequently, given that the

reference non-multimorbid HD patient is generally healthier in the subpopulation with the highest educational attainment, it is plausible that those who do become multimorbid in this subpopulation experience a comparatively higher relative mortality hazard rate. Hence, it is important to bear in mind that the hazard ratio reflects the increased relative mortality hazard rate associated with having a specific multimorbid disease portfolio compared to only having HD when interpreting this inverse social gradient. Importantly, the inverse social gradient does not directly translate to increased mortality with higher education on the risk scale, as illustrated in Fig. 7(c). Social disparities are extensively documented across various aspects of multimorbidity, including prevalence [21], healthcare utilization [47], and transitions between disease portfolios [28]. Our results contribute to this by revealing an inverse social gradient, concerning the isolated effect of combinations of chronic diseases on mortality within a nationwide HD population.

As clinical practice such as guidelines, screening, testing and treatment for chronic diseases have evolved over the period 1995-2015, our analysis was adjusted for calendar time at HD diagnosis. We systematically assessed its influence on the most frequently observed disease portfolios. Generally, we observed increased survival for patients diagnosed more recently, compared to earlier (98 out of the 100 most common portfolios). However, an inverse trend, indicating decreased survival over calendar time, was observed for a few disease portfolios, particularly for the portfolio [dementia], and in many cases when dementia is combined with diabetes and/or stroke. It is well known that demographic changes have caused an increase in the prevalence of dementia over the years [48], but since the model is conditional on the disease portfolio, an increased prevalence of dementia over time does not in itself explain the result. We currently lack an explanation for this result, and plan to further investigate it in future research.

Interpretations

The present study illustrates that the complexity of addressing effects of multiple chronic conditions in a large, temporal dataset requires consideration of the individual's complete disease portfolio. The extended Cox model utilized throughout this work was chosen because it allows for modelling time-varying variables in a survival context. Additionally, it has the advantage of making no assumptions regarding the distribution of the survival times, i.e., the underlying hazard function is left unspecified [49]. Few assumptions are, however, made about the hazard function, namely the relationship between covariates and the hazard function. By examining Schoenfeld residuals, we found that in some cases, the proportional hazards assumption was not fully supported [50], meaning that the effects might vary across time. It is therefore essential to interpret the presented effects as weighted averages of the true, possibly time-varying effects across the entire observation period [51]. There are previous studies on the effect of multimorbidity on time to death within HD populations [5,7]. However, the analyses conducted in these studies do not acknowledge that a patient's multimorbidity state is likely to change dynamically through time, i.e., that it is time dependent. The differences in prevalence at time $t=0$ and the end of observation (Fig. 2) in the present study illustrate much progression in disease portfolios. Thus, it is essential to consider this when performing a temporal statistical analysis. When interpreting effects, it is crucial to have the population in mind. As the study population was selected and followed from the time of HD diagnosis, the considered individuals are generally ill, compared to, e.g., an individual without any chronic diseases. Furthermore, with Denmark being a European welfare state, the population differs from many other countries where individuals may have to pay for examinations, thus effects might not be directly comparable due to variations in treatment accessibility.

It is crucial to elaborate further on the contrasts associated with the presented effect estimates. The estimates presented compare a non-multimorbid HD patient to a multimorbid HD patient with a specific disease portfolio. In the OME model, the effect of comparing, e.g., a cancer and COPD-

diagnosed HD patient to a non-multimorbid HD patient would be the same as comparing an HD patient with cancer, COPD, and depression to an HD patient with depression only. In other words, the effect of a disease combination in an additive model can be interpreted as having the specific combination of diagnoses in the disease portfolio versus not having it. In the presence of high-order interactions, however, the interpretation is solely the increased (or decreased) effect comparing an individual with the particular disease portfolio to an individual without it. This is due to the possibility of interactions with other variables, which modify the effects of the disease combination.

The scenarios in Fig. 9 were created to illustrate the workings of the extended Cox model by illustrating how the model estimates the mortality risk over time for the hypothetical HD-diagnosed individuals. However, one should be careful in interpreting these scenarios. They cannot be used prognostically to forecast, as time points of portfolio expansions are never known at the time of HD, as that would be conditioning on future events. These scenarios were solely constructed to represent how the model depicts the mortality risk of a “typical” HD patient over time. The figures help illustrate how the interaction effects on the log-hazard scale relate to the risk of mortality on the probability scale.

For the results presented in this paper, it is essential to emphasize that the effects and interactions uncovered represent associations, not causal relationships. While our results provide valuable insights into the relationships among the chronic diseases, they should be interpreted as observational associations, which can be informative for hypothesis generation and risk assessment for individual portfolios. Furthermore, a considerable group of individuals had missing educational attainment information in this study. In our analyses, we modelled missing values as separate categories. We also estimated the final ALL model under the multiple imputation framework [52], which led to similar results as those presented.

Strengths and Weaknesses

The main strength of the present study is the entire Danish population of HD diagnosed individuals observed over a long period using register data. Danish register data is generally of high quality and fully representative of the entire Danish population [53]. Additionally, the use of algorithmic diagnoses processing both ICD-10 diagnosis history and ATC medicine history ensures that the HD population covers both the primary and secondary parts of the Danish healthcare system. There are, however, several limitations associated with this study. Given the observational nature of the study, our results do not enable us to draw causal conclusions. In addition, despite the algorithmic diagnoses previously being shown reliable [18], a chronic disease’s true onset comes before diagnosis. This is less of a challenge when diagnoses are considered in a cross-sectional study than in a longitudinal setting. Therefore, as timestamps for true disease onsets are not possible, it is crucial to interpret the longitudinal effects associated with a diagnosis in the context of exactly a diagnosis, i.e., the detection of disease, where the individual may have been ill for some time prior to that.

Conclusions

In conclusion, we emphasize the importance of considering a patient’s entire disease portfolio when assessing or modelling risk, avoiding oversimplified silo-based generalizations about the effect of individual diseases. This study highlights the importance of modelling interaction effects when chronic diseases co-occur. Omitting these interactions can result in underestimation of the elevated mortality risk associated with multimorbidity in patients with HD. Through our analysis of a comprehensive nationwide longitudinal dataset of 766,596 patients with HD, we identified sex- and socioeconomic disparities in disease portfolio hazard ratios. Notably, an inverse socioeconomic gradient was systematically observed for the most common and complex disease portfolios meaning

increased mortality hazard rate with multimorbidity relative to no multimorbidity as education attainment level increases. However, absolute mortality risk still decreased with increasing educational attainment due to baseline effects of education. Cancer was present in all disease portfolios with the highest mortality impact. Excluding cancer, disease portfolios including psychiatric chronic diseases were of highest mortality impact. We identified interactions among all considered co-occurring chronic diseases. We found that stroke, osteoporosis, chronic obstructive pulmonary disease, dementia and depression were integral components of the most complex interactions of highest order. When these chronic diseases co-occur in the heart disease patient, their contribution to the patient's risk profile depends on multiple factors, encouraging a holistic view of the patient's entire disease portfolio along with their demographic and socioeconomic risk factors.

Acknowledgements

This study was funded by Greater Copenhagen Health Science Partners (GCHSP) as part of the Clinical Academic Group Prognostication of Acute Recovery Capacity – in an Aging Population (ACUTE-CAG). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. We would like to thank Guillermina Eslava for stimulating discussions and useful suggestions on earlier drafts of the manuscript.

Conflicts of Interests

All authors declare no financial or non-financial competing interests or conflicts of interests.

Data availability

The data used in this study are not publicly available as it consists of sensitive, individual-level information in the form of national register data. According to the Danish data protection legislation, the authors are not allowed to share these sensitive data by request. Instead, the data are available for research purposes upon request to the Danish Health Authority.

Code availability

In order to analyse the data the following software was used: R (version 4.2.2) with packages "survival" {v. 3.5-5}, "lava" {v. 1.7.1}, "glmnet" {v. 4.1-6} and "multcomp" {v. 1_4-20}.

Author contribution statement

A.F., A.S. and N.N.H. developed the design and concept of the study. H.G.JL., N.N.H. and O.A. made substantial contributions to the preparation of data. N.N.H. made all software implementations. N.N.H. conducted the statistical analysis with assistance from A.S. N.N.H. wrote the initial draft of the manuscript in collaboration with A.F. and A.S. All authors made substantial contributions to the interpretation of results. All authors read and approved the final version of the manuscript.

Bibliography

1. Nowbar AN, Gitto M, Howard JP, Francis DP, Al-Lamee R. Mortality from ischemic heart disease: Analysis of data from the world health organization and coronary artery disease risk factors from NCD risk factor collaboration. *Circ Cardiovasc Qual Outcomes* Lippincott Williams and Wilkins; 2019 Jun 1;12(6). PMID:31163980
2. Forman DE, Maurer MS, Boyd C, Brindis R, Salive ME, Horne FMF, Bell SP, Fulmer T, Reuben DB, Zieman S, Rich MW. Multimorbidity in Older Adults With Cardiovascular Disease. *J Am Coll Cardiol Elsevier*; 2018 May 15;71(19):2149–2161. PMID:29747836

3. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: Implications for understanding health and health services. *Ann Fam Med Annals of Family Medicine, Inc*; 2009;7(4):357–363. PMID:19597174
4. Buddeke J, Bots ML, Van Dis I, Visseren FLJ, Hollander M, Schellevis FG, Vaartjes I. Comorbidity in patients with cardiovascular disease in primary care: A cohort study with routine healthcare data. *British Journal of General Practice Royal College of General Practitioners*; 2019;69(683):E398–E406. PMID:31064742
5. Gimeno-Miguel A, Gracia Gutiérrez A, Poblador-Plou B, Coscollar-Santaliestra C, Pérez-Calvo JI, Divo MJ, Calderón-Larrañaga A, Prados-Torres A, Ruiz-Laiglesia FJ. Multimorbidity patterns in patients with heart failure: An observational Spanish study based on electronic health records. *BMJ Open BMJ Publishing Group*; 2019 Dec 23;9(12). PMID:31874886
6. Chen YH, Karimi M, Rutten-Van Mölken MPMH. The disease burden of multimorbidity and its interaction with educational level. *PLoS One Public Library of Science*; 2020 Dec 1;15(12 December). PMID:33270760
7. Sariaslan A, Sharpe M, Larsson H, Wolf A, Lichtenstein P, Fazel S. Psychiatric comorbidity and risk of premature mortality and suicide among those with chronic respiratory diseases, cardiovascular diseases, and diabetes in Sweden: A nationwide matched cohort study of over 1 million patients and their unaffected siblings. *PLoS Med Public Library of Science*; 2022 Jan 1;19(1). PMID:35085232
8. Van Deursen VM, Urso R, Laroche C, Damman K, Dahlström U, Tavazzi L, Maggioni AP, Voors AA. Co-morbidities in patients with heart failure: An analysis of the European heart Failure Pilot Survey. *Eur J Heart Fail Oxford University Press*; 2014;16(1):103–111. PMID:24453099
9. Eriksson B, Wändell P, Dahlström U, Näsman P, Lund LH, Edner M. Comorbidities, risk factors and outcomes in patients with heart failure and an ejection fraction of more than or equal to 40% in primary care- and hospital care-based outpatient clinics. *Scand J Prim Health Care Taylor and Francis Ltd*; 2018 Apr 3;36(2):207–215. PMID:29633886
10. Haue AD, Armenteros JJA, Holm PC, Eriksson R, Moseley PL, Køber L V., Bundgaard H, Brunak S. Temporal patterns of multi-morbidity in 570157 ischemic heart disease patients: a nationwide cohort study. *Cardiovasc Diabetol BioMed Central Ltd*; 2022 Dec 1;21(1). PMID:35641964
11. Willadsen TG, Siersma V, Nicolaisdóttir DR, Køster-Rasmussen R, Reventlow S, Rozing M. The effect of disease onset chronology on mortality among patients with multimorbidity: A Danish nationwide register study. *Journal of Multimorbidity and Comorbidity SAGE Publications*; 2022 Jan;12:263355652211220. doi: 10.1177/26335565221122025
12. France EF, Wyke S, Gunn JM, Mair FS, McLean G, Mercer SW. Multimorbidity in primary care: A systematic review of prospective cohort studies. *British Journal of General Practice* 2012 Apr;62(597). PMID:22520918
13. The Academy of Medical Sciences. Multimorbidity: a priority for global health research. 2018.
14. Muth C, Kirchner H, Van Den Akker M, Scherer M, Glasziou PP. Current guidelines poorly address multimorbidity: Pilot of the interaction matrix method. *J Clin Epidemiol Elsevier Inc.*; 2014 Nov 1;67(11):1242–1250. PMID:25216898
15. Muth C, Blom JW, Smith SM, Johnell K, Gonzalez-Gonzalez AI, Nguyen TS, Brueckle MS, Cesari M, Tinetti ME, Valderas JM. Evidence supporting the best clinical management of patients with multimorbidity and polypharmacy: a systematic guideline review and expert consensus. *J Intern Med. Blackwell Publishing Ltd*; 2019. p. 272–288. PMID:30357955
16. Pedersen CB. The Danish civil registration system. *Scand J Public Health* 2011 Jul;39(7):22–25. PMID:21775345
17. Commission E. Data protection rules as a trust-enabler in the EU and beyond–taking stock. *European Commission Communication* 2019;
18. Robinson K, Magtengaard, Lau CJuel, Jeppesen Maj, Vind A, Bonnerup, Glümer Charlotte. *Kroniske sygdomme : hvordan opgøres kroniske sygdomme. Forskningscenter for forebyggelse og sundhed*; 2012. ISBN:9788799450206

19. Christensen AI, Davidsen M, Ekholm O, Pedersen PV, Juel K. Danskernes sundhed: Den nationale sundhedsprofil 2013. Sundhedsstyrelsen; 2014;
20. Jensen HAR, Davidsen M, Ekholm O, Christensen AI. Danskernes sundhed-Den nationale sundhedsprofil 2017. Sundhedsstyrelsen; 2018;
21. Schiøtz ML, Stockmarr A, Høst D, Glümer C, Frølich A. Social disparities in the prevalence of multimorbidity - A register-based population study. BMC Public Health BioMed Central Ltd.; 2017 May 10;17(1). PMID:28486983
22. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol Taylor & Francis; 2015;449–490.
23. Mors O, Perto GP, Mortensen PB. The Danish psychiatric central research register. Scand J Public Health Sage Publications Sage UK: London, England; 2011;39(7_suppl):54–57.
24. Kildemoes HW, Sørensen HT, Hallas J. The Danish national prescription registry. Scand J Public Health Suppl JSTOR; 2011;38–41.
25. Sahl Andersen J, De Fine Olivarius N, Krasnik A. The Danish national health service register. Scand J Public Health Sage Publications Sage UK: London, England; 2011;39(7_suppl):34–37.
26. Jensen VM, Rasmussen AW. Danish education registers. Scand J Public Health 2011 Jul;39(7):91–94. PMID:21775362
27. Helweg-Larsen K. The Danish register of causes of death. Scand J Public Health 2011 Jul;39(7):26–29. PMID:21775346
28. Holm NN, Frølich A, Andersen O, Juul-Larsen HG, Stockmarr A. Longitudinal models for the progression of disease portfolios in a nationwide chronic heart disease population. PLoS One Public Library of Science; 2023 Apr 1;18(4 April). PMID:37079591
29. Berrington de González A, Cox DR. Interpretation of interaction: A review. Ann Appl Stat Institute of Mathematical Statistics; 2007 Dec 4;1(2). doi: 10.1214/07-aos124
30. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological) Wiley Online Library; 1972;34(2):187–202.
31. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika Oxford University Press; 1994;81(3):515–526.
32. Brambor T, Clark WR, Golder M. Understanding interaction models: Improving empirical analyses. Political analysis Cambridge University Press; 2006;14(1):63–82.
33. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. The effect of ignoring statistical interactions in regression analyses conducted in epidemiologic studies: an example with survival analysis using Cox proportional hazards regression model. Epidemiology (Sunnyvale) NIH Public Access; 2015;6(1).
34. Meinshausen N, Bühlmann P. Stability selection. J R Stat Soc Series B Stat Methodol Oxford University Press; 2010;72(4):417–473.
35. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological) Wiley; 1996 Jan;58(1):267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
36. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med Wiley Online Library; 1997;16(4):385–395.
37. Koyanagi A, Köhler-Forsberg O, Benros ME, Munk Laursen T, Haro JM, Nordentoft M, Hjorthøj C. Mortality in unipolar depression preceding and following chronic somatic diseases. Acta Psychiatr Scand Blackwell Publishing Ltd; 2018 Dec 1;138(6):500–508. PMID:29761489
38. Kugathasan P, Stubbs B, Aagaard J, Jensen SE, Munk Laursen T, Nielsen RE. Increased mortality from somatic multimorbidity in patients with schizophrenia: a Danish nationwide cohort study. Acta Psychiatr Scand Blackwell Publishing Ltd; 2019 Oct 1;140(4):340–348. PMID:31355419
39. Zhang Y, Chen Y, Ma L. Depression and cardiovascular disease in elderly: Current understanding. Journal of Clinical Neuroscience. Churchill Livingstone; 2018. p. 1–5. PMID:29066229
40. Sturgeon KM, Deng L, Bluethmann SM, Zhou S, Trifiletti DM, Jiang C, Kelly SP, Zaorsky NG. A

- population-based study of cardiovascular disease mortality risk in US cancer patients. *Eur Heart J* 2019 Dec 21;40(48):3889–3897. doi: 10.1093/eurheartj/ehz766
41. González-Rodríguez A, Labad J, Seeman M V. Schizophrenia and cancer. *Curr Opin Support Palliat Care* LWW; 2020;14(3):232–238.
 42. Niazi M, Galehdar N, Jamshidi M, Mohammadi R, Moayyedkazemi A. A Review of the Role of Statins in Heart Failure Treatment. *Curr Clin Pharmacol* Bentham Science Publishers Ltd.; 2019 Aug 4;15(1):30–37. PMID:31376825
 43. Lee MMY, Sattar N, McMurray JJV, Packard CJ. Statins in the Prevention and Treatment of Heart Failure: a Review of the Evidence. *Curr Atheroscler Rep*. Current Medicine Group LLC 1; 2019. PMID:31350612
 44. Vallejo-Vaz AJ, Stevens CAT, Lyons ARM, Dharmayat KI, Freiburger T, Hovingh GK, Mata P, Raal FJ, Santos RD, Soran H, others. Global perspective of familial hypercholesterolaemia: a cross-sectional study from the EAS Familial Hypercholesterolaemia Studies Collaboration (FHSC). *The Lancet Elsevier*; 2021;398(10312):1713–1725.
 45. Rinonapoli G, Ruggiero C, Meccariello L, Bisaccia M, Ceccarini P, Caraffa A. Osteoporosis in men: A review of an underestimated bone condition. *Int J Mol Sci*. MDPI AG; 2021. p. 1–22. PMID:33672656
 46. Davies NM, Dickson M, Smith GD, Van Den Berg GJ, Windmeijer F. The causal effects of education on health outcomes in the UK Biobank. *Nat Hum Behav* Nature Publishing Group; 2018 Feb 1;2(2):117–125. PMID:30406209
 47. Frølich A, Ghith N, Schiøtz M, Jacobsen R, Stockmarr A. Multimorbidity, healthcare utilization and socioeconomic status: A register-based study in Denmark. *PLoS One Public Library of Science*; 2019 Aug 1;14(8). PMID:31369580
 48. Nichols E, Szeke CEI, Vollset SE, Abbasi N, Abd-Allah F, Abdela J, Aichour MTE, Akinyemi RO, Alahdab F, Asgedom SW, others. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol Elsevier*; 2019;18(1):88–106.
 49. Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. Springer Science & Business Media; 2012.
 50. Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. 1994. Available from: <https://www.jstor.org/stable/2337123>
 51. Stensrud MJ, Hernán MA. Why test for proportional hazards? *JAMA American Medical Association*; 2020;323(14):1401–1402.
 52. Van Buuren S. Flexible imputation of missing data. CRC press; 2018.
 53. Erlangsen A, Fedyszyn I. Danish nationwide registers for public health and health-related research. *Scand J Public Health*. 2015. p. 333–339. PMID:25759376

Supplemental material

Supplemental Table 1. Algorithmic diagnoses. Algorithms used to define the 15 diagnoses.

Chronic condition	ICD-10 from NPR and PCRR	Definition
Heart disease	I20, I21, I23-I25, I50, I11, I13	(DIAG) ^a and/or (MEDICINE) ^b with ATC: C01A, C01B, C01D, C01E
Stroke	G45, G46, I60-I69	(DIAG) ^a

Hypertension	I10-I13, I15	(DIAG) ^a and/or (MEDICINE) ^b with ATC: C07B, C03A, C03B, C03E, C03X and/or (MEDICINE) ^b with ATC: C03C, C03D, C07A, C09 if they do not have contacts with ICD-10 codes I20.0, I21, I25.1, I50 and/or (MEDICINE) ^b with ATC: C08 if they do not have contacts with ICD-10 codes I20-25
High cholesterol	E78.0, E78.2, E78.4, E78.5	(DIAG) ^a and/or (MEDICINE) ^b with ATC: C10
Allergies	J30.1-J30.9	(DIAG) ^a and/or (MEDICINE) ^b with ATC: V01AA11, V01AA02, V01AA03, V01AA05, R01AC, R01AD, R06A, S01G, R01BA52. Two prescriptions within a period of two years.
Joint disease	M05, M06.0, M06.8, M07.0, M07.1, M10.0, M10.9	(DIAG) ^a
Osteoporosis	M80-M82 and/or for persons aged 45 years or older S22.0, S22.1, S32.0, S32.7, S32.8, S42.2, S42.4, S42.7-S42.9, S52.5-S52.9, S62.0, S62.1, S72	(DIAG) ^a and/or (MEDICINE) ^b with ATC: M05B, G03XC01, H05AA02, H05AA03
Osteoarthritis	M15-M19	(DIAG) ^a
Back pain	M40-M54	(DIAG) ^a
Cancer	C00-C43, C45-C97	(DIAG) ^a
Chronic obstructive pulmonary disease (COPD)	J40-J44, J47, J96	(DIAG) ^a All patients aged 35 years or older at contact. And/or (MEDICINE) ^b with ATC: V03AN01 if they do not have contacts with ICD-10 codes J45 or J46, also aged 35 years or older at the date of distribution. And/or (SERVICE) ^c with lab services 807113 (lung spirometer test), 807121 (lung function test), also aged 35 years or older at the time of service.
Schizophrenia	F20-F22, F25, F28, F29, F31	(DIAG) ^a and/or (MEDICINE) ^b N05AX13, N05AX12, N05AH03, N05AX08
Dementia	F00, G30, F01, F02.0, F03.9, G31.8B, G31.8E, G31.9, G31.0B	(DIAG) ^a All patients aged 60 years or older at contact. And/or (MEDICINE) ^b with ATC: N06D, also aged 35 years or older at the

	date of distribution.
Long term use of antidepressants (Depression)	(MEDICINE) ^b with ATC: N06A. Note: At least three different prescriptions with at least 2 years between the first and last one. Patients with previous schizophrenia or dementia are excluded.
Diabetes	(DIAG) ^a and/or (MEDICINE) ^b with ATC: A10. A10BA02 is excluded for females aged 20-40 years at the date of distribution

^a DIAG: All patients at any age unless otherwise specified, who had a hospital inpatient or outpatient encounter with one of the ICD-10 codes specified in the column ICD-10 from NPR and PCRR. Primary (A) and secondary (B) diagnoses are considered. Diagnosis timestamp is taken to be at the midway-point of the first encounter fulfilling the requirements.

^b MEDICINE: All patients at any age unless otherwise specified, who had a minimum of two medicine prescriptions from DNPR on two separate days with the stated ATC codes within a period of one year. Diagnosis timestamp is taken to be the date of medicine distribution when the criteria was fulfilled.

^c SERVICE: All patients who had a minimum of two healthcare services from NHSR with either of the stated lab service codes on two separate days within a period of one year. As healthcare services are reported on a weekly basis, diagnosis timestamp was taken to be at the last day of the week. If a patient fulfills multiple of the criteria DIAG, MEDICINE and SERVICE, the chronic disease diagnosis timestamp is taken to be the first occurring.

Supplemental Table 2. Prevalence of diagnoses according to sex.

Chronic condition	Male	Female	Total
Prevalence at $t = 0$ (time of HD diagnosis)			
Heart disease	100.00 (406792)	100.00 (359804)	100.00 (766596)
Hypertension	64.97 (264306)	74.69 (268721)	69.53 (533027)
High cholesterol	32.50 (132213)	23.52 (84640)	28.29 (216853)
Allergies	14.51 (59013)	21.16 (76149)	17.63 (135162)
COPD	11.03 (44873)	12.08 (43462)	11.52 (88335)
Diabetes	11.66 (47426)	10.12 (36401)	10.93 (83827)
Depression	6.38 (25955)	12.30 (44255)	9.16 (70210)
Osteoporosis	4.13 (16800)	14.81 (53307)	9.14 (70107)
Stroke	8.81 (35834)	9.06 (32596)	8.93 (68430)
Cancer	6.87 (27934)	7.26 (26122)	7.05 (54056)
Osteoarthritis	5.70 (23193)	8.20 (29496)	6.87 (52689)
Back pain	5.26 (21409)	6.39 (23005)	5.79 (44414)
Dementia	1.39 (5665)	2.58 (9277)	1.95 (14942)
Schizophrenia	1.34 (5466)	1.92 (6926)	1.62 (12392)
Joint disease	1.34 (5455)	1.55 (5580)	1.44 (11035)
Lifetime prevalence ^a			

Heart disease	100.00 (406792)	100.00 (359804)	100.00 (766596)
Hypertension	77.71 (316100)	85.11 (306224)	81.18 (622324)
High cholesterol	51.12 (207983)	37.94 (136502)	44.94 (344485)
Allergies	25.24 (102687)	32.99 (118701)	28.88 (221388)
COPD	25.07 (101975)	24.32 (87506)	24.72 (189481)
Osteoporosis	12.11 (49269)	33.35 (120002)	22.08 (169271)
Stroke	21.15 (86021)	22.12 (79606)	21.61 (165627)
Diabetes	22.74 (92519)	18.75 (67450)	20.87 (159969)
Depression	15.87 (64542)	25.33 (91156)	20.31 (155698)
Cancer	20.38 (82923)	17.92 (64485)	19.23 (147408)
Osteoarthritis	12.63 (51361)	16.47 (59268)	14.43 (110629)
Back pain	11.52 (46846)	13.43 (48332)	12.42 (95178)
Dementia	6.34 (25806)	10.25 (36890)	8.18 (62696)
Schizophrenia	3.56 (14462)	5.22 (18778)	4.34 (33240)
Joint disease	4.26 (17330)	3.73 (13436)	4.01 (30766)

Prevalence are reported as % (n).

^aLifetime prevalence is to be understood in the sense that the HD diagnosed individual has been observed to obtain the diagnosis in the time period 1995-2015.

Supplemental Table 3 Male hazard ratios the 10 most common disease portfolio dyads, triads, tetrads and pentads. The results are presented for the ALL model at the four educational attainment levels: None, short, medium and long and correspond to the situation presented in Fig. 5. The reference group comprises males with only HD and the corresponding educational attainment level. Results are also presented for the additive OME model. In each disease portfolio group, the disease portfolio HR estimates are presented in order of prevalence, with the upper rows being more prevalent than the lower rows. HT = hypertension, HC = high cholesterol, AL = allergies, COPD = chronic obstructive pulmonary disease, CAN = cancer, ST = stroke, DIA = diabetes, OP = osteoporosis, DEP = depression, BP = back pain, OA = osteoarthritis.

Dyads									
Portfolio ^a	No Education		Short Education		Medium Education		Long Education		Additive model
	HR	CI (99.9%)	HR	CI (99.9%)	HR	CI (99.9%)	HR	CI (99.9%)	HR CI (99.9%)
[HT]	0.94	(0.90-0.98)	1.00	(0.95-1.05)	1.00	(0.90-1.10)	0.97	(0.88-1.09)	1.14 (1.12-1.15)
[HC]	0.39	(0.36-0.41)	0.38	(0.36-0.41)	0.39	(0.36-0.42)	0.39	(0.36-0.43)	0.64 (0.64-0.65)
[AL]	0.79	(0.75-0.83)	0.79	(0.75-0.83)	0.79	(0.75-0.83)	0.79	(0.75-0.83)	0.92 (0.91-0.94)
[COPD]	2.79	(2.63-2.95)	2.97	(2.80-3.16)	3.08	(2.77-3.42)	3.17	(2.79-3.60)	2.23 (2.20-2.25)
[CAN]	7.03	(6.58-7.50)	7.34	(6.83-7.88)	7.84	(6.78-9.07)	6.78	(5.74-8.02)	3.19 (3.15-3.23)
[ST]	2.81	(2.66-2.97)	2.97	(2.80-3.15)	3.11	(2.80-3.46)	3.04	(2.70-3.41)	1.85 (1.83-1.88)
[DIA]	2.01	(1.93-2.10)	2.04	(1.95-2.13)	2.16	(2.01-2.33)	2.30	(2.11-2.51)	1.48 (1.47-1.50)
[OP]	2.25	(2.12-2.39)	2.26	(2.12-2.41)	2.35	(2.13-2.60)	2.40	(2.13-2.71)	1.61 (1.59-1.63)
[DEP]	1.35	(1.26-1.44)	1.44	(1.35-1.54)	1.55	(1.38-1.75)	1.41	(1.24-1.60)	1.20 (1.18-1.21)
[BP]	0.94	(0.90-0.99)	0.96	(0.91-1.01)	1.01	(0.92-1.11)	1.02	(0.91-1.14)	0.96 (0.94-0.97)
Triads									
[HT, HC]	0.52	(0.49-0.55)	0.55	(0.52-0.58)	0.55	(0.50-0.62)	0.55	(0.49-0.62)	0.73 (0.72-0.75)
[HT, AL]	0.89	(0.85-0.94)	0.95	(0.90-1.00)	0.95	(0.85-1.05)	0.93	(0.83-1.03)	1.05 (1.03-1.07)
[HT, DIA]	1.90	(1.79-2.01)	2.04	(1.91-2.17)	2.15	(1.91-2.42)	2.25	(1.97-2.56)	1.69 (1.66-1.72)
[HT, COPD]	2.99	(2.83-3.17)	3.39	(3.19-3.62)	3.50	(3.06-4.00)	3.53	(3.02-4.12)	2.53 (2.49-2.58)
[ST, HT]	2.53	(2.37-2.69)	2.84	(2.65-3.04)	2.96	(2.58-3.40)	2.83	(2.43-3.30)	2.11 (2.07-2.15)
[HT, CAN]	6.18	(5.84-6.53)	6.59	(6.21-7.00)	7.29	(6.51-8.16)	6.72	(5.93-7.61)	3.63 (3.15-3.23)

[HT, OP]	2.24	(2.10-2.39)	2.39	(2.23-2.57)	2.48	(2.16-2.84)	2.48	(2.12-2.89)	1.83	(1.80-1.87)
[HT, DEP]	1.38	(1.30-1.48)	1.57	(1.46-1.69)	1.69	(1.46-1.95)	1.50	(1.28-1.76)	1.36	(1.33-1.39)
[HT, OA]	0.73	(0.69-0.77)	0.78	(0.73-0.82)	0.77	(0.70-0.86)	0.76	(0.67-0.85)	0.95	(0.93-0.97)
[HC, AL]	0.30	(0.28-0.33)	0.30	(0.28-0.33)	0.31	(0.28-0.34)	0.31	(0.28-0.34)	0.60	(0.59-0.61)
Tetrads										
[HT, HC, DIA]	1.03	(0.98-1.09)	1.11	(1.04-1.18)	1.18	(1.05-1.33)	1.25	(1.09-1.43)	1.09	(1.07-1.11)
[HT, HC, AL]	0.49	(0.47-0.52)	0.52	(0.49-0.56)	0.53	(0.47-0.59)	0.52	(0.46-0.59)	0.68	(0.66-0.69)
[ST, HT, HC]	1.38	(1.29-1.47)	1.54	(1.44-1.66)	1.63	(1.41-1.88)	1.58	(1.34-1.85)	1.36	(1.33-1.39)
[HT, HC, COPD]	1.72	(1.62-1.83)	1.94	(1.81-2.08)	2.03	(1.75-2.34)	2.07	(1.75-2.45)	1.63	(1.60-1.67)
[HT, HC, CAN]	3.80	(3.57-4.03)	4.04	(3.79-4.31)	4.51	(3.99-5.11)	4.21	(3.67-4.84)	2.34	(2.29-2.40)
[HT, HC, OA]	0.45	(0.43-0.48)	0.48	(0.45-0.51)	0.48	(0.43-0.54)	0.48	(0.42-0.54)	0.61	(0.60-0.63)
[HT, AL, COPD]	2.64	(2.48-2.80)	2.99	(2.80-3.20)	3.08	(2.69-3.53)	3.11	(2.65-3.64)	2.34	(2.29-2.39)
[HT, HC, DEP]	0.79	(0.74-0.85)	0.89	(0.83-0.97)	0.97	(0.83-1.13)	0.87	(0.74-1.04)	0.88	(0.86-0.90)
[HT, HC, OP]	1.29	(1.20-1.39)	1.37	(1.27-1.48)	1.44	(1.24-1.66)	1.45	(1.23-1.72)	1.18	(1.15-1.21)
[HT, AL, DEP]	1.38	(1.29-1.48)	1.57	(1.45-1.69)	1.68	(1.45-1.95)	1.49	(1.27-1.76)	1.26	(1.23-1.29)
Pentads										
[HT, HC, AL, DIA]	1.03	(0.97-1.10)	1.11	(1.04-1.19)	1.18	(1.05-1.34)	1.25	(1.09-1.44)	1.01	(0.98-1.03)
[HT, HC, AL, COPD]	1.52	(1.42-1.62)	1.71	(1.59-1.84)	1.79	(1.54-2.07)	1.82	(1.54-2.16)	1.51	(1.47-1.55)
[ST, HT, HC, DIA]	2.34	(2.18-2.51)	2.65	(2.45-2.86)	2.96	(2.55-3.45)	3.06	(2.58-3.62)	2.02	(1.97-2.07)
[HT, HC, COPD, DIA]	2.80	(2.61-3.00)	3.20	(2.96-3.45)	3.54	(3.03-4.13)	3.84	(3.21-4.60)	2.43	(2.37-2.49)
[HT, HC, AL, DEP]	0.79	(0.73-0.85)	0.89	(0.82-0.97)	0.97	(0.83-1.13)	0.87	(0.73-1.04)	0.81	(0.79-0.83)
[ST, HT, HC, AL]	1.31	(1.22-1.40)	1.47	(1.36-1.58)	1.55	(1.34-1.79)	1.50	(1.27-1.76)	1.26	(1.23-1.29)
[HT, HC, AL, OA]	0.43	(0.40-0.46)	0.45	(0.42-0.49)	0.46	(0.41-0.51)	0.45	(0.40-0.52)	0.57	(0.55-0.58)
[HT, HC, DEP, DIA]	1.48	(1.37-1.59)	1.69	(1.55-1.84)	1.95	(1.66-2.29)	1.86	(1.55-2.24)	1.30	(1.27-1.34)
[HT, HC, OA, DIA]	0.90	(0.84-0.96)	0.96	(0.89-1.03)	1.02	(0.90-1.16)	1.08	(0.94-1.25)	0.91	(0.88-0.93)
[HT, HC, CAN, DIA]	5.24	(4.90-5.62)	5.64	(5.24-6.07)	6.69	(5.84-7.67)	6.66	(5.71-7.76)	3.48	(3.39-3.56)

^a All portfolios contain the HD diagnosis.

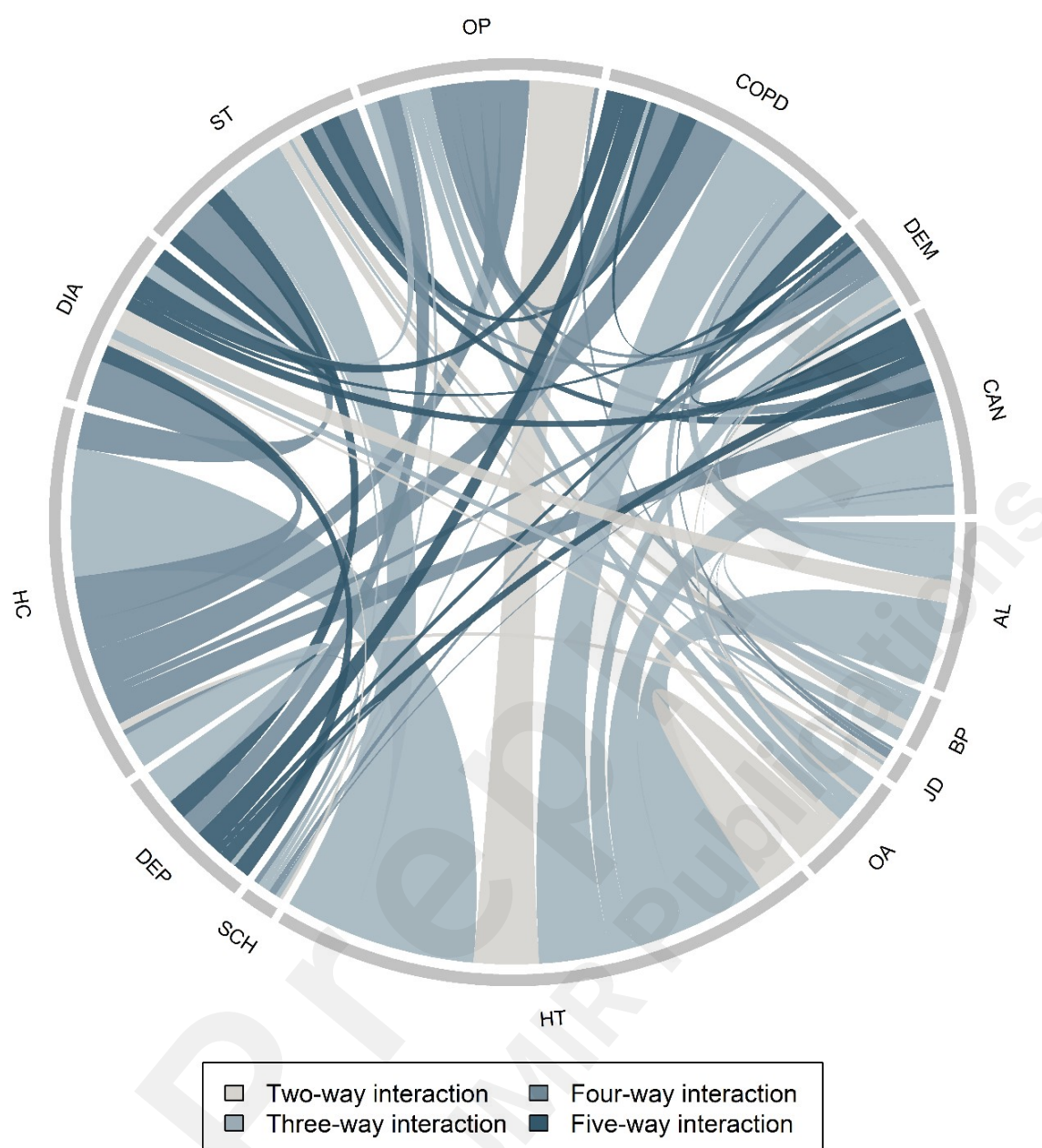
Supplemental Table 4 Female hazard ratios the 10 most common disease portfolio dyads, triads, tetrads and pentads. The results are presented for the ALL model at the four educational attainment levels: None, short, medium and long and correspond to the situation presented in Fig. 6. The reference group comprises females with only HD and the corresponding educational attainment level. Results are also presented for the additive OME model. In each disease portfolio group, the disease portfolio HR estimates are presented in order of prevalence, with the upper rows being more prevalent than the lower rows.

HT = hypertension, HC = high cholesterol, AL = allergies, COPD = chronic obstructive pulmonary disease, CAN = cancer, ST = stroke, DIA = diabetes, OP = osteoporosis, DEP = depression, BP = back pain, OA = osteoarthritis.

Dyads										
Portfolio ^a	No Education		Short Education		Medium Education		Long Education		Additive model	
	HR	CI (99.9%)	HR	CI (99.9%)	HR	CI (99.9%)	HR	CI (99.9%)	HR	CI (99.9%)
[HT]	0.92	(0.88-0.97)	0.98	(0.93-1.04)	0.97	(0.88-1.08)	0.95	(0.85-1.07)	1.14	(1.12-1.15)
[HC]	0.39	(0.36-0.41)	0.38	(0.36-0.41)	0.39	(0.36-0.42)	0.39	(0.36-0.43)	0.64	(0.64-0.65)
[AL]	0.74	(0.71-0.78)	0.74	(0.71-0.78)	0.74	(0.71-0.78)	0.74	(0.71-0.78)	0.92	(0.91-0.94)
[COPD]	3.56	(3.35-3.79)	3.80	(3.55-4.06)	3.94	(3.53-4.39)	4.05	(3.56-4.62)	2.23	(2.20-2.25)
[CAN]	8.07	(7.53-8.66)	8.43	(7.80-9.11)	9.01	(7.76-10.46)	7.79	(6.57-9.24)	3.19	(3.15-3.23)
[ST]	3.54	(3.33-3.75)	3.76	(3.49-4.04)	4.69	(4.05-5.42)	4.14	(3.45-4.96)	1.85	(1.83-1.88)
[DIA]	2.32	(2.21-2.42)	2.34	(2.23-2.46)	2.49	(2.30-2.69)	2.65	(2.42-2.90)	1.48	(1.47-1.50)
[OP]	1.81	(1.71-1.92)	1.82	(1.70-1.94)	1.89	(1.71-2.09)	1.93	(1.71-2.18)	1.61	(1.59-1.63)

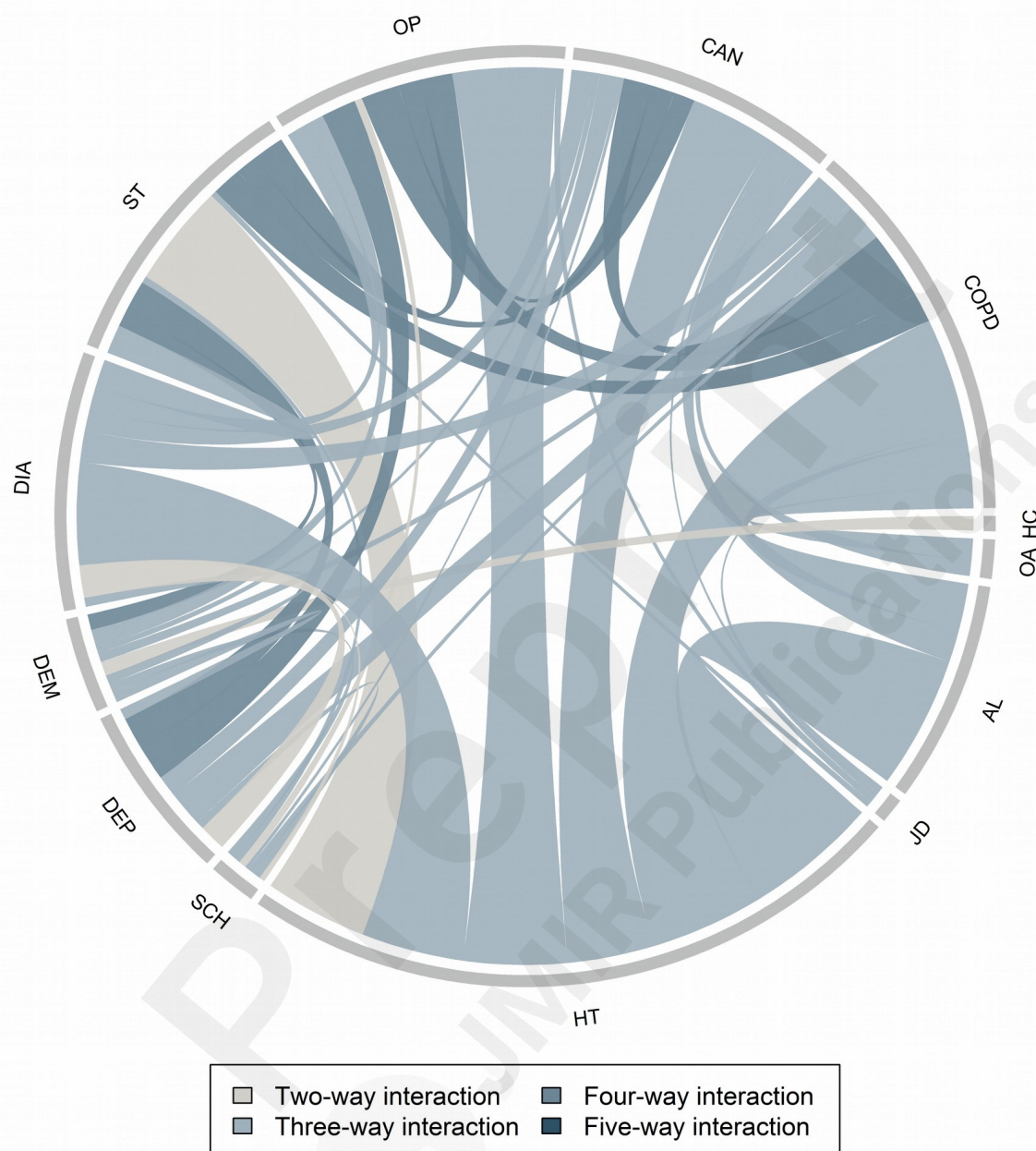
[DEP]	1.23	(1.15-1.31)	1.31	(1.23-1.41)	1.42	(1.26-1.60)	1.29	(1.13-1.47)	1.20	(1.18-1.21)
[BP]	0.85	(0.80-0.89)	0.86	(0.81-0.91)	0.91	(0.83-1.01)	0.91	(0.81-1.02)	0.96	(0.94-0.97)
Triads										
[HT, HC]	0.51	(0.48-0.54)	0.54	(0.51-0.57)	0.54	(0.48-0.61)	0.54	(0.47-0.61)	0.73	(0.72-0.75)
[HT, AL]	0.83	(0.78-0.87)	0.88	(0.83-0.93)	0.87	(0.79-0.97)	0.86	(0.76-0.96)	1.05	(1.03-1.07)
[HT, DIA]	2.13	(2.00-2.27)	2.29	(2.13-2.47)	2.42	(2.14-2.74)	2.53	(2.20-2.90)	1.69	(1.66-1.72)
[HT, COPD]	3.74	(3.52-3.98)	4.25	(3.95-4.57)	4.38	(3.81-5.03)	4.41	(3.75-5.18)	2.53	(2.49-2.58)
[ST, HT]	3.11	(2.90-3.34)	3.52	(3.23-3.83)	4.37	(3.68-5.18)	3.77	(3.06-4.64)	2.11	(2.07-2.15)
[HT, CAN]	6.94	(6.52-7.39)	7.41	(6.91-7.95)	8.19	(7.27-9.24)	7.55	(6.62-8.62)	3.63	(3.15-3.23)
[HT, OP]	1.76	(1.65-1.88)	1.88	(1.74-2.03)	1.95	(1.70-2.24)	1.95	(1.66-2.28)	1.83	(1.80-1.87)
[HT, DEP]	1.24	(1.16-1.32)	1.40	(1.30-1.52)	1.51	(1.30-1.75)	1.34	(1.14-1.58)	1.36	(1.33-1.39)
[HT, OA]	0.71	(0.67-0.76)	0.76	(0.71-0.81)	0.76	(0.68-0.84)	0.74	(0.66-0.83)	0.95	(0.93-0.97)
[HC, AL]	0.29	(0.26-0.31)	0.29	(0.26-0.31)	0.29	(0.26-0.32)	0.29	(0.26-0.32)	0.60	(0.59-0.61)
Tetrads										
[HT, HC, DIA]	1.16	(1.09-1.24)	1.25	(1.16-1.34)	1.33	(1.17-1.51)	1.40	(1.22-1.62)	1.09	(1.07-1.11)
[HT, HC, AL]	0.46	(0.43-0.48)	0.48	(0.45-0.52)	0.49	(0.43-0.55)	0.48	(0.42-0.55)	0.68	(0.66-0.69)
[ST, HT, HC]	1.70	(1.58-1.82)	1.91	(1.75-2.09)	2.40	(2.02-2.86)	2.10	(1.70-2.59)	1.36	(1.33-1.39)
[HT, HC, COPD]	2.15	(2.01-2.30)	2.43	(2.25-2.63)	2.54	(2.18-2.94)	2.59	(2.18-3.08)	1.63	(1.60-1.67)
[HT, HC, CAN]	4.27	(3.99-4.57)	4.54	(4.21-4.90)	5.08	(4.46-5.78)	4.74	(4.10-5.48)	2.34	(2.29-2.40)
[HT, HC, OA]	0.44	(0.41-0.47)	0.47	(0.44-0.50)	0.47	(0.42-0.53)	0.47	(0.41-0.53)	0.61	(0.60-0.63)
[HT, AL, COPD]	2.15	(2.01-2.30)	2.43	(2.25-2.63)	2.54	(2.18-2.94)	2.59	(2.18-3.08)	2.34	(2.29-2.39)
[HT, HC, DEP]	0.71	(0.66-0.76)	0.80	(0.73-0.87)	0.87	(0.74-1.01)	0.78	(0.65-0.93)	0.88	(0.86-0.90)
[HT, HC, OP]	1.02	(0.95-1.09)	1.08	(0.99-1.17)	1.13	(0.97-1.31)	1.14	(0.96-1.36)	1.18	(1.15-1.21)
[HT, AL, DEP]	1.16	(1.09-1.25)	1.32	(1.22-1.43)	1.42	(1.22-1.65)	1.26	(1.07-1.49)	1.26	(1.23-1.29)
Pentads										
[HT, HC, AL, DIA]	1.01	(1.03-1.18)	1.18	(1.09-1.27)	1.26	(1.11-1.43)	1.33	(1.15-1.54)	1.01	(0.98-1.03)
[HT, HC, AL, COPD]	1.79	(1.67-1.92)	2.03	(1.87-2.20)	2.11	(1.82-2.45)	2.15	(1.81-2.57)	1.51	(1.47-1.55)
[ST, HT, HC, DIA]	3.31	(3.07-3.58)	3.77	(3.43-4.14)	5.02	(4.18-6.03)	4.68	(3.75-5.84)	2.02	(1.97-2.07)
[HT, HC, COPD, DIA]	3.68	(3.41-3.98)	4.21	(3.85-4.61)	4.66	(3.97-5.47)	5.07	(4.21-6.10)	2.43	(2.37-2.49)
[HT, HC, AL, DEP]	0.66	(0.62-0.72)	0.75	(0.69-0.82)	0.82	(0.70-0.96)	0.73	(0.62-0.88)	0.81	(0.79-0.83)
[ST, HT, HC, AL]	1.52	(1.41-1.64)	1.72	(1.57-1.88)	2.15	(1.81-2.57)	1.88	(1.52-2.33)	1.26	(1.23-1.29)
[HT, HC, AL, OA]	0.40	(0.37-0.42)	0.42	(0.39-0.45)	0.42	(0.37-0.48)	0.42	(0.37-0.48)	0.57	(0.55-0.58)
[HT, HC, DEP, DIA]	1.52	(1.40-1.64)	1.73	(1.58-1.90)	2.00	(1.70-2.36)	1.92	(1.59-2.31)	1.30	(1.27-1.34)
[HT, HC, OA, DIA]	1.01	(0.94-1.08)	1.08	(1.00-1.17)	1.15	(1.01-1.31)	1.22	(1.05-1.41)	0.91	(0.88-0.93)
[HT, HC, CAN, DIA]	5.90	(5.45-6.38)	6.34	(5.81-6.92)	7.53	(6.51-8.70)	7.48	(6.36-8.80)	3.48	(3.39-3.56)

^a All portfolios contain the HD diagnosis.



Supplemental Fig. 1 Graphical representation of disease-disease interactions in the disease interactions only (DIO) model. A ribbon connects chronic diseases that have any significant interaction ($p \leq 0.001$) between them. The connection's width corresponds to the number of HD-diagnosed individuals obtaining both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between two diseases. The ribbon chart is ordered by number of connections between diseases, starting from AL with five connections all the way to CAN which interacts with all of the additional diseases.

AL = allergies, BP = back pain, JD = joint disease, OA = osteoarthritis, HT = hypertension, SCH = schizophrenia, DEP = depression, HC = high cholesterol, DIA = diabetes, ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, CAN = cancer.



Supplemental Fig. 2 Graphical representation of disease-disease interactions in the stable (S) model. A ribbon connects chronic diseases that have any significant interaction ($p \leq 0.001$) between them. The connection's width corresponds to the number of HD-diagnosed individuals obtaining both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between two diseases. The ribbon chart is ordered by number of connections between diseases, starting from HC with one connection all the way to COPD which interacts with 11 of the additional diseases. Note that back pain does not interact with any chronic disease in this model.

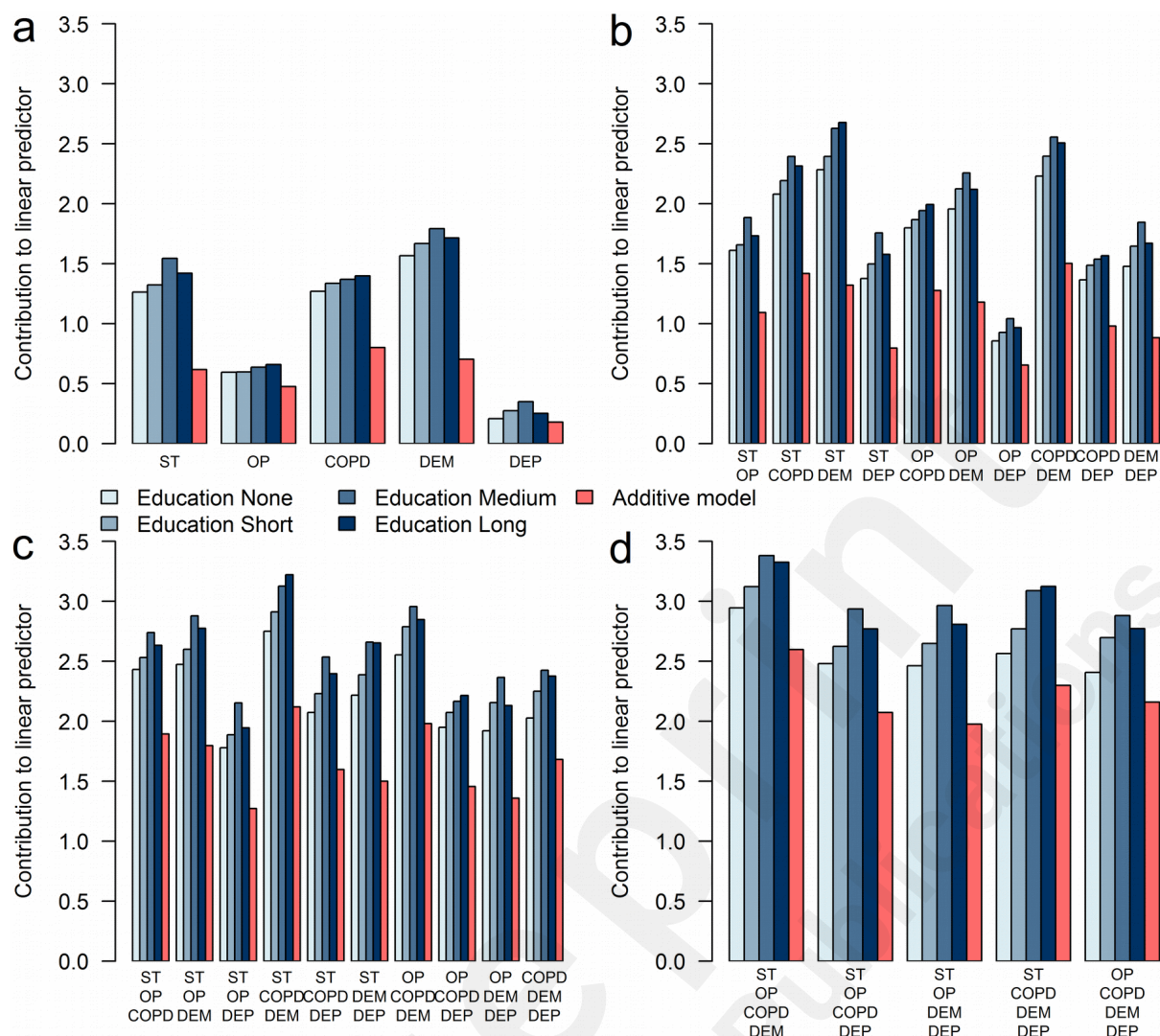
AL = allergies, JD = joint disease, OA = osteoarthritis, HT = hypertension, SCH = schizophrenia, DEP = depression, HC = high cholesterol, DIA = diabetes, ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, CAN = cancer.

Supplemental Table 5 Diagnosis-diagnosis interactions identified across the all interactions (ALL) model, the disease interactions only (DIO) model, and the stable (S) model. A cell in the table indicates under which models arising from the different variable selection procedures an interaction between the row and column condition is identified. Due to symmetry, only half of the table is presented.

ST = stroke, HT = hypertension, HC = high cholesterol, AL = allergies, JD = joint disease, OP = osteoporosis, OA = osteoarthritis, BP = back pain, CAN = cancer, COPD = chronic obstructive pulmonary disease, DEM = dementia, SCH = schizophrenia, DEP = depression, DIA = diabetes.

Chronic condition	HT	HC	AL	JD	OP	OA	BP	CAN	COPD	DEM	SCH	DEP	DIA
ST	ALL, DIO, S	ALL, DIO		ALL, DIO, S	ALL, DIO, S	DIO	ALL, DIO	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S
HT		ALL, DIO	ALL, DIO, S		ALL, DIO, S	ALL, DIO		ALL, DIO, S	ALL, DIO, S	ALL, DIO		ALL	S
HC				ALL, DIO	ALL, DIO	ALL	ALL	ALL, DIO	ALL, DIO	ALL, DIO, S	ALL, DIO	ALL, DIO	ALL, DIO
AL								ALL, DIO, S	ALL, DIO, S	DIO		ALL	ALL, DIO
JD					ALL, DIO, S			ALL, DIO, S	ALL, DIO, S	ALL, DIO			ALL, DIO
OP						ALL, DIO	ALL, DIO	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S
OA							DIO	ALL, DIO, S	DIO, S	ALL, DIO			
BP								ALL, DIO					ALL, DIO
CAN									ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	ALL, DIO, S

COPD										ALL, DIO, S	ALL, DIO, S	ALL, DIO, S	AL L, DIO , S
DEM											ALL, DIO, S	ALL, DIO, S	AL L, DIO , S
SCH												ALL, DIO, S	AL L, DIO , S
DEP													AL L, DIO , S



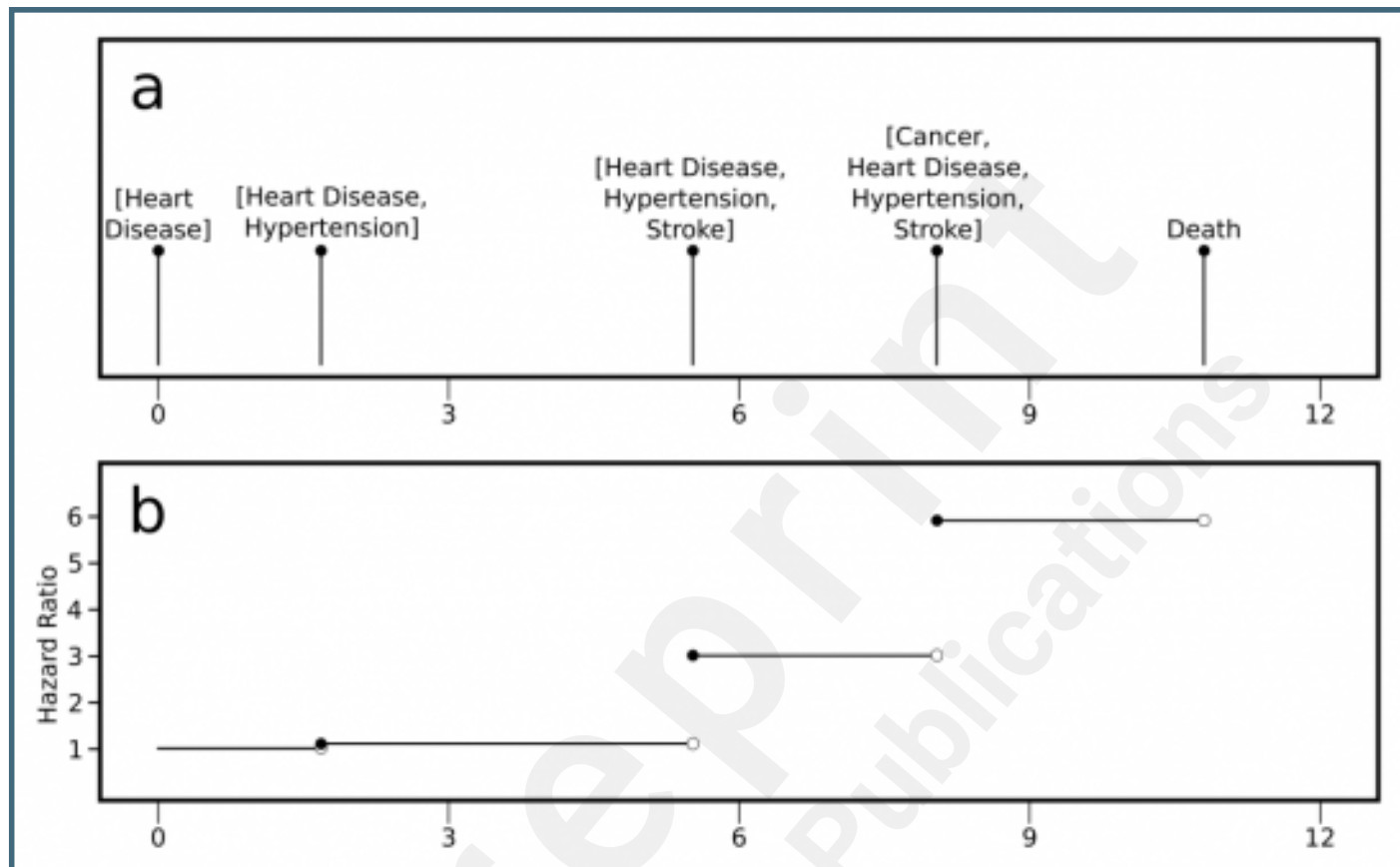
Supplemental Fig. 3 Effects of disease portfolio dyads (a), triads (b), tetrads (c) and pentads (d) involving stroke, osteoporosis, COPD, dementia and depression. Effects are displayed for females of varying educational attainment levels at the log hazard rate scale. Comparisons are made to a female of the corresponding educational attainment level who solely has HD. Effects are presented for the ALL model (different shades of blue) and the OME model (red). All comparisons are made at mean age and calendar time. Note that the HD condition is present in all disease portfolios.

ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, DEP = depression.

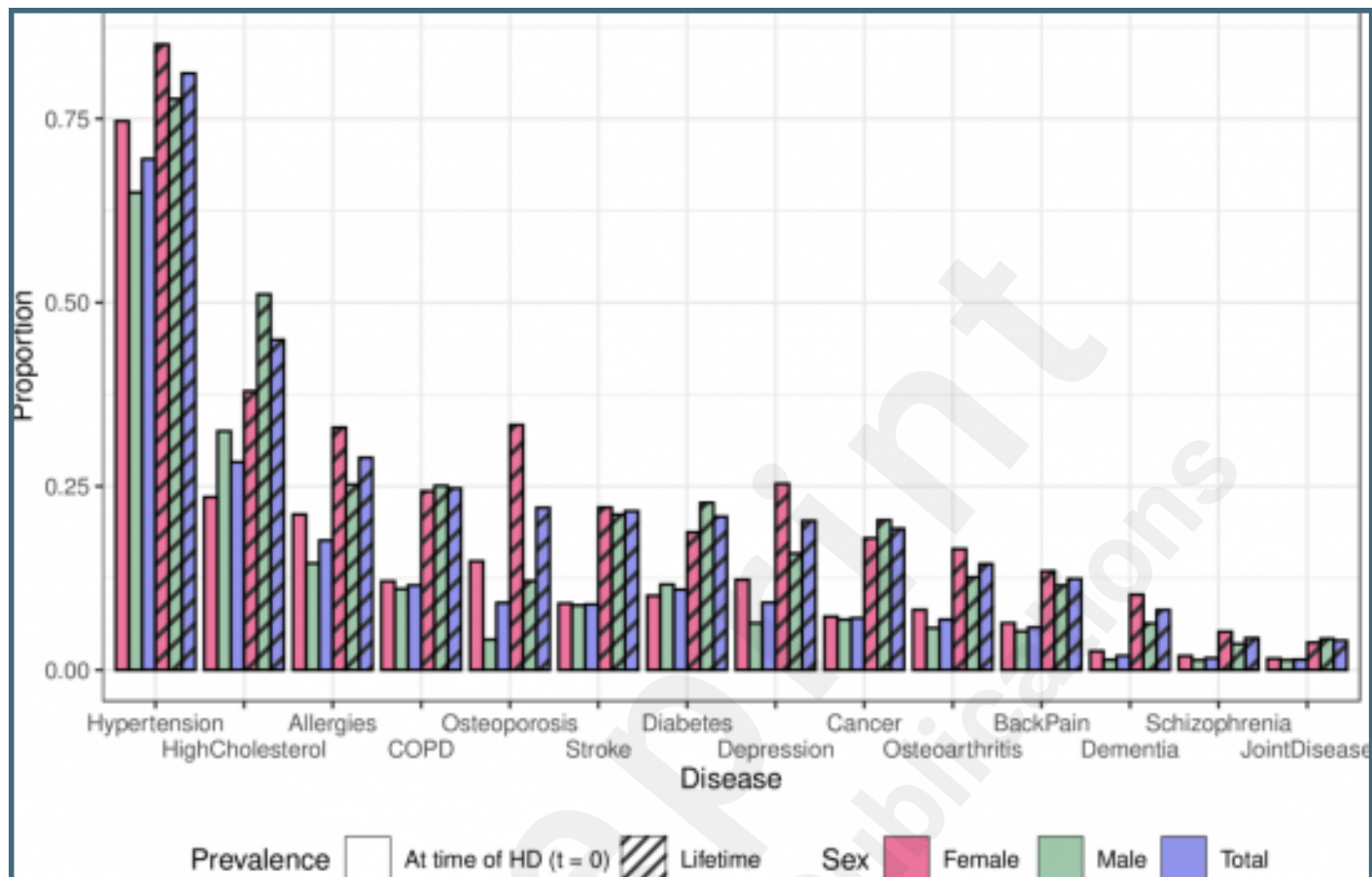
Supplementary Files

Figures

Example of how the statistical model works. (a) Illustration of an event sequence where a hypothetical HD patient gets the diagnosis of HD at time 0 and subsequently hypertension, stroke and cancer diagnoses at different times (measured in years following HD) before death. (b) The corresponding longitudinal development of the hazard ratio of the subject, relative to a theoretical non-multimorbid patient who solely has HD.

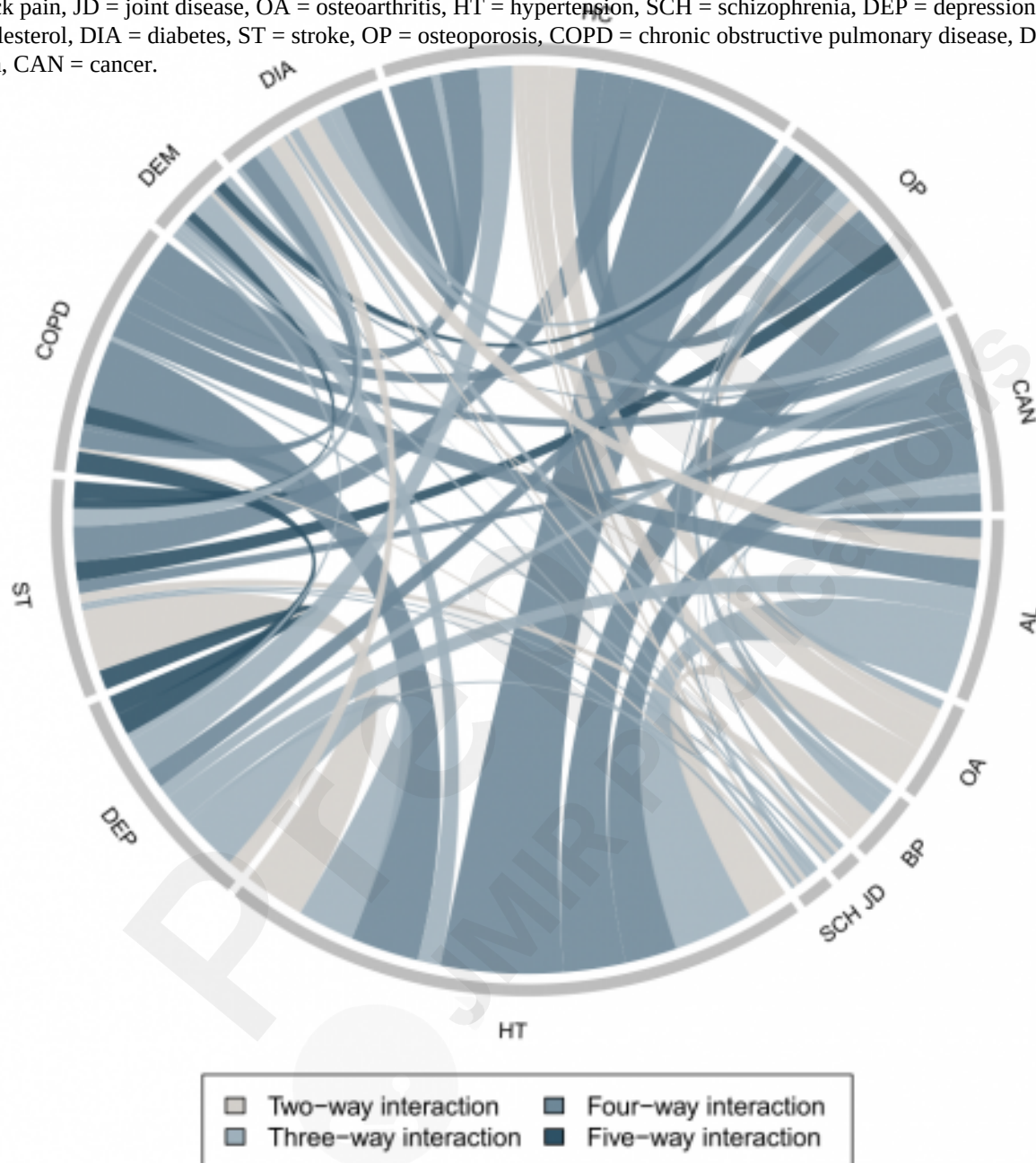


Diagnosis prevalence according to sex. Prevalence is reported at the time of HD diagnosis and for the entire span of the observed disease trajectories (Lifetime).

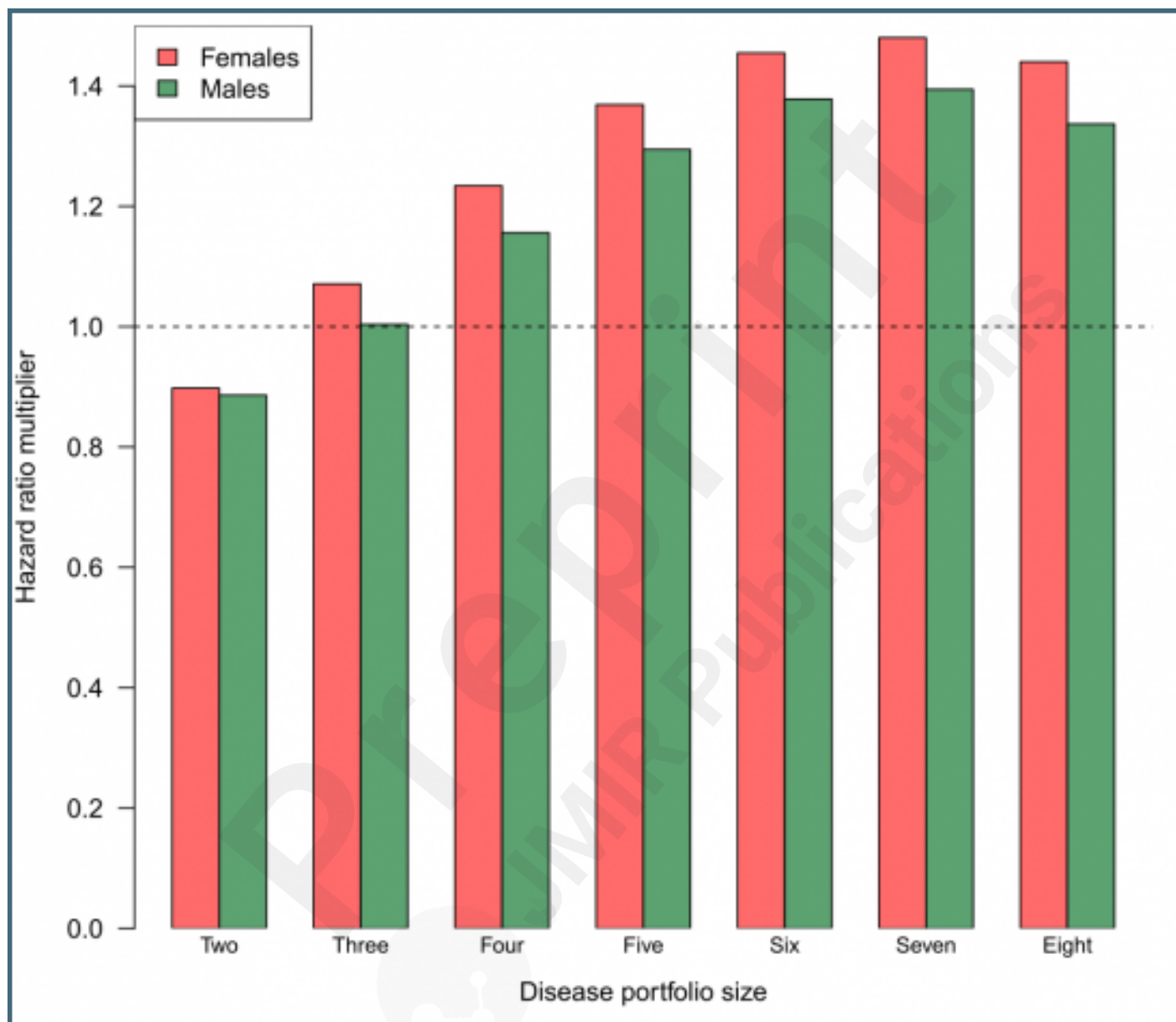


Preprint
JMIR Publications

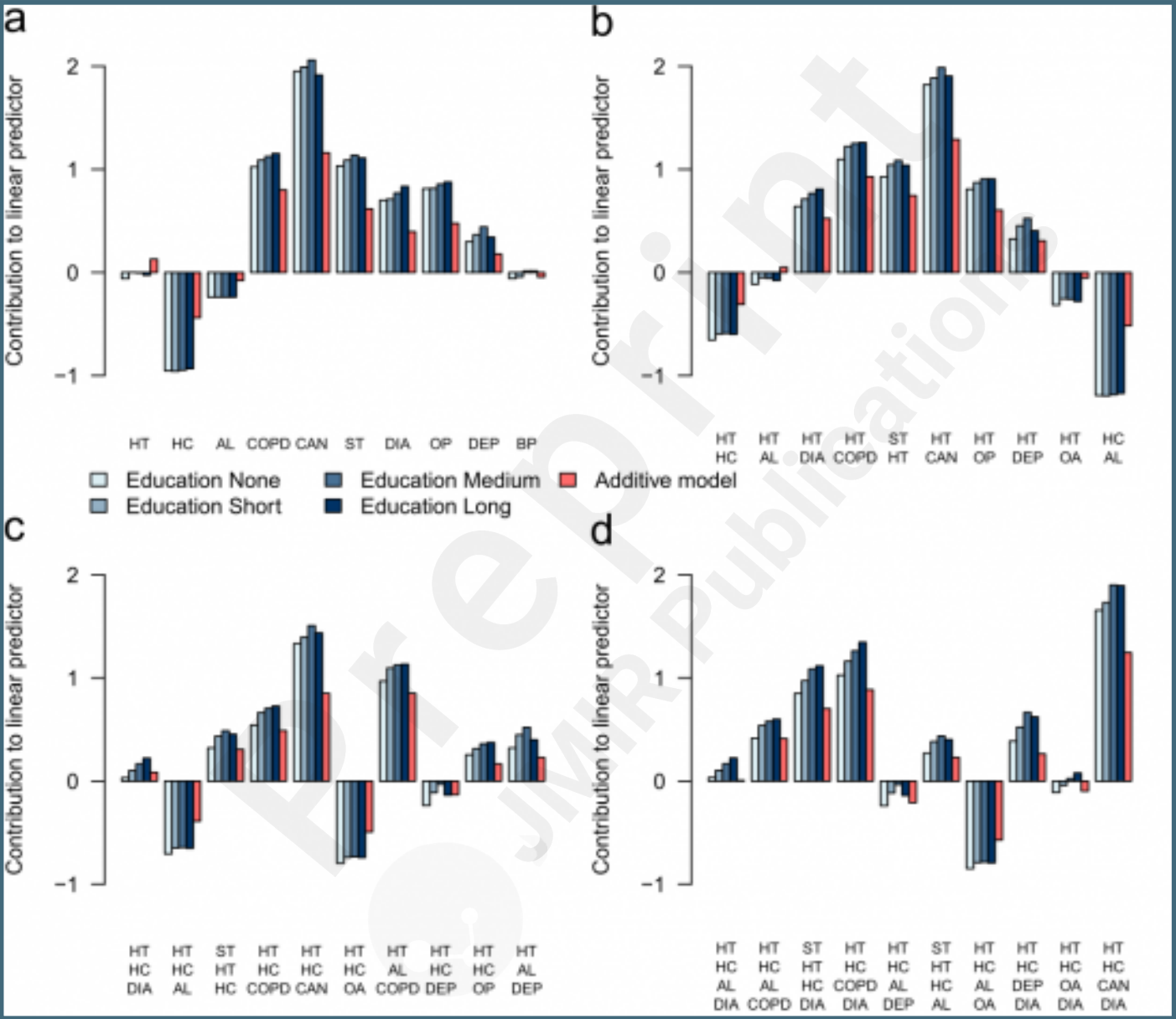
Graphical representation of disease-disease interactions in the all interactions (ALL) model. A ribbon connects chronic diseases that have any significant interaction ($p \leq 0.001$) between them. The connection's width corresponds to the number of HD-diagnosed individuals obtaining both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between two diseases. The ribbon chart is ordered by number of connections between diseases, starting from AL with five connections all the way to CAN which interacts with all of the additional diseases. AL = allergies, BP = back pain, JD = joint disease, OA = osteoarthritis, HT = hypertension, SCH = schizophrenia, DEP = depression, HC = high cholesterol, DIA = diabetes, ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, CAN = cancer.



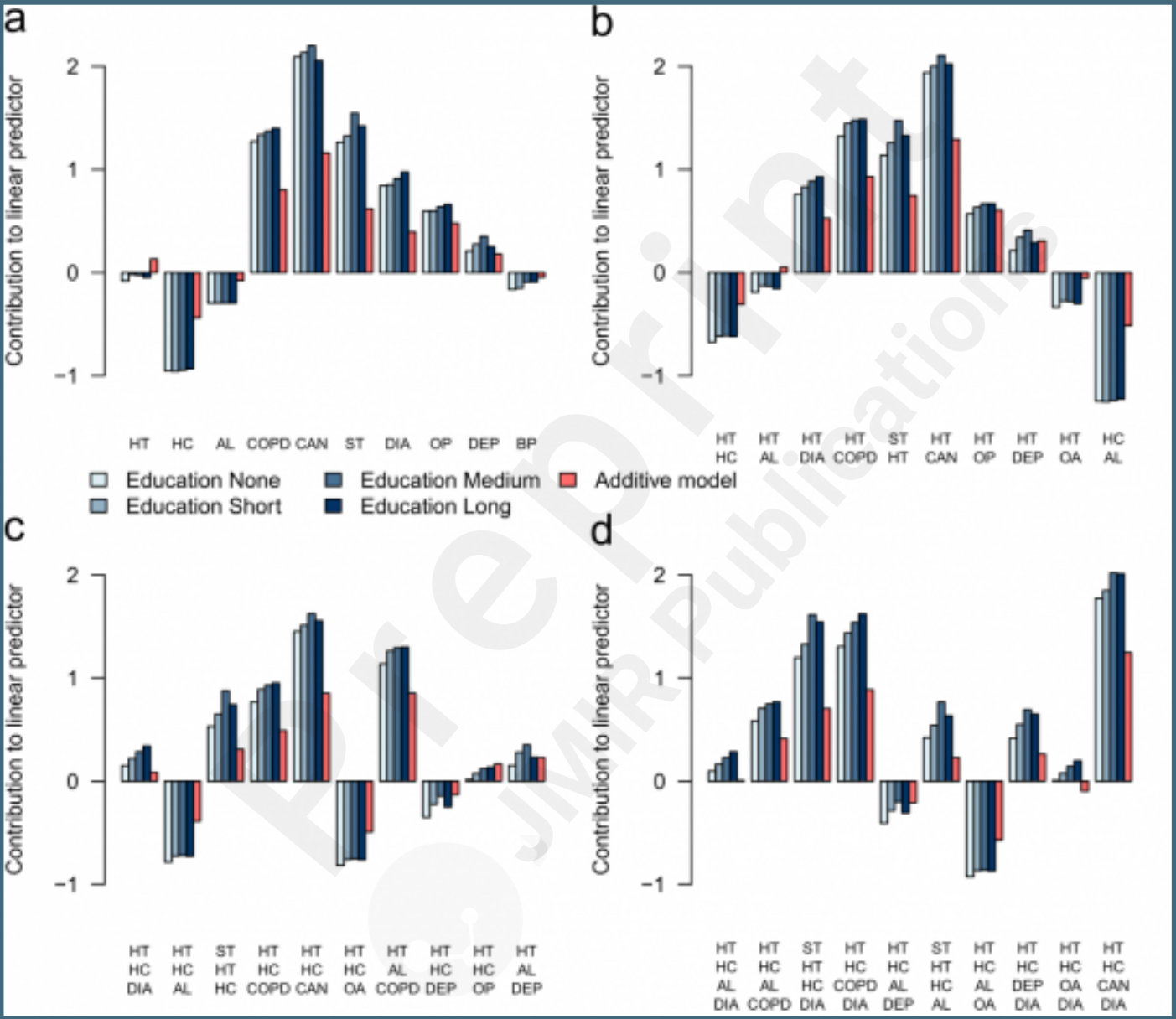
Difference in effect estimates for disease portfolios of increasing sizes for females and males. Each bar represents a weighted average of the differences in effects between the additive OME model and the ALL model on the hazard scale $\exp(\beta_{\text{ALL}} - \beta_{\text{OME}})$. The bars thus indicate the average multiplier required to convert the HR from the OME model into the HR from the ALL model. The weights were determined based on the occurrence of each specific disease portfolio across the different educational attainment levels for each sex.



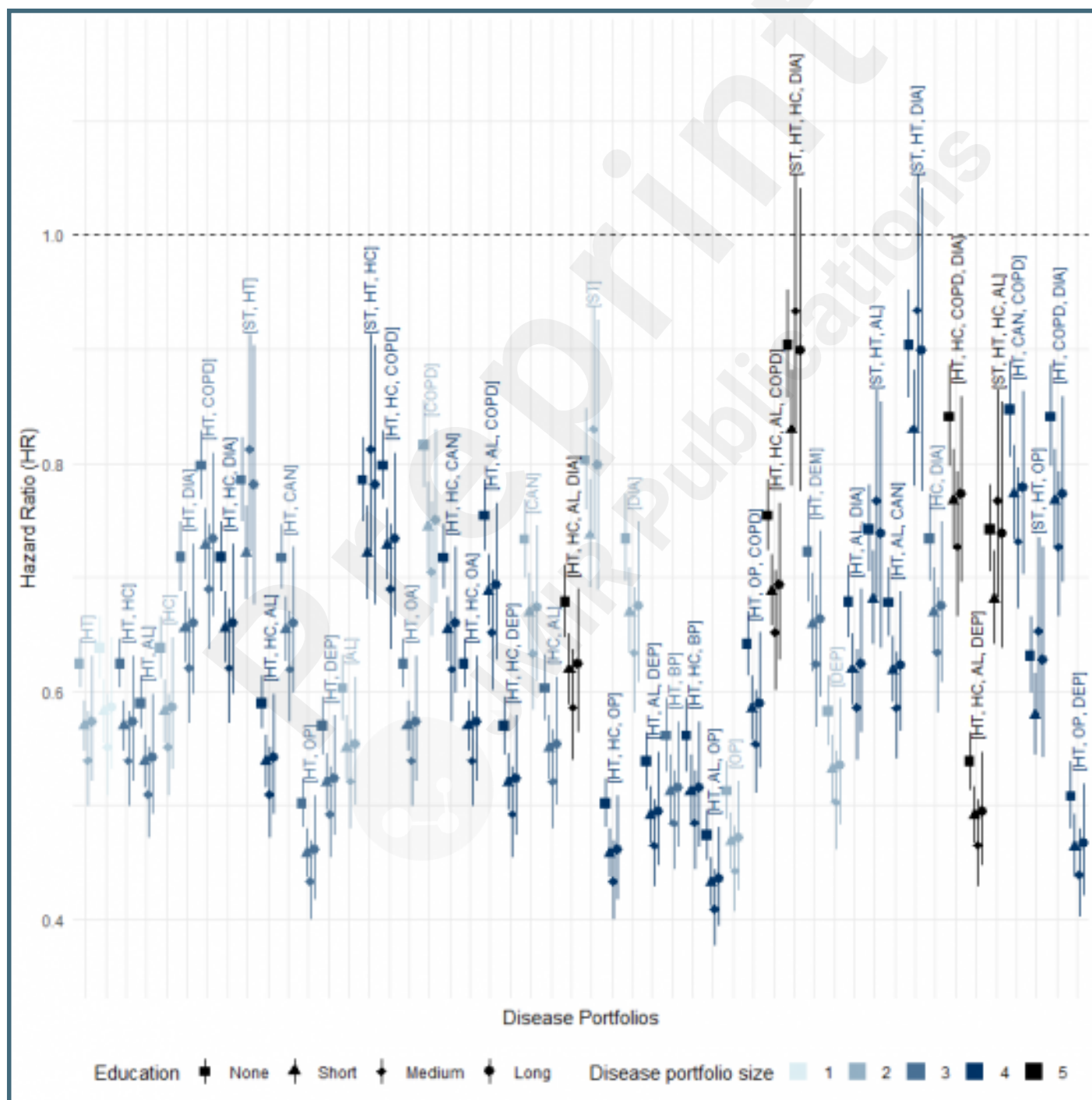
Effects of the ten most frequent disease portfolio dyads (a), triads (b), tetrads (c) and pentads (d). Effects are displayed for males of varying educational attainment levels at the log hazard rate scale. Comparisons are made to a male of the corresponding educational attainment level who solely has HD. Effects are presented for the ALL model (different shades of blue) and the OME model (red). All comparisons are made at mean age and calendar time. Note that HD is present in all disease portfolios. HT = hypertension, HC = high cholesterol, AL = allergies, COPD = chronic obstructive pulmonary disease, CAN = cancer, ST = stroke, DIA = diabetes, OP = osteoporosis, DEP = depression, BP = back pain, OA = osteoarthritis.



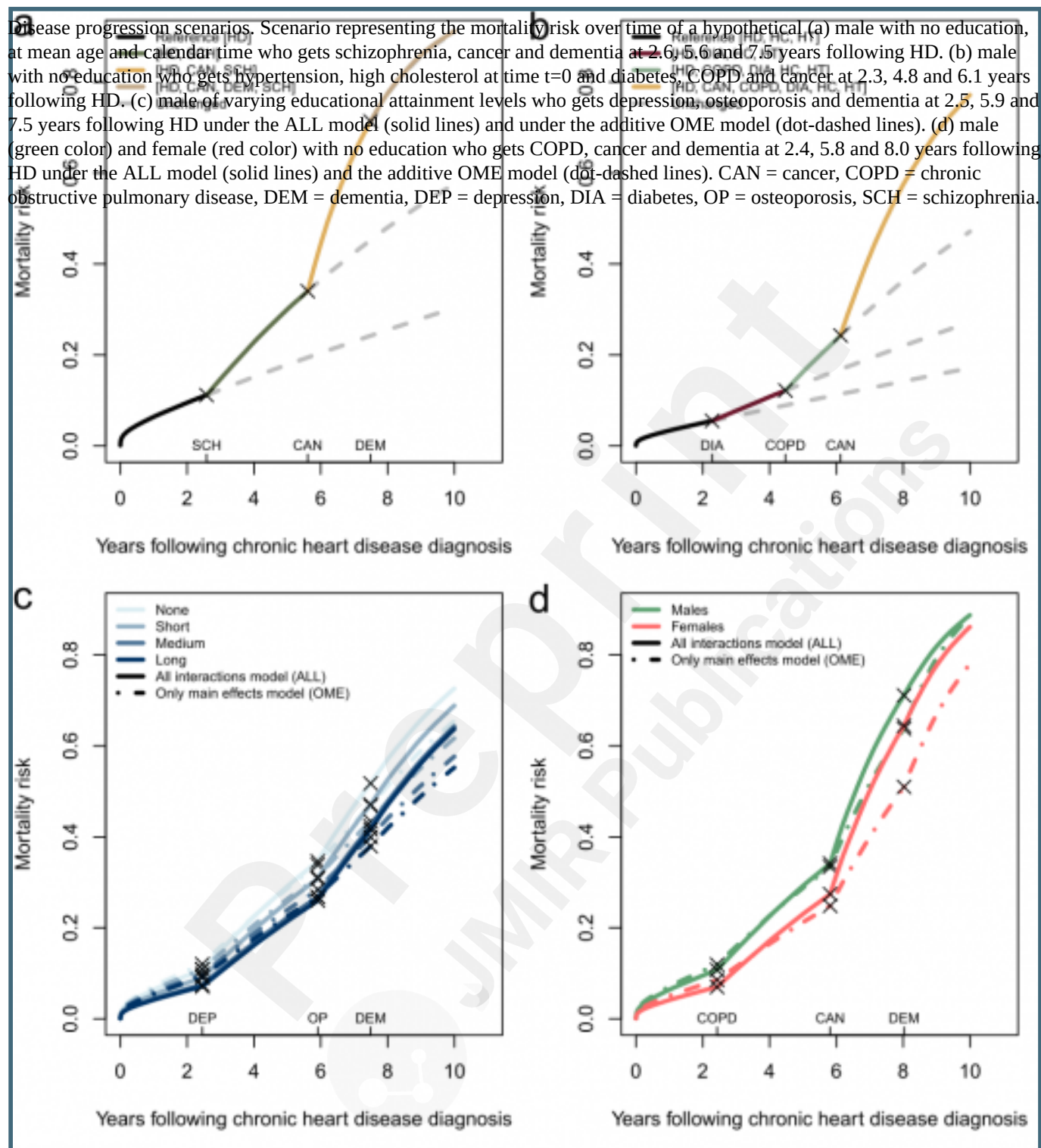
Effects of the ten most frequent disease portfolio dyads (a), triads (b), tetrads (c) and pentads (d). Effects are displayed for females of varying educational attainment levels at the log hazard rate scale. Comparisons are made to a female of the corresponding educational attainment level who solely has HD. Effects are presented for the ALL model (different shades of blue) and the OME model (red). All comparisons are made at mean age and calendar time. Note that HD is present in all disease portfolios. HT = hypertension, HC = high cholesterol, AL = allergies, COPD = chronic obstructive pulmonary disease, CAN = cancer, ST = stroke, DIA = diabetes, OP = osteoporosis, DEP = depression, BP = back pain, OA = osteoarthritis.



Hazard ratios of female sex (vs male) by disease portfolio and educational attainment level. Estimates for the 50 most common disease portfolios are shown with 99.9% confidence intervals. The estimates are presented for each of the educational attainment levels: None, short, medium and long, indicated by different shapes and always in ascending order from none to long. The reference group comprises males with the same disease portfolio and educational attainment level. The disease portfolios are ordered by prevalence from left to right with [HT] the most frequent disease portfolio. Note that all portfolios contain the HD condition, so it is not labelled in the plot. Therefore, the disease portfolio without a label in the plot (the second from the left) correspond to the disease portfolio with only HD. HT = hypertension, HC = high cholesterol, AL = allergies, DIA = diabetes, COPD = chronic obstructive pulmonary disease, ST = stroke, CAN = cancer, OP = osteoporosis, DEP = depression, OA = osteoarthritis, BP = back pain, DEM = dementia.



Preprint
JMIR Publications



Multimedia Appendixes

Graphical representation of disease-disease interactions in the disease interactions only (DIO) model. A ribbon connects chronic diseases that have any significant interaction ($p \leq 0.001$) between them. The connection's width corresponds to the number of HD-diagnosed individuals obtaining both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between two diseases. The ribbon chart is ordered by number of connections between diseases, starting from AL with five connections all the way to CAN which interacts with all of the additional diseases. AL = allergies, BP = back pain, JD = joint disease, OA = osteoarthritis, HT = hypertension, SCH = schizophrenia, DEP = depression, HC = high cholesterol, DIA = diabetes, ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, CAN = cancer.

URL: <http://asset.jmir.pub/assets/6a0009a4433817fca34094d5ab203a97.png>

Graphical representation of disease-disease interactions in the stable (S) model. A ribbon connects chronic diseases that have any significant interaction ($p \leq 0.001$) between them. The connection's width corresponds to the number of HD-diagnosed individuals obtaining both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between two diseases. The ribbon chart is ordered by number of connections between diseases, starting from HC with one connection all the way to COPD which interacts with 11 of the additional diseases. Note that back pain does not interact with any chronic disease in this model. AL = allergies, JD = joint disease, OA = osteoarthritis, HT = hypertension, SCH = schizophrenia, DEP = depression, HC = high cholesterol, DIA = diabetes, ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, CAN = cancer.

URL: <http://asset.jmir.pub/assets/5a33fddb478e62551d11958ecd387c5e.png>

Effects of disease portfolio dyads (a), triads (b), tetrads (c) and pentads (d) involving stroke, osteoporosis, COPD, dementia and depression. Effects are displayed for females of varying educational attainment levels at the log hazard rate scale. Comparisons are made to a female of the corresponding educational attainment level who solely has HD. Effects are presented for the ALL model (different shades of blue) and the OME model (red). All comparisons are made at mean age and calendar time. Note that the HD condition is present in all disease portfolios. ST = stroke, OP = osteoporosis, COPD = chronic obstructive pulmonary disease, DEM = dementia, DEP = depression.

URL: <http://asset.jmir.pub/assets/8d80ba005fe93e09922507573ff45aa3.png>