cancers

MDPI

Article

Multi-Morbidity and Risk of Breast Cancer among Women in the UK Biobank Cohort

Afi Mawulawoe Sylvie Henyoh [1,*], Rodrigue S. Allodji [1], Florent de Vathaire [1], Marie-Christine Boutron-Ruault [2], Neige M. Y. Journy [1] and Thi-Van-Trinh Tran [1,*]

[1] Radiation Epidemiology Group, Center for Research in Epidemiology and Population Health, INSERM U1018, Paris Sud-Paris Saclay University, Gustave Roussy, 94800 Villejuif, France
[2] Health across Generations Team, Center for Research in Epidemiology and Population Health, INSERM U1018, Paris Sud-Paris Saclay University, Gustave Roussy, 94800 Villejuif, France
* Correspondence: afi.henyoh@irsn.fr (A.M.S.H.); thi-van-trinh.tran@nih.gov (T.-V.-T.T.)

Simple Summary: (Multi-)Morbidity shares common biological mechanisms or risk factors with breast cancer. However, the risk of breast cancer among women with (multi-)morbidity remains unclear. In this study, using data of 239,436 female participants aged 40–69 of the UK Biobank cohort, we identified five chronic disease patterns: no-predominant morbidity, psychiatric morbidities, respiratory/immunological morbidities, cardiovascular/metabolic morbidities, and unspecific morbidities. After a median follow-up of 7 years, 5326 women developed breast cancer. We found no association between breast cancer risk and either the number of chronic diseases or chronic disease patterns, apart from an increased risk among women aged younger than 50 with a psychiatric pattern. Women with any multi-morbidity were more likely to die or to be diagnosed with other cancers. Our findings suggest that multi-morbidity may not be a key factor to help identify patients at an increased risk of breast cancer.

Abstract: (Multi-)Morbidity shares common biological mechanisms or risk factors with breast cancer. This study aimed to investigate the association between the number of morbidities and patterns of morbidity and the risk of female breast cancer. Among 239,436 women (40–69 years) enrolled in the UK Biobank cohort who had no cancer history at baseline, we identified 35 self-reported chronic diseases at baseline. We assigned individuals into morbidity patterns using agglomerative hierarchical clustering analysis. We fitted Cox models to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for breast cancer risk. In total, 58.4% of women had at least one morbidity, and the prevalence of multi-morbidity was 25.8%. During a median 7-year follow-up, there was no association between breast cancer risk (5326 cases) and either the number of morbidities or the identified clinically relevant morbidity patterns: no-predominant morbidity (reference), psychiatric morbidities (HR = 1.04, 95%CI 0.94–1.16), respiratory/immunological morbidities (HR = 0.98, 95%CI 0.90–1.07), cardiovascular/metabolic morbidities (HR = 0.93, 95%CI 0.81–1.06), and unspecific morbidities (HR = 0.98, 95%CI 0.89–1.07), overall. Among women younger than 50 years of age only, however, there was a significant association with psychiatric morbidity patterns compared to the no-predominant morbidity pattern (HR = 1.25, 95%CI 1.02–1.52). The other associations did not vary when stratifying by age at baseline and adherence to mammography recommendations. In conclusion, multi-morbidity was not a key factor to help identify patients at an increased risk of breast cancer.

Keywords: morbidity; morbidity patterns; breast cancer; incidence; cohort study; multiple correspondence analysis; cluster analysis

Citation: Henyoh, A.M.S.; Allodji, R.S.; de Vathaire, F.; Boutron-Ruault, M.-C.; Journy, N.M.Y.; Tran, T.-V.-T. Multi-Morbidity and Risk of Breast Cancer among Women in the UK Biobank Cohort. *Cancers* **2023**, *15*, 1165. https://doi.org/10.3390/cancers15041165

Academic Editor: Jahar Bhowmik

Received: 30 December 2022
Revised: 7 February 2023
Accepted: 8 February 2023
Published: 11 February 2023

1. Introduction

Breast cancer is the most common female cancer, with 2,088,849 new cases worldwide in 2018, accounting for 11.6% of incident cancer cases [1]. Despite decades of intensive

research effort, only about 70% of the disease occurrence is explained by well-established risk factors [2]. Most of the identified risk factors are not readily modifiable [2–4], leading to a need for additional research to better understand etiologic processes.

In developed countries, most breast cancer cases are diagnosed among women of middle age or older [5], coinciding with the occurrence of other long-term morbidities [6,7]. Previous studies have suggested associations between breast cancer risk and specific chronic diseases, such as endocrine disorders [8,9], inflammatory conditions [10], autoimmune diseases [11], and cardiovascular diseases (CVDs) [12], especially among postmenopausal women. The underlying mechanisms of these associations could involve shared common physiopathological pathways (e.g., estrogen-related pathways, inflammation pathways) [13–16], shared genetic predispositions, shared risk factors (e.g., obesity, physical inactivity) [17], and medications (e.g., aspirin) [18].

As people get older, they often develop two or more chronic diseases. With an aging population, the number of people experiencing several multi-morbidities is rising globally [6,7,19–21]. In the general population, co-existing morbidities could be classified into common clinically meaningful patterns [22,23]. Sharing underlying biological mechanisms and/or sets of risk factors, the morbidities in the same cluster often interact mutually, which complicates treatments and management and increases the risk of adverse events above and beyond the sum of the risk of individual disease [24]. Being diagnosed with multi-morbidity is also associated with an increased likelihood of being subjected to breast cancer screening [25–27], which may lead to increased surveillance of breast cancer incidence. Thus, it is necessary to consider patterns of morbidity, in addition to associations with single chronic diseases, with breast cancer risk.

However, to date, there is no epidemiological evidence as to whether and to what extent breast cancer risk varies according to different patterns of morbidity. In this context, our study aimed to investigate the association between the number of morbidities and patterns of morbidity and the risk of female breast cancer.

2. Materials and Methods

2.1. Data Source and Study Design

The UK Biobank is a prospective population-based cohort that recruited 273,375 women, aged 40 to 69 years, from March 2006 to July 2010 [28]. Individuals were invited to participate on a voluntary basis and provided electronic informed consent for data provision and linkage. The baseline data assessment included self-reported data on personal and family medical history, lifestyle, hormone-related factors, and sociodemographic characteristics. Additional anthropometric measurements were performed. The cohort additionally retrieved individual information from the national cancer and death registries.

2.2. Study Population

We excluded women with any cancer diagnosis prior to baseline except non-melanoma skin cancer (n = 29,332), women who underwent a mastectomy prior to baseline (n = 2457), and women with less than one year of follow-up (n = 2150), leaving 239,436 women in the final analysis (Figure 1).

2.3. Baseline Morbidity Identification

Based on an established list of morbidities, which was originally designed by Barnett et al. [19] to measure multi-morbidity in a large population-based dataset and subsequently validated in the UK Biobank cohort (Appendix A, Table A1) [29], we defined 35 morbidities based on baseline self-reported health conditions (Figure 2). For each woman, we computed the total number of morbidities and categorized them as none/one/multi-morbidity (at least two morbidities).

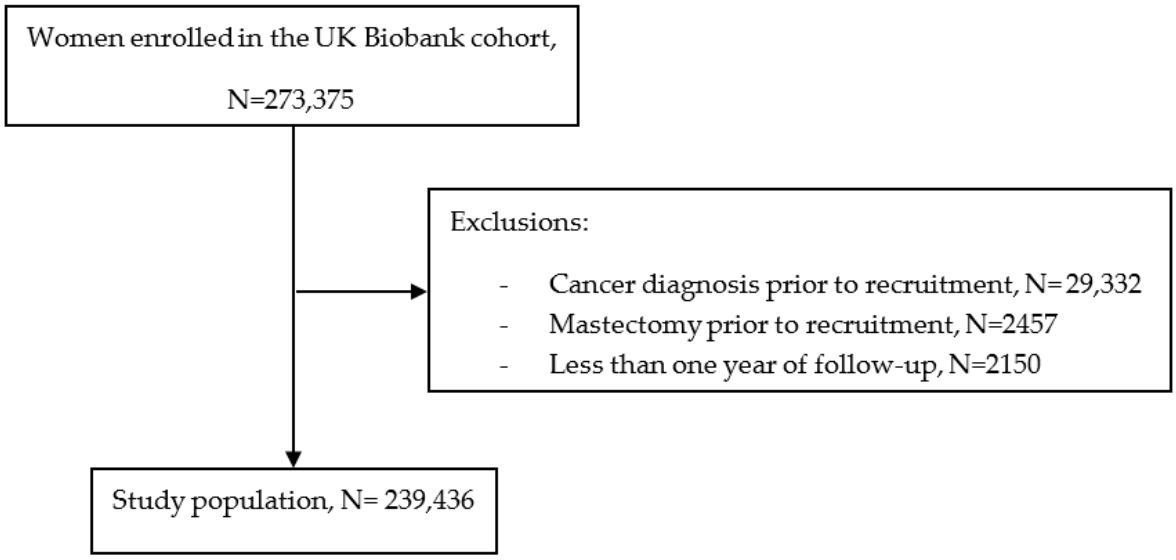

Figure 1. Flow chart of the study population.

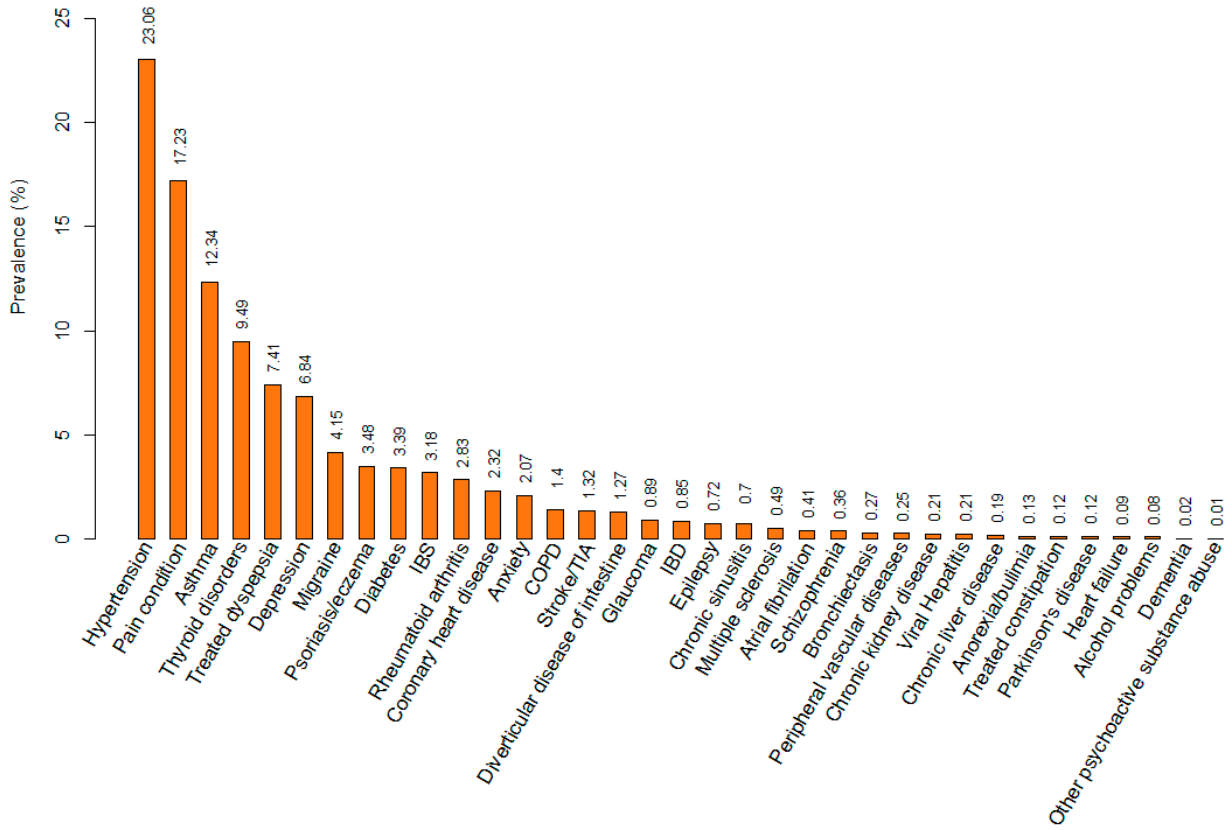

Figure 2. Morbidities identified among cancer-free UK Biobank women at recruitment.

2.4. Breast Cancer Ascertainment

We defined breast cancer as a diagnosis of invasive or in situ breast cancer, using the international classification of diseases (ICD) versions 9 and 10 (ICD-10: C50 or D05; ICD-9: 174 or 2330). We considered only breast cancer cases that were the first cancer diagnosed.

2.5. Baseline Confounding Factors

All confounding factors (age at menarche, age at menopause, menopausal hormone therapy use, oral contraceptive use, parity and age at first birth, body mass index (BMI),

ethnicity, Townsend score, level of physical activity, alcohol consumption) were measured/collected at baseline. We selected well-established breast cancer risk factors based on previous studies [30,31]. We also selected variables that were statistically significantly associated with both morbidity and breast cancer risk (p-value < 0.05) as confounders if their inclusion in the age-adjusted Cox models changed the hazard ratio by 5% or more [32]. See Appendix A, Table A2 for more details on the variables of interest, their definition, and information sources.

2.6. Statistical Analysis

2.6.1. Multiple Correspondence Analysis (MCA) and Cluster Analysis (See Appendix B)

Among 35 baseline self-report morbidities, we included only morbidities with a prevalence of more than 1% (Figure 2) to obtain stable clustering results [33]. We used MCA [34,35] and cluster analysis to identify morbidity patterns. MCA can produce the input data for the cluster analysis, while reducing noise by excluding unnecessary dimensions that do not contribute significantly to the cluster's classification. We determined the optimal number of dimensions to extract based on the elbow rule in the Scree plot [34] and Horn's parallel analysis for common factor analysis [36].

Using the numerical outputs of the MCA, we performed agglomerative hierarchical clustering (AHC) preceded by K-means clustering with 2000 initial cluster seeds [37], through the HCPC function of the Factominer package in R. This method allowed us to reduce the required memory allocations [38,39]. We considered the distance between points in Euclidean space as the distance metric [40], and Ward's method was used to create homogeneous clusters by fusion [36]. We chose the optimal number of clusters, i.e., the identified morbidity patterns and assessed cluster quality, using the Davies–Bouldin [41] and the GAP indexes [42]. The optimal number of clusters was the one that corresponded to the minimum value of the Davies–Bouldin index and to the maximum Gap statistics index.

Within each cluster, we computed the observed/expected ratios ("O/E-ratios") for each single morbidity, i.e., the ratio between the prevalence of a given condition in a cluster and its prevalence in the overall study population. Similarly, we computed the exclusivity of each single morbidity, i.e., the number of individuals that had a given morbidity in a cluster over the number of individuals with the same morbidity in the whole study population. A morbidity was considered part of a given morbidity cluster when its O/E-ratio was $\geq$2 and its exclusivity was $\geq$25% [23,43]. We named the morbidity patterns based on the predominant morbidities in the clusters.

2.6.2. Association among the Number of Morbidities, Morbidity Patterns, and Breast Cancer Risk

The follow-up time started at the date of first registration at a UK Biobank center and ended at the date of the first cancer diagnosis (any cancer diagnosis, except non-melanoma skin cancer) or mastectomy, death, loss to follow-up, or 31 March 2016, whichever came first. We fitted Cox proportional hazard models to estimate hazard ratios and 95% confidence intervals (95%CIs) of breast cancer risk associated with each single pre-existing baseline morbidity included in the cluster analysis, the number of morbidities, and the morbidity patterns. The timescale was the follow-up time.

We graphically assessed the proportional hazards assumption using scaled Schoenfeld residuals plots and log linearity assumption (for quantitative covariates) using Martingale residuals plots and deviance residuals plots. The final multivariable Cox models were adjusted for age at baseline, age at menarche, age at menopause, menopausal hormone therapy use, oral contraception use, parity and age at first birth, BMI, ethnicity, the Townsend score, level of physical activity, and alcohol consumption.

We tested the modifying effects of age at baseline, the adherence to the recommendations for breast cancer screening, the BMI, the socioeconomic status, the physical activity level, and the menopause status at baseline with the likelihood ratio test. We conducted several sensitivity analyses: (i) we restricted analyses to menopausal women; (ii) we

considered only invasive breast cancer as the outcome; (iii) we used the attained age as the timescale; (iv) we considered death and diagnosis of non-breast cancer as competing risks, using sub-distribution hazards models [44]; (v) we extracted 11 MCA dimensions, which accounted for more than 70% of the total variability among the study population, as recommended by Higgs [45]; we also extracted all dimensions, assuming they were all significant, and kept different numbers of clusters (3 and 4 clusters) with both 11 and all dimensions extracted.

All statistical analyses were performed using R version 4.1.0.

3. Results

In the study population, the median age at baseline was 57.7 years (interquartile range [IQR]: 50.2, 63.2). At least one morbidity was present in 58.4% of women at baseline, and the prevalence of multi-morbidity was 25.8%. Hypertension was the most prevalent morbidity (23.1%), followed by painful conditions (17.2%) and asthma (12.3%). The prevalence of obesity was 23.5%, and 23.5% of women had menopause after the age of 51 at baseline. Most women were postmenopausal (73.5%) and were adherent to breast cancer screening recommendations (66.6%) at baseline, as assessed at recruitment (Table 1, Figure 2). During a median follow-up time of 7.1 years (IQR: 6.4, 7.8), 5,326 women developed breast cancer (2.0%).

3.1. Description of Morbidity Patterns

We considered the first five MCA dimensions (see Supplementary, Figures S1–S3), which explained 39% of the total variance, as input to the clustering algorithms. We identified five baseline morbidity patterns (see Supplementary, Figures S4 and S5 & Table S1), named as follows: Pattern 1—no-predominant morbidity, pattern 2—psychiatric morbidities, pattern 3—respiratory/immunological morbidities, pattern 4—cardiovascular/metabolic morbidities, pattern 5—unspecific morbidities (see Table 1).

3.1.1. Pattern 1: No-Predominant Morbidity [n = 159,083 (66.4%), 3534 Breast Cancer Cases (2.0% of Cases)]

The median age at baseline was 57.4 years (IQR: 49.9, 63.0), and the median follow-up time was 7.1 years (IQR: 6.4, 7.8). There was no morbidity with an O/E ratio $\geq$ 2. The main features of this pattern were the low rate of multi-morbidity (6.9%) and the high rate of the absence of morbidity (62.6%).

3.1.2. Pattern 2: Psychiatric Morbidities [n = 16,627 (7.0%), 381 Breast Cancer Cases (2.0% of Cases)]

The median age at baseline was 55.7 years (IQR: 48.7, 61.7), and the median follow-up time was 7.0 years (IQR: 6.3, 7.8). Women with this pattern were predominantly diagnosed with anxiety and depression disorders.

3.1.3. Pattern 3: Respiratory/Immunological Morbidities [n = 27,920 (11.7%), 611 Breast Cancer Cases (2.0% of cases)]

The median age at baseline was 56.7 years (IQR: 49.1, 62.8), and the median follow-up time was 7.1 years (IQR: 6.4, 7.8). Women with this pattern were predominantly diagnosed with psoriasis/eczema, COPD, and asthma.

3.1.4. Pattern 4: Cardiovascular/Metabolic Morbidities [n = 11,041 (4.6%), 246 Breast Cancer Cases (2.0% of cases)]

The median age at baseline was 62.6 years (IQR: 57.2, 66.4), and the median follow-up time was 7.0 years (IQR: 6.3, 7.8). Women with this pattern were predominantly diagnosed with diabetes, stroke, and coronary–heart disease. The main features of this pattern were the high proportions of elderly (about 65% were 65 years or older at baseline), multi-morbidity (96.6%), and deprived people (37.1% of women with this pattern were in the quintile with the highest levels of deprivation).

Table 1. Characteristics of the overall study population and the identified baseline morbidity patterns.

Characteristics	Overall Study Population N = 239,436	Pattern 1: No-Predominant Morbidity N = 159,083	Pattern 2: Psychiatric Morbidities N = 16,627	Pattern 3: Respiratory/Immunological Morbidities N = 27,920	Pattern 4: Cardiovascular/Metabolic Morbidities N = 11,041	Pattern 5: Unspecific Morbidities N = 24,765	*p*-Value *
Year of follow-up, median (IQR)	7.1 (6.4, 7.8)	7.1 (6.4, 7.8)	7.0 (6.3, 7.8)	7.1 (6.4, 7.8)	7.0 (6.3, 7.8)	7.1 (6.4, 7.8)	<0.001
Breast cancer cases, n (%)	5326 (2)	3534 (2)	381 (2)	611 (2)	246 (2)	554 (2)	0.97
Number of comorbid conditions, n (%)							<0.001
None	99,614 (41.6)	99,614 (62.6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
One	77,994 (32.6)	48,489 (30.5)	6260 (37.6)	14,283 (51.2)	379 (3.4)	8583 (34.7)	
Two	38,424 (16.0)	10,145 (6.4)	5974 (35.9)	9156 (32.8)	4717 (42.7)	8432 (34.0)	
Three and more	23,404 (9.8)	835 (0.5)	4393 (26.4)	4481 (16.0)	5945 (53.8)	7750 (31.3)	
Morbidity, n (%)							
Stroke and transient ischemic attack (TIA)	3149 (1.3)	833 (0.5)	90 (0.5)	181 (0.6)	1761 (15.9)	284 (1.1)	<0.001
Diabetes	8122 (3.4)	1429 (0.9)	63 (0.4)	282 (1.0)	5924 (53.7)	424 (1.7)	<0.001
Coronary heart disease	5566 (2.3)	796 (0.5)	53 (0.3)	329 (1.2)	3978 (36.0)	410 (1.7)	<0.001
Migraine	9947 (4.2)	247 (0.2)	940 (5.7)	686 (2.5)	102 (0.9)	7972 (32.2)	<0.001
Diverticular disease of intestine	3048 (1.3)	0 (0.0)	10 (0.1)	109 (0.4)	247 (2.2)	2682 (10.8)	<0.001
Irritable bowel syndrome	7622 (3.2)	32 (0.0)	642 (3.9)	195 (0.7)	276 (2.5)	6477 (26.2)	<0.001
Rheumatoid arthritis	6778 (2.8)	0 (0.0)	81 (0.5)	201 (0.7)	4 (0.0)	6492 (26.2)	<0.001
Treated dyspepsia	17,733 (7.4)	6427 (4.0)	1704 (10.2)	2053 (7.4)	1807 (16.4)	5742 (23.2)	<0.001
Psoriasis or eczema	8344 (3.5)	0 (0.0)	773 (4.6)	5823 (20.9)	190 (1.7)	1558 (6.3)	<0.001
Chronic obstructive respiratory disease (COPD)	3355 (1.4)	0 (0.0)	7 (0.0)	3333 (11.9)	2 (0.0)	13 (0.1)	<0.001
Asthma	29,541 (12.3)	0 (0.0)	2311 (13.9)	21,708 (77.8)	2473 (22.4)	3049 (12.3)	<0.001
Anxiety	4964 (2.1)	0 (0.0)	4460 (26.8)	113 (0.4)	216 (2.0)	175 (0.7)	<0.001
Depression	16,368 (6.8)	0 (0.0)	13,362 (80.4)	424 (1.5)	1157 (10.5)	1425 (5.8)	<0.001
Thyroid disorders	22,718 (9.5)	13,277 (8.3)	1776 (10.7)	2213 (7.9)	2806 (25.4)	2646 (10.7)	<0.001
Hypertension	55,223 (23.1)	31,013 (19.5)	3647 (21.9)	6112 (21.9)	8505 (77.0)	5946 (24.0)	<0.001
Pain conditions	41,258 (17.2)	21,363 (13.4)	3665 (22.0)	4767 (17.1)	3132 (28.4)	8331 (33.6)	<0.001
Age at baseline, median (IQR)	57.7 (50.2, 63.2)	57.4 (49.9, 63.0)	55.7 (48.7, 61.7)	56.7 (49.1, 62.8)	62.6 (57.2, 66.4)	59.2 (51.9, 64.0)	<0.001
Family history of breast cancer, n (%)	25,330 (10.6)	16,858 (10.6)	1765 (10.6)	2885 (10.3)	1102 (10.0)	2720 (11.0)	0.035
BMI, n (%)							<0.001
<18.5	1803 (0.8)	1215 (0.8)	115 (0.7)	225 (0.8)	23 (0.2)	225 (0.9)	
18.5–25	92,857 (38.8)	66,570 (41.8)	5644 (33.9)	10,139 (36.3)	1547 (14.0)	8957 (36.2)	
25–30	87,381 (36.5)	58,431 (36.7)	6067 (36.5)	10,161 (36.4)	3581 (32.4)	9141 (36.9)	
>30	56,150 (23.5)	31,992 (20.1)	4725 (28.4)	7282 (26.1)	5799 (52.5)	6352 (25.6)	
Unknown	1245 (0.5)	875 (0.6)	76 (0.5)	113 (0.4)	91 (0.8)	90 (0.4)	
Adherence to breast cancer screening, n (%)							<0.001
<50 years of age	58,722 (24.5)	40,371 (25.4)	4902 (29.5)	7745 (27.7)	873 (7.9)	4831 (19.5)	
>50 years of age, >3 years ago	7929 (3.3)	5072 (3.2)	554 (3.3)	889 (3.2)	545 (4.9)	869 (3.5)	
>50 years of age, in the last 3 years	159,407 (66.6)	104,789 (65.9)	10,158 (61.1)	17,761 (63.6)	8937 (80.9)	17,762 (71.7)	
>50 years of age, never	8013 (3.3)	5384 (3.4)	631 (3.8)	943 (3.4)	348 (3.2)	707 (2.9)	
>50 years of age, unknown	5365 (2.2)	3467 (2.2)	382 (2.3)	582 (2.1)	338 (3.1)	596 (2.4)	
Age at menarche, median (IQR)	13.0 (12.0, 14.0)	13 (12.0, 14.0)	13 (12.0, 14.0)	13.0 (12.0, 14.0)	13 (12.0, 14.0)	13 (12.0, 14.0)	<0.001
Age at menopause [µ], median (IQR)	50.0 (47.0, 52.0)	50.0 (47.0, 52.0)	50.0 (45.5, 52.0)	50.0 (46.0, 52.0)	50.0 (45.0, 52.0)	50.0 (46.0, 52.0)	<0.001
Menopause status at baseline, n (%)							<0.001
Still had periods	63,488 (26.5)	44,275 (27.8)	4979 (29.9)	8152 (29.2)	951 (8.6)	5131 (20.7)	
Had menopause before the age of 45	25,659 (10.7)	14,768 (9.3)	2095 (12.6)	3356 (12.0)	2024 (18.3)	3416 (13.8)	
Had menopause between the age of 45 and 54	129,114 (53.9)	85,911 (54.0)	8332 (50.1)	14,084 (50.4)	6796 (61.6)	13,991 (56.5)	
Had menopause after the age of 54	21,175 (8.8)	14,129 (8.9)	1221 (7.3)	2328 (8.3)	1270 (11.5)	2227 (9.1)	

Table 1. *Cont.*

Characteristics	Overall Study Population N = 239,436	Pattern 1: No-Predominant Morbidity N = 159,083	Pattern 2: Psychiatric Morbidities N = 16,627	Pattern 3: Respiratory/Immunological Morbidities N = 27,920	Pattern 4: Cardiovascular/Metabolic Morbidities N = 11,041	Pattern 5: Unspecific Morbidities N = 24,765	*p*-Value *
Menopausal hormone therapy use ᵘ, n (%)							<0.001
Never	85,613 (48.7)	59,734 (52.0)	4572 (39.3)	8935 (45.2)	4485 (44.4)	7887 (40.2)	
Ever, less than 5 years duration	31,000 (17.6)	19,322 (16.8)	2553 (21.9)	3683 (18.6)	1620 (16.1)	3822 (19.5)	
Ever, 5 years and longer duration	47,233 (26.8)	28,799 (25.1)	3496 (30.0)	5759 (29.1)	2898 (28.7)	6281 (32.0)	
Ever, unknown duration	11,229 (6.4)	6386 (5.6)	975 (8.4)	1314 (6.6)	1004 (9.9)	1550 (7.9)	
Unknown status	874 (0.5)	567 (0.5)	52 (0.4)	77 (0.4)	84 (0.8)	94 (0.5)	
Oral contraception use, n (%)							<0.001
Never	44,767 (18.7)	29,175 (18.3)	2795 (16.8)	4818 (17.3)	3147 (28.5)	4832 (19.5)	
Ever, less than 10 years duration	87,270 (36.4)	57,671 (36.3)	5929 (35.7)	10,134 (36.3)	4074 (36.9)	9462 (38.2)	
Ever, 10 years and longer duration	84,462 (35.3)	57,626 (36.2)	6117 (36.8)	10,315 (36.9)	2505 (22.7)	7899 (31.9)	
Ever, unknown duration	22,542 (9.4)	14,354 (9.0)	1758 (10.6)	2628 (9.4)	1270 (11.5)	2532 (10.2)	
Unknown status	395 (0.2)	257 (0.2)	28 (0.2)	25 (0.1)	45 (0.4)	40 (0.2)	
Parity and age at first birth, n (%)							<0.001
None of live birth	44,601 (18.6)	29,572 (18.6)	3575 (21.5)	5497 (19.7)	1614 (14.6)	4343 (17.5)	
At least one birth before 30	150,386 (62.8)	98,115 (61.7)	10,088 (60.7)	17,341 (62.1)	8183 (74.1)	16,659 (67.3)	
At least one birth after age 30	43,302 (18.1)	30,569 (19.2)	2910 (17.5)	5003 (17.9)	1154 (10.5)	3666 (14.8)	
Unknown	1147 (0.5)	827 (0.5)	54 (0.3)	79 (0.3)	90 (0.8)	97 (0.4)	
Levels of physical activities, n (%)							<0.001
Low	76,618 (32.0)	47,554 (29.9)	5964 (35.9)	9211 (33.0)	4867 (44.1)	9022 (36.4)	
Moderate	85,403 (35.7)	57,868 (36.4)	5758 (34.6)	9893 (35.4)	3341 (30.3)	8543 (34.5)	
High	77,415 (32.3)	53,661 (33.7)	4905 (29.5)	8816 (31.6)	2833 (25.7)	7200 (29.1)	
Alcohol consumption, n (%)							<0.001
Never	22,751 (9.5)	12,842 (8.1)	1952 (11.7)	2650 (9.5)	2201 (19.9)	3106 (12.5)	
Once or twice a week or less	128,606 (53.7)	84,178 (52.9)	8816 (53.0)	14,979 (53.6)	6553 (59.4)	14,080 (56.9)	
Three times a week or more	87,417 (36.5)	61,568 (38.7)	5819 (35.0)	10,247 (36.7)	2255 (20.4)	7528 (30.4)	
Unknown	662 (0.3)	495 (0.3)	40 (0.2)	44 (0.2)	32 (0.3)	51 (0.2)	
Ethnicity, n (%)							<0.001
White	224,792 (93.9)	149,010 (93.7)	15,960 (96.0)	26,260 (94.1)	9802 (88.8)	23,760 (95.9)	
Asia	5200 (2.2)	3615 (2.3)	192 (1.2)	558 (2.0)	508 (4.6)	327 (1.3)	
Black and Caribbean	4286 (1.8)	2975 (1.9)	146 (0.9)	491 (1.8)	427 (3.9)	247 (1.0)	
Other/unknown	5158 (2.2)	3483 (2.2)	329 (2.0)	611 (2.2)	304 (2.8)	431 (1.7)	
Region, n (%)							<0.001
England	212,190 (88.6)	140,684 (88.4)	15,006 (90.3)	24,840 (89.0)	9744 (88.3)	21,916 (88.5)	
Scotland	17,382 (7.3)	11,914 (7.5)	1022 (6.1)	1786 (6.4)	837 (7.6)	1823 (7.4)	
Wales	9864 (4.1)	6485 (4.1)	599 (3.6)	1294 (4.6)	460 (4.2)	1026 (4.1)	
Socioeconomic status based on Townsend Score, n (%)							<0.001
Interquartile 1	59,168 (24.7)	40,773 (25.6)	3653 (22.0)	6715 (24.1)	1904 (17.2)	6123 (24.7)	
Interquartile 2	58,909 (24.6)	40,010 (25.2)	3918 (23.6)	6477 (23.2)	2333 (21.1)	6171 (24.9)	
Interquartile 3	59,853 (25.0)	39,856 (25.1)	4195 (25.2)	6949 (24.9)	2708 (24.5)	6145 (24.8)	
Interquartile 4	61,506 (25.7)	38,444 (24.2)	4861 (29.2)	7779 (27.9)	4096 (37.1)	6326 (25.5)	

IQR: Interquartile range. * *p*-value expresses the presence of statistically significant differences among the five morbidity patterns identified (Kruskal–Wallis test for continuous variables, Pearson's χ^2 test for categorical). ᵘ Post-menopausal women only.

3.1.5. Pattern 5: Unspecific Morbidities [n = 24,765 (10.3%), 554 Breast Cancer Cases (2.0%)]

The median age at baseline was 59.2 years (IQR: 51.9, 64.0), and the median follow-up time was 7.1 years (IQR: 6.4, 7.8). Women with this pattern were predominantly diagnosed with migraine, diverticular intestine disease, inflammatory bowel disease, rheumatoid disease, and threated dyspepsia.

3.2. Breast Cancer Risk According to the Number of Morbidities and Morbidity Patterns

In both age-adjusted and fully adjusted models, no significant association was found between either the number of morbidities or any morbidity pattern and breast cancer risk, but there was a 12% increased risk associated with self-reported depression (Tables 2 and 3). The results did not vary significantly with age at baseline (p-value interaction = 0.43 and 0.07, for the analyses on the number of morbidities and morbidity patterns, respectively) and adherence to recommendations for breast cancer screening among women aged 50 and older (p-value interaction = 0.44 and 0.84, for the analyses on the number of morbidities and morbidity patterns, respectively), although we found an increased risk among women aged of up to 50 years in the psychiatric morbidities pattern (HR= 1.25; 95%CI: 1.02–1.52) (Figures 3 and 4). The results remained consistent after accounting for competing risks (Table 4), when considering attained age as the timescale in the Cox models (Supplementary, Table S2) and in other sensitivity analyses (see Supplementary, Tables S3 and S4).

Table 2. Association between preexisting single diseases at baseline and breast cancer risk.

Pre-Existing Disease at Baseline	Number of Breast Cancer Cases/Person Years	Age-Adjusted Model HR (95%CI)	Multivariable Model HR (95%CI)
Hypertension			
No	3979/1,287,967	1.00 (Reference)	1.00 (Reference)
Yes	1347/383,417	1.06 (0.99–1.13)	1.03 (0.97–1.11)
Pain condition			
No	4336/1,386,565	1.00 (Reference)	1.00 (Reference)
Yes	990/284,820	1.06 (0.98–1.13)	1.04 (0.97–1.12)
Asthma			
No	4692/1,465,134	1.00 (Reference)	1.00 (Reference)
Yes	634/206,250	0.97 (0.89–1.05)	0.96 (0.88–1.04)
Thyroid disorders			
No	4836/1,513,768	1.00 (Reference)	1.00 (Reference)
Yes	490/157,617	0.94 (0.85–1.03)	0.93 (0.85–1.02)
Treated dyspepsia			
No	4898/1,548,551	1.00 (Reference)	1.00 (Reference)
Yes	428/122,834	1.04 (0.95–1.15)	1.04 (0.94–1.15)
Depression			
No	4927/1,557,562	1.00 (Reference)	1.00 (Reference)
Yes	399/113,821	1.13 (1.02–1.26)	1.12 (1.01–1.24)
Migraine			
No	5099/1,601,276	1.00 (Reference)	1.00 (Reference)
Yes	227/70,109	1.04 (0.91–1.18)	1.05 (0.91–1.19)
Psoriasis			
No	5131/1,612,546	1.00 (Reference)	1.00 (Reference)
Yes	195/58,839	1.06 (0.92–1.22)	1.04 (0.90–1.2)
Diabetes			
No	5138/1,616,001	1.00 (Reference)	1.00 (Reference)
Yes	188/55,384	1.02 (0.88–1.18)	0.99 (0.85–1.15)
Irritable bowel syndrome			
No	5157/1,617,608	1.00 (Reference)	1.00 (Reference)
Yes	169/53,776	0.98 (0.84–1.15)	0.99 (0.85–1.15)
Rheumatoid arthritis			
No	5181/1,624,015	1.00 (Reference)	1.00 (Reference)
Yes	145/473,698	0.92 (0.78–1.09)	0.92 (0.78–1.09)

Table 2. *Cont.*

Pre-Existing Disease at Baseline	Number of Breast Cancer Cases/Person Years	Age-Adjusted Model HR (95%CI)	Multivariable Model HR (95%CI)
Coronary heart disease			
No	5227/1,632,796	1.00 (Reference)	1.00 (Reference)
Yes	99/38,589	0.72 (0.59–0.88)	0.73 (0.60–0.89)
Anxiety			
No	5221/1,637,202	1.00 (Reference)	1.00 (Reference)
Yes	105/34,183	0.97 (0.80–1.18)	0.96 (0.79–1.17)
COPD			
No	5245/1,648,455	1.00 (Reference)	1.00 (Reference)
Yes	81/22,930	1.05 (0.84–1.30)	1.07 (0.86–1.33)
Stroke			
No	5260/1,649,817	1.00 (Reference)	1.00 (Reference)
Yes	66/21,568	0.89 (0.70–1.14)	0.91 (0.71–1.16)
Diverticular disease of intestine			
No	5258/1,650,114	1.00 (Reference)	1.00 (Reference)
Yes	68/21,271	0.92 (0.72–1.17)	0.9 (0.71–1.15)

HR: hazard ratio; CI: confidence interval; the fully adjusted model was adjusted for age at baseline (continuous), age at menarche (continuous), age at menopause (still had periods; had menopause before the age of 45 years; had menopause between the age of 45 and 54; had menopause after the age of 55), menopausal hormone therapy use (never; ever, less than 5-year duration; ever, 5 years and longer; ever, unknown duration), oral contraceptive use (never; ever, less than 10-year duration; ever, at least 10-year duration; ever, unknown duration; unknown status), parity and age at first birth (no live birth; at least one birth before age 30; at least one birth after age 30), BMI (continuous), ethnicity (Asian; Black/Caribbean; White; others/unknown), Townsend score (continuous); level of physical activity (low; moderate; high), alcohol consumption (never; twice a week or less; three times a week or more; unknown status).

Table 3. Associations among number of morbidities, morbidity patterns, and breast cancer risk.

	Study Population (n = 239,436)			Postmenopausal Women Only (n = 175,949)		
Characteristics	**Breast Cancer Cases/Person-Years**	**Age-Adjusted Models HR (95%CI)**	**Fully Adjusted Models HR (95%CI)**	**Breast Cancer Cases/Person Years**	**Age-Adjusted Models HR (95%CI)**	**Fully Adjusted Models HR (95%CI)**
Number of morbidities						
No morbidity	2131/69,8776	1.00 (Reference)	1.00 (Reference)	1451/454,566	1.00 (Reference)	1.00 (Reference)
One morbidity	1736/54,3974	1.01 (0.95–1.08)	1.00 (0.94–1.07)	1361/408,943	1.02 (0.95–1.10)	1 (0.93–1.08)
Multi-morbidities	1459/428,635	1.04 (0.97–1.02)	1.03 (0.96–1.11)	1268/359,844	1.06 (0.98–1.14)	1.02 (0.94–1.1)
Two morbidities	911/266,831	1.05 (0.97–1.14)	1.04 (0.96–1.13)	786/218,780	1.08 (0.99–1.18)	1.04 (0.95–1.14)
3+ morbidities	548/161,804	1.03 (0.93–1.13)	1.01 (0.92–1.12)	482/141,065	1.02 (092–1.14)	0.97 (0.87–1.08)
Morbidity patterns						
No-predominant morbidity	3534/1,110,979	1.00 (Reference)	1.00 (Reference)	2670/798,572	1.00 (Reference)	1.00 (Reference)
Psychiatric morbidities	381/115,476	1.06 (0.95–1.18)	1.04 (0.94–1.16)	264/80,575	1.00 (0.88–1.14)	0.98 (0.86–1.11)
Respiratory/immunological morbidities	611/195,129	0.99 (0.91–1.08)	0.98 (0.9–1.07)	467/137,526	1.02 (0.92–1.12)	1.01 (0.91–1.11)
Cardiovascular/metabolic morbidities	246/75,843	0.94 (0.83–1.07)	0.93 (0.81–1.06)	232/69,252	0.96 (0.84–1.10)	0.91 (0.79–1.05)
Unspecific morbidities	554/173,957	0.98 (0.89–1.07)	0.98 (0.89–1.07)	447/137,429	0.96 (0.87–1.06)	0.95 (0.86–1.05)

HR: hazard ratio; CI: confidence interval; the fully adjusted model was adjusted for age at baseline (continuous), age at menarche (continuous), age at menopause (still had periods; had menopause before the age of 45 years; had menopause between the age of 45 and 54; had menopause after the age of 55; others/unknown), Townsend score (continuous); level of physical activity (low; moderate; high), alcohol consumption (never; once or twice a week or less; three times a week or more; unknown status), menopausal hormone therapy use (never; ever, less than 5-year duration; ever, 5 years and longer; ever, unknown duration), oral contraceptive use (never; ever, less than 10-year duration; ever, at least 10-year duration; ever, unknown duration; unknown status), parity and age at first birth (no live birth; at least one birth before age 30; at least one birth after age 30), BMI (continuous), ethnicity (Asian; Black/Caribbean; White).

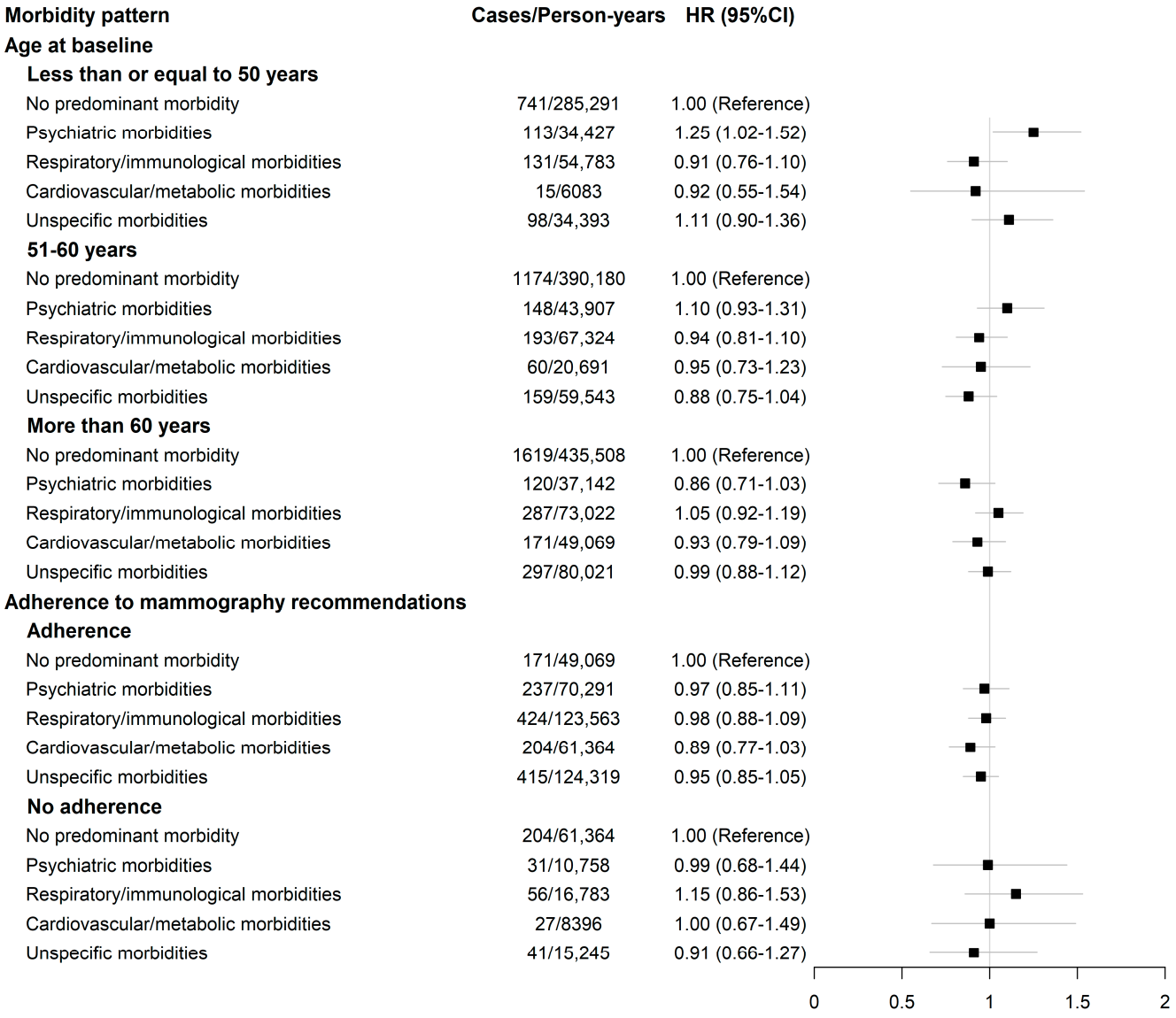

Figure 3. Associations between morbidity clusters and breast cancer risk, according to age-groups and the adherence to breast cancer screening recommendations. HR: hazard ratio; CI: confidence interval. The adherence to mammography included only women older than 50 years. The model was adjusted for age at menarche (continuous), age at menopause (still had periods; had menopause before the age of 45 years; had menopause between the age of 45 and 54; had menopause after the age of 55), menopausal hormone therapy use (never; ever, less than 5-year duration; ever, 5 years and longer; ever, unknown duration), oral contraceptive use (never; ever, less than 10-year duration; ever, at least 10-year duration; ever, unknown duration; unknown status), parity and age at first birth (no live birth; at least one birth before age 30; at least one birth after age 30), BMI (continuous), ethnicity (Asian; Black/Caribbean; White; others/unknown), Townsend score (continuous); level of physical activity (low; moderate; high), alcohol consumption (never; twice a week or less; three times a week or more; unknown status).

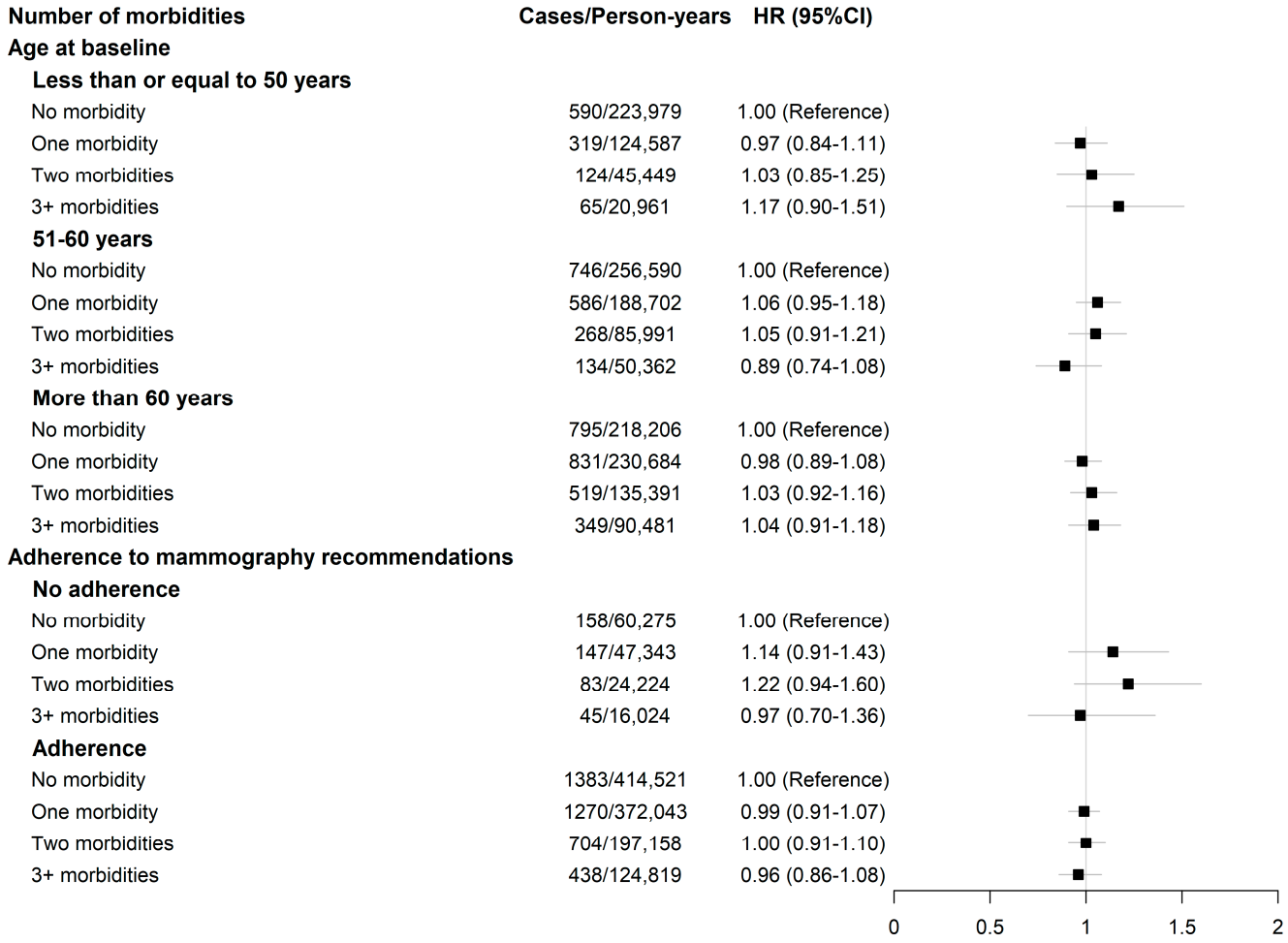

Figure 4. Associations between morbidity clusters and breast cancer risk, according to age groups and the adherence to breast cancer screening recommendations. HR: hazard ratio; CI: confidence interval. The adherence to mammography included only women older than 50 years. The model was adjusted for age at menarche (continuous), age at menopause (still had periods; had menopause before the age of 45 years; had menopause between the age of 45 and 54; had menopause after the age of 55), menopausal hormone therapy use (never; ever, less than 5-year duration; ever, 5 years and longer; ever, unknown duration), oral contraceptive use (never; ever, less than 10-year duration; ever, at least 10-year duration; ever, unknown duration; unknown status), parity and age at first birth (no live birth; at least one birth before age 30; at least one birth after age 30), BMI (continuous), ethnicity (Asian; Black/Caribbean; White; others/unknown), Townsend score (continuous); level of physical activity (low; moderate; high), alcohol consumption (never; twice a week or less; three times a week or more; unknown status).

Table 4. Association between morbidity patterns and breast cancer risk, counting death and first diagnosed non-breast cancer cases as a competing risk.

Event	Morbidity Pattern	Cases/Person-Years	Hazard Ratio (95%CI)
Breast cancer as first diagnosed cancer			
	No-predominant morbidity	3534/1,110,979	1.00 (Reference)
	Psychiatric morbidities	381/115,476	1.04 (0.94–1.16)
	Respiratory/immunological morbidities	611/195,129	0.98 (0.90–1.07)
	Cardiovascular/metabolic morbidities	246/758,423	0.93 (0.81–1.06)
	Unspecific morbidities	554/173,957	0.98 (0.89–1.07)

Table 4. *Cont.*

Event	Morbidity Pattern	Cases/Person-Years	Hazard Ratio (95%CI)
Non-breast cancer as first diagnosed cancer			
	No-predominant morbidity	4964/1,110,979	1.00 (Reference)
	Psychiatric morbidities	485/115,476	0.96 (0.88–1.06)
	Respiratory/immunological morbidities	1041/195,129	1.18 (1.11–1.27)
	Cardiovascular/metabolic morbidities	561/758,423	1.19 (1.09–1.30)
	Unspecific morbidities	862/173,957	1.00 (0.93–1.07)
	Death		
	No-predominant morbidity	645/1,110,979	1.00 (Reference)
	Psychiatric morbidities	126/115,476	1.82 (1.50–2.21)
	Respiratory/immunological morbidities	203/195,129	1.68 (1.44–1.97)
	Cardiovascular/metabolic morbidities	242/758,423	3.06 (2.61–3.58)
	Unspecific morbidities	205/173,957	1.65 (1.41–1.94)

HR: hazard ratio; CI: confidence interval. The model was adjusted for age at baseline (continuous), age at menarche (continuous), age at menopause (still had periods; had menopause before the age of 45 years; had menopause between the age of 45 and 54; had menopause after the age of 55), menopausal hormone therapy use (never; ever, less than 5-year duration; ever, 5 years and longer; ever, unknown duration), oral contraceptive use (never; ever, less than 10-year duration; ever, at least 10-year duration; ever, unknown duration; unknown status), parity and age at first birth (no live birth; at least one birth before age 30; at least one birth after age 30), BMI (continuous), ethnicity (Asian; Black/Caribbean; White; others/unknown), Townsend score (continuous); level of physical activity (low; moderate; high), alcohol consumption (never; twice a week or less; three times a week or more; unknown status).

4. Discussion

Among female participants in the UK Biobank cohort, 58.4% had at least one chronic disease, while 25.8% had two or more simultaneous morbidities. Hypertension was the most prevalent disease (23.1%) at baseline. We found five morbidity patterns: no-predominant morbidity, psychiatric morbidities, respiratory/immunological morbidities, cardiovascular/metabolic morbidities, and unspecific morbidities. There was a 1.12-fold increased risk among women who self-reported depression and a 25% increased risk of breast cancer associated with a psychiatric morbidity pattern compared to that with the no-predominant morbidity pattern, among women younger than 50 only. We did not observe other significant associations between either the number of morbidities or any morbidity pattern and the risk of breast cancer, which did not vary according to adherence to breast cancer screening recommendations, socioeconomic status, BMI, physical activity level, or menopausal status.

Despite heterogeneous findings in previous studies on morbidities across different populations and settings, several morbidity patterns often emerge in the literature, which were also observed in our study [22,23,46,47]. The pattern of cardiovascular/metabolic morbidities has been extensively described previously, as there are established etiologic associations among diabetes, stroke, heart failure, and heart disease, with an interlinked pathophysiology and common risk factors, such as obesity, physical inactivity, and smoking [48]. For the pattern of psychiatric morbidities, although little is known about the pathogenesis of depression and anxiety, these two frequent mental illnesses share a largely overlapping set of risk factors with breast cancer, including female sex, genetic predisposition, family history, and environmental influence (childhood adversity, low socioeconomic status) [49,50]. Depression and anxiety are also common coexisting conditions among patients with chronic comorbidities, including cancer [51,52]. Consistent with our findings, a recent nationwide population-based study has shown that mental disorders were associated with a subsequent higher risk of cancer, although the causal link remains a topic of debate [52]. The diseases included in the respiratory pattern, such as chronic obstructive pulmonary disease and asthma, involve a prolonged inflammatory response and the sharing of risk factors, such as smoking, an unhealthy diet, physical inactivity, and high alcohol consumption. However, combinations among asthma, COPD, and psoriasis and eczema are less common. Thus, these patterns found in our clustering analysis not only represent a clinically relevant morbidity status in women in the UK Biobank cohort but

also reflect distinct profiles of (known or unknown), shared genetics, and behavioral and environmental risk factors, both of which might increase the risk of developing cancer.

Indeed, to our knowledge, our study is the first to investigate the association between morbidity patterns and breast cancer risk. We found no association between either the number of morbidities or morbidity patterns and breast cancer risk, regardless of the women's age at baseline, and socioeconomic characteristics, apart from an increased risk among women aged less than 50 having multiple psychiatric diseases. Analyses stratified based on adherence to breast cancer screening recommendations did not modify our main results, suggesting that surveillance bias is not an important modifying factor in the association between breast cancer risk and morbidities. Previously, there was only a case-control study reporting results on the association between multi-morbidity and breast cancer risk. The findings indicated that an increasing number of morbidities measured with the Charlson comorbidity index (CCI) was associated with an increasing breast cancer risk (46,324 cases) after a 10-year follow-up of women aged 45–85, but no association was found for individual morbidities [53]. However, they were not able to control for confounding factors other than age at baseline and to account for surveillance bias. For comparison purposes, we applied the same methods in an additional analysis by using the CCI (Tables S5 and S6), and we did not find a significant association between the Charlson morbidity number and breast cancer risk after adjusting for well-known risk factors.

There are several hypotheses to explain the null results. First, women with morbidity could experience other serious long-term outcomes before a breast cancer diagnosis. Indeed, when accounting for death and malignancies other than breast cancer as competing risks, we found that compared to that in women with no predominant morbidity, women with other patterns were more likely to die and/or to be diagnosed with other cancers. This is particularly pronounced among women with cardiovascular/metabolic and respiratory/immunological morbidities. Second, given the different biologic characteristics of divergent breast cancer subtypes [54] and the complexity of multi-morbidity mechanisms and risk factors, the risk estimations could vary across individual associations, and the possible opposing effects could drive the combined estimates toward null. For instance, BMI, a common risk factor of various morbidities, is strongly associated with hormone receptor-positive tumors, but not a triple-negative or core basal phenotype [55]. A high BMI is a risk factor of postmenopausal breast cancer, but a protective factor of premenopausal breast cancer. Type 2 diabetes is an independent risk factor of breast cancer risk in postmenopausal women, but no increased risk was observed for premenopausal women [8]. In our study, when restricting analyses to postmenopausal women only, the null associations remained consistent. Previous large prospective cohorts reported that low socioeconomic positions, a contributing factor of psychiatric morbidities, were found to be associated with a lower risk of ER+ breast cancer but a higher risk of the ER- subtype [56,57]. Meanwhile, adverse life events, such as childhood abuse and divorce, were associated with a higher risk of ER+, but not ER-, breast cancer [57,58]. Third, our null results could also suggest that the underlying common biological pathways among morbidities in an individual pattern and their shared risk factors were not a key factor explaining breast cancer risk after accounting for established breast cancer risk factors.

Strength and limitations: The UK Biobank cohort is a large population-based cohort with a high follow-up rate and important number of breast cancer cases. The cohort includes a wide range of information on personal medical history, reproductive factors, lifestyle factors, socioeconomic status, and family medical history, with low levels of missing data. Nevertheless, there are several limitations that must be acknowledged. Assuming that the prevalence of having at least one morbidity in women in the UK Biobank cohort is slightly lower than what has been found (42.2%, 33.8%) in previous studies of Barnett and Gondek, respectively (since these studies have included data of both women and men in the analyses, which could lead to a potential underestimation of the morbidity prevalence), this suggests the occurrence of "healthy" volunteer bias (i.e., UK Biobank participants are more likely to be in good health conditions than the general population) [59,60]. However, since

our study focuses on investigating breast cancer risk in relation to morbidity and not on estimating disease prevalence rates and many people with a wide range of morbidities and risk factors are included in the cohort, the risk estimations are unlikely to be biased [59,60]. We used self-reported health condition data, which were not externally validated, and the UK Biobank did not include information on morbidity severity. There was no longitudinal updated morbidity status and thus no possibility to study changes in morbidity patterns during follow-up. We also missed details on the breast cancer stage, grade, and receptor status. This did not allow us to further study the surveillance biases related to the disease stage and grade or to investigate potential pathways related to tumor receptor status.

5. Conclusions

Female participants in the UK Biobank cohort can be classified into five morbidity patterns: no-predominant morbidity, psychiatric morbidities, respiratory/immunological morbidities, cardiovascular/metabolic morbidities, and unspecific morbidities. We found a significant increased risk among women aged younger than 50 with a psychiatric diseases pattern, but there was no other significant association among the number of morbidities, the morbidity patterns, and the risk of breast cancer in this population. Our findings suggest that multimorbidity is not a decisive factor to help identify patients at increased risk of breast cancer.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers15041165/s1, Figure S1: Individuals' point cloud (a) on the first two and (b) on the first and third axes. Figure S2: Morbidity categories' point cloud (a) on the first two and (b) on the first and third axes. Figure S3: suggested number of MCA dimensions: (a) Scree plot; (b) Horn's parallel analysis. Figure S4: Suggested number of morbidity clusters: (a) Gap-statistic method; (b) elbow method. Figure S5: Dendrogram of five morbidity clusters solution. Table S1: Suggested number of morbidity patterns based on the Davies-Bouldin index. Table S2: Cox models using attained age as timescale. Table S3: Cox models using invasive breast cancer only as outcome. Table S4: Others modified effects. Table S5: Association between breast cancer risk and the preexisting Charlson single pre-existing diseases at baseline. Table S6: Association between Charlson comorbidity index at baseline and breast cancer risk.

Author Contributions: Conceptualization, T.-V.-T.T.; methodology, T.-V.-T.T. and A.M.S.H.; statistical analysis, T.-V.-T.T. and A.M.S.H.; supervision, T.-V.-T.T. and N.M.Y.J.; original drafting of the paper, A.M.S.H., T.-V.-T.T.; writing-review-and editing, A.M.S.H., R.S.A., F.d.V., M.-C.B.-R., N.M.Y.J. and T.-V.-T.T.; interpretation of the results, A.M.S.H., R.S.A., F.d.V., M.-C.B.-R., N.M.Y.J. and T.-V.-T.T. All authors have read and agreed to the published version of the manuscript.

Funding: Afi HENYOH received an internship stipend from INSERM. Thi-Van-Trinh TRAN received a doctoral grant from the Paris Sud-Paris Saclay University.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was performed under generic ethical approval obtained by the UK Biobank from the National Health Service National Research Ethics Service (approval letter ref 16/NW/0274, 13 May 2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: This work has been conducted using the UK Biobank Resource under Application Number 35032. Bona-fide researchers can apply to use the UK Biobank dataset by registering and applying at http://www.ukbiobank.ac.uk/register-apply accessed on 1 March 2022.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Baseline long-term health condition groupings.

Morbidity ^	Conditions Included
1. Painful conditions *	Back pain Joint pain Headaches (not migraine) Sciatica Plantar fasciitis Carpal tunnel syndrome Fibromyalgia Arthritis Shingles Disc problem Prolapsed disc/slipped disc Spine arthritis/spondylitis Ankylosing spondylitis Back problem Osteoarthritis Gout Cervical spondylosis Trigeminal neuralgia Disc degeneration Trapped nerve/compressed nerve
2. Hypertension	Hypertension Essential hypertension
3. Depression *	Depression Postnatal depression
4. Asthma	Asthma
5. Coronary heart disease	Heart attack/MI Angina
6. Treated dyspepsia	Gastro-esophageal reflux (GORD)/gastric reflux Esophagitis/Barrett's esophagus Gastric stomach ulcers Gastric erosions/gastritis Duodenal ulcer Dyspepsia/indigestion Hiatus hernia *Helicobacter pylori*
7. Diabetes	Diabetic nephropathy Diabetic neuropathy/ulcers Diabetes Type 1 diabetes Type 2 diabetes Diabetic eye disease
8. Thyroid disorders	Thyroid problem (not cancer) Hyperthyroidism/thyrotoxicosis Hypothyroidism/myxedema Graves' disease Thyroid goitre Thyroiditis

Table A1. *Cont.*

Morbidity ^	Conditions Included
9. Rheumatoid arthritis, other inflammatory polyarthropathies, systemic connective tissue disorders and systemic autoimmune disorders	Myositis/myopathy Systemic lupus erythematosus Connective tissue disorder Sjogren's syndrome/sicca syndrome Dermatopolymyositis Scleroderma/systemic sclerosis Rheumatoid arthritis Psoriatic arthropathy Dermatomyositis Polymyositis Polymyalgia rheumatica Malabsorption/celiac disease
10. Chronic obstructive pulmonary disease (COPD)	COPD/chronic obstructive airways disease Emphysema/chronic bronchitis Emphysema
11. Anxiety, other neurotic, stress-related, and somatoform disorders *	Anxiety/panic attacks Nervous breakdown Post-traumatic stress disorder Obsessive compulsive disorder Stress Insomnia Psychological/psychiatric problem
12. Irritable bowel syndrome	Irritable bowel syndrome
13. Alcohol problems *	Alcohol dependency Alcoholic liver disease/alcoholic cirrhosis
14. Other psychoactive substance abuse *	Opioid dependency Other substance abuse/dependency
15. Treated constipation	Constipation
16. Stroke and transient ischemic attack (TIA)	Stroke TIA Subarachnoid hemorrhage Brain hemorrhage Ischemic stroke
17. Chronic kidney disease	Polycystic kidney Diabetic nephropathy Renal/kidney failure Renal failure requiring dialysis Renal failure not requiring dialysis Kidney nephropathy Immunoglobulin A (IgA) nephropathy
18. Diverticular disease of intestine	Diverticular disease/diverticulitis
19. Atrial fibrillation	Atrial fibrillation
20. Peripheral vascular disease	Peripheral vascular disease Leg claudication/intermittent claudication
21. Heart failure	Cardiomyopathy Hypertrophic cardiomyopathy Heart failure/pulmonary edema
22. Prostate disorders	Prostate problem (not cancer) Enlarged prostate Benign prostatic hypertrophy
23. Glaucoma	Glaucoma
24. Epilepsy	Epilepsy

Table A1. *Cont.*

Morbidity ^	**Conditions Included**
25. Dementia	Dementia/Alzheimer/cognitive impairment
26. Schizophrenia (and related non-organic psychosis) and bipolar disorder *	Schizophrenia Mania/bipolar disorder/manic depression
27. Psoriasis or eczema	Eczema/dermatitis Psoriasis
28. Inflammatory bowel disease	Inflammatory bowel disease Crohn's disease Ulcerative colitis
29. Migraine	Migraine
30. Chronic sinusitis	Chronic sinusitis
31. Anorexia or bulimia *	Anorexia, bulimia/other eating disorder
32. Bronchiectasis	Bronchiectasis
33. Parkinson's disease	Parkinson's disease
34. Multiple sclerosis	Multiple sclerosis
35. Viral Hepatitis	Infective/viral hepatitis Hepatitis B Hepatitis C Hepatitis D Hepatitis E
36. Chronic liver disease	Esophageal varices Non infective hepatitis Liver failure/cirrhosis Primary biliary cirrhosis
37. Osteoporosis ~	Osteoporosis
38. Chronic fatigue syndrome ~	Chronic fatigue syndrome
39. Endometriosis ~	Endometriosis
40. Meniere disease ~	Meniere disease
41. Pernicious Anemia ~	Pernicious anemia
42. Polycystic ovaries ~	Polycystic ovaries
43. Cancer	Lifetime diagnosis

^ Self-report lifetime diagnosis by doctor recorded by nurse-led interview (UK Biobank data field 20002), except cancer diagnosis that was reported by touch-screen questionnaire (UK Biobank data field 2453). The list of disease groupings was based on Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B: Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012, 380(9836):37–43 [19]. ~ Plus other conditions considered long-term, requiring medication, and that had a prevalence of $\geq$0.1% in the whole UK Biobank cohort. * Painful and psychiatric conditions were not included in the morbidity count for this study; this resulted in a total of 36 morbidities included.

Table A2. Definition of confounding factors and effect modifiers.

Risk Factors	Coding	Information Source	Testing for Confounding Effect	Testing for Modification Effect
Socio-demographic and economic characteristics				
Age at baseline	Continuous	SR-Q	Yes	Yes
Occupation	Administrative and Secretarial Occupations Associate Professional and Technical Occupations Elementary Occupations Managers and Senior Officials Personal Service Occupations Process, Plant, and Machine Operatives Professional Occupations Sales and Customer Service Occupations Skilled Trades Occupations Unknown	SR-Q	Yes	No
Race	Asian Black and Caribbean White Other/Unknown	SR-Q	Yes	No
Townsend score	Continuous	UK data service	Yes	Yes
Hormone-related factors				
Age at menarche	Continuous	SR-Q	Yes	No
Age at menopause	Still had periods Had menopause before the age of 45 years Had menopause between the age of 45 and 54 Had menopause after the age of 55	SR-Q	Yes	Yes
Menopausal hormone therapy use	Never Ever, less than 5-year duration Ever, 5 years and longer Ever, unknown duration	Reporting menopause (periods stopped) (SR-Q)OR Reporting use of menopausal hormone therapy (SR-Q)OR Undergoing a bilateral oophorectomy (SR-I)OR $\geq$51 years of age at baseline	Yes	No
Oral contraception use	Never Ever, less than 10-year duration; Ever, at least 10-year duration; Ever, unknown duration; Unknown status	SR-Q	Yes	No
Parity and age at first birth	No live birth At least one birth before age 30 At least one birth after age 30	SR-Q	Yes	No

Table A2. *Cont.*

Risk Factors	Coding	Information Source	Testing for Confounding Effect	Testing for Modification Effect
Health and health care-related factors				
BMI	Continuous	PM	Yes	Yes
Level of physical activity	Low Moderate High	SR-Q	Yes	Yes
Alcohol consumption	Never Twice a week or less Three times a week or more Unknown status	SR-Q	Yes	No
Adherence to mammography guidelines	Never Ever, last use since more than 3 years ago Ever, in the last 3 years Ever, unknown time of last use	SR-Q	Yes	Yes

SR-Q: self-reported data from questionnaire, SR-I: self-reported data from trained nurses lead interviews, lead BM: body mass index, PM: physical measurement.

Appendix B

Appendix B.1 Multiple Correspondence Analysis (MCA)

MCA is a data analysis technique used to detect and represent underlying structures in sets of nominal categorical data. It allows identifying groups with similar characteristics and shows, in a multidimensional space, relationships between dichotomous or categorical variables (in our case, morbidities) that would be difficult to observe in a contingency table [34,35]. So far, MCA allows individuals to be directly represented as points (coordinates) in a geometric space through the transformation of original binary data to continuous ones. We performed MCA based on the indicator matrix, also called a complete disjunctive table, which is an individual's × variables matrix, where the rows represent individual, and the columns are dummy variables representing categories of morbidity variables.

Appendix B.2 Cluster Analysis

- Agglomerative hierarchical clustering (AHC)

AHC is a commonly used method for cluster analysis in big data research and data mining aiming to establish a hierarchy of clusters [38,39]. As such, HCA attempts to group, inside a heterogeneous population, subjects with similar features into clusters based on similarity or dissimilarity measures. Initially, each observation belongs to one of N disjoint single patterns. The algorithm then sequentially joined the two closest, in terms of the Euclidean distance, until after $(N - 1)$ steps, all observations belong to a single pattern of size N [40].

- Ward's method for cluster analysis

This approach is based on a classical sum-square criterion and produces clusters that minimize the decrease in between-cluster inertia, therefore minimizing the within-cluster inertia at each merging step [61]. The hierarchical grouping process can be graphically summarized by a tree-like graphical representation called a dendrogram. Similar objects are linked, and their position in the diagram is determined by the level of similarity/dissimilarity between the objects [62]. Thus, clusters were identified by taking into account the similarity distances between the morbidities among the study population, which subsequently allowed the patterns to be formed.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [CrossRef]
2. Tamimi, R.M.; Spiegelman, D.; Smith-Warner, S.A.; Wang, M.; Pazaris, M.; Willett, W.C.; Eliassen, A.H.; Hunter, D.J. Population Attributable Risk of Modifiable and Nonmodifiable Breast Cancer Risk Factors in Postmenopausal Breast Cancer. *Am. J. Epidemiol.* **2016**, *184*, 884–893. [CrossRef]
3. van Gemert, W.A.; Lanting, C.I.; Goldbohm, R.A.; van den Brandt, P.A.; Grooters, H.G.; Kampman, E.; Kiemeney, L.a.L.M.; van Leeuwen, F.E.; Monninkhof, E.M.; de Vries, E.; et al. The proportion of postmenopausal breast cancer cases in the Netherlands attributable to lifestyle-related risk factors. *Breast Cancer Res. Treat.* **2015**, *152*, 155–162. [CrossRef]
4. Maas, P.; Barrdahl, M.; Joshi, A.D.; Auer, P.L.; Gaudet, M.M.; Milne, R.L.; Schumacher, F.R.; Anderson, W.F.; Check, D.; Chattopadhyay, S.; et al. Breast Cancer Risk From Modifiable and Nonmodifiable Risk Factors Among White Women in the United States. *JAMA Oncol.* **2016**, *2*, 1295–1302. [CrossRef]
5. Heer, E.; Harper, A.; Escandor, N.; Sung, H.; McCormack, V.; Fidler-Benaoudia, M.M. Global burden and trends in premenopausal and postmenopausal breast cancer: A population-based study. *Lancet Glob. Health* **2020**, *8*, e1027–e1037. [CrossRef]
6. Xu, X.; Mishra, G.D.; Jones, M. Evidence on multimorbidity from definition to intervention: An overview of systematic reviews. *Ageing Res. Rev.* **2017**, *37*, 53–68. [CrossRef] [PubMed]
7. Cassell, A.; Edwards, D.; Harshfield, A.; Rhodes, K.; Brimicombe, J.; Payne, R.; Griffin, S. The epidemiology of multimorbidity in primary care: A retrospective cohort study. *Br. J. Gen. Pract.* **2018**, *68*, e245–e251. [CrossRef] [PubMed]
8. Boyle, P.; Boniol, M.; Koechlin, A.; Robertson, C.; Valentini, F.; Coppens, K.; Fairley, L.-L.; Boniol, M.; Zheng, T.; Zhang, Y.; et al. Diabetes and breast cancer risk: A meta-analysis. *Br. J. Cancer* **2012**, *107*, 1608–1617. [CrossRef] [PubMed]
9. Tran, T.-V.-T.; Kitahara, C.M.; de Vathaire, F.; Boutron-Ruault, M.-C.; Journy, N. Thyroid dysfunction and cancer incidence: A systematic review and meta-analysis. *Endocr. Relat. Cancer* **2020**, *27*, 245–259. [CrossRef]

10. Tsai, M.-S.; Chen, H.-P.; Hung, C.-M.; Lee, P.-H.; Lin, C.-L.; Kao, C.-H. Hospitalization for Inflammatory Bowel Disease is Associated with Increased Risk of Breast Cancer: A Nationwide Cohort Study of an Asian Population. *Ann. Surg. Oncol.* **2015**, *22*, 1996–2002. [CrossRef]
11. Schairer, C.; Pfeiffer, R.M.; Gadalla, S.M. Autoimmune diseases and breast cancer risk by tumor hormone-receptor status among elderly women. *Int. J. Cancer* **2018**, *142*, 1202–1208. [CrossRef]
12. Han, H.; Guo, W.; Shi, W.; Yu, Y.; Zhang, Y.; Ye, X.; He, J. Hypertension and breast cancer risk: A systematic review and meta-analysis. *Sci. Rep.* **2017**, *7*, 44877. [CrossRef] [PubMed]
13. Krashin, E.; Piekiełko-Witkowska, A.; Ellis, M.; Ashur-Fabian, O. Thyroid Hormones and Cancer: A Comprehensive Review of Preclinical and Clinical Studies. *Front. Endocrinol.* **2019**, *10*, 59. [CrossRef] [PubMed]
14. Hall, L.C.; Salazar, E.P.; Kane, S.R.; Liu, N. Effects of thyroid hormones on human breast cancer cell proliferation. *J. Steroid Biochem. Mol. Biol.* **2008**, *109*, 57–66. [CrossRef] [PubMed]
15. Mantovani, A.; Allavena, P.; Sica, A.; Balkwill, F. Cancer-related inflammation. *Nature* **2008**, *454*, 436–444. [CrossRef]
16. Xue, F.; Michels, K.B. Diabetes, metabolic syndrome, and breast cancer: A review of the current evidence. *Am. J. Clin. Nutr.* **2007**, *86*, s823–s835. [CrossRef] [PubMed]
17. Mehta, L.S.; Watson, K.E.; Barac, A.; Beckie, T.M.; Bittner, V.; Cruz-Flores, S.; Dent, S.; Kondapalli, L.; Ky, B.; Okwuosa, T.; et al. Cardiovascular Disease and Breast Cancer: Where These Entities Intersect: A Scientific Statement From the American Heart Association. *Circulation* **2018**, *137*, e30–e66. [CrossRef] [PubMed]
18. Algra, A.M.; Rothwell, P.M. Effects of regular aspirin on long-term cancer incidence and metastasis: A systematic comparison of evidence from observational studies versus randomised trials. *Lancet Oncol.* **2012**, *13*, 518–527. [CrossRef]
19. Barnett, K.; Mercer, S.W.; Norbury, M.; Watt, G.; Wyke, S.; Guthrie, B. Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study. *Lancet* **2012**, *380*, 37–43. [CrossRef] [PubMed]
20. Prados-Torres, A.; Poblador-Plou, B.; Gimeno-Miguel, A.; Calderón-Larrañaga, A.; Poncel-Falcó, A.; Gimeno-Feliú, L.A.; González-Rubio, F.; Laguna-Berna, C.; Marta-Moreno, J.; Clerencia-Sierra, M.; et al. Cohort Profile: The Epidemiology of Chronic Diseases and Multimorbidity. The EpiChron Cohort Study. *Int. J. Epidemiol.* **2018**, *47*, 382–384f. [CrossRef]
21. Britt, H.C.; Harrison, C.M.; Miller, G.C.; Knox, S.A. Prevalence and patterns of multimorbidity in Australia. *Med. J. Aust.* **2008**, *189*, 72–77. [CrossRef] [PubMed]
22. Prados-Torres, A.; Calderón-Larrañaga, A.; Hancco-Saavedra, J.; Poblador-Plou, B.; van den Akker, M. Multimorbidity patterns: A systematic review. *J. Clin. Epidemiol.* **2014**, *67*, 254–266. [CrossRef]
23. Vetrano, D.L.; Roso-Llorach, A.; Fernández, S.; Guisado-Clavero, M.; Violán, C.; Onder, G.; Fratiglioni, L.; Calderón-Larrañaga, A.; Marengoni, A. Twelve-year clinical trajectories of multimorbidity in a population of older adults. *Nat. Commun.* **2020**, *11*, 3223. [CrossRef] [PubMed]
24. The Emerging Risk Factors Collaboration Association of Cardiometabolic Multimorbidity With Mortality. *JAMA* **2015**, *314*, 52–60. [CrossRef] [PubMed]
25. Yasmeen, S.; Hubbard, R.A.; Romano, P.S.; Zhu, W.; Geller, B.M.; Onega, T.; Yankaskas, B.C.; Miglioretti, D.L.; Kerlikowske, K. Risk of Advanced-Stage Breast Cancer Among Older Women with Comorbidities. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **2012**, *21*, 1510–1519. [CrossRef]
26. Jensen, L.F.; Pedersen, A.F.; Andersen, B.; Vestergaard, M.; Vedsted, P. Non-participation in breast cancer screening for women with chronic diseases and multimorbidity: A population-based cohort study. *BMC Cancer* **2015**, *15*, 798. [CrossRef] [PubMed]
27. Warner, D.F.; Koroukian, S.M.; Schiltz, N.K.; Smyth, K.A.; Cooper, G.S.; Owusu, C.; Stange, K.C.; Berger, N.A. Complex Multimorbidity and Breast Cancer Screening Among Midlife and Older Women: The Role of Perceived Need. *Gerontologist* **2019**, *59*, S77–S87. [CrossRef]
28. Sudlow, C.; Gallacher, J.; Allen, N.; Beral, V.; Burton, P.; Danesh, J.; Downey, P.; Elliott, P.; Green, J.; Landray, M.; et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **2015**, *12*, e1001779. [CrossRef]
29. Nicholl, B.I.; Mackay, D.; Cullen, B.; Martin, D.J.; Ul-Haq, Z.; Mair, F.S.; Evans, J.; McIntosh, A.M.; Gallagher, J.; Roberts, B.; et al. Chronic multisite pain in major depression and bipolar disorder: Cross-sectional study of 149,611 participants in UK Biobank. *BMC Psychiatry* **2014**, *14*, 350. [CrossRef]
30. Anderson, K.N.; Schwab, R.B.; Martinez, M.E. Reproductive Risk Factors and Breast Cancer Subtypes: A Review of the Literature. *Breast Cancer Res. Treat.* **2014**, *144*, 1–10. [CrossRef]
31. McPherson, K.; Steel, C.M.; Dixon, J.M. ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ* **2000**, *321*, 624–628. [CrossRef] [PubMed]
32. Wang, Y.-X.; Arvizu, M.; Rich-Edwards, J.W.; Stuart, J.J.; Manson, J.E.; Missmer, S.A.; Pan, A.; Chavarro, J.E. Menstrual cycle regularity and length across the reproductive lifespan and risk of premature mortality: Prospective cohort study. *BMJ* **2020**, *371*, m3464. [CrossRef] [PubMed]
33. Violán, C.; Roso-Llorach, A.; Foguet-Boreu, Q.; Guisado-Clavero, M.; Pons-Vigués, M.; Pujol-Ribera, E.; Valderas, J.M. Multimorbidity patterns with K-means nonhierarchical cluster analysis. *BMC Fam. Pract.* **2018**, *19*, 108. [CrossRef] [PubMed]
34. Sourial, N.; Wolfson, C.; Zhu, B.; Quail, J.; Fletcher, J.; Karunananthan, S.; Bandeen-Roche, K.; Béland, F.; Bergman, H. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *J. Clin. Epidemiol.* **2010**, *63*, 638–646. [CrossRef]

35. García-Gil, M.; Blanch, J.; Comas-Cufí, M.; Daunis-i-Estadella, J.; Bolíbar, B.; Martí, R.; Ponjoan, A.; Alves-Cabratosa, L.; Ramos, R. Patterns of statin use and cholesterol goal attainment in a high-risk cardiovascular population: A retrospective study of primary care electronic medical records. *J. Clin. Lipidol.* **2016**, *10*, 134–142. [CrossRef] [PubMed]
36. Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. PsycNET. Available online: https://psycnet.apa.org/doiLanding?doi=10.1037%2Fa0023353 (accessed on 21 July 2021).
37. Xu, R.; Wunsch, D. Survey of Clustering Algorithms. *IEEE Trans. Neural Netw.* **2005**, *16*, 645–678. [CrossRef]
38. Petushkova, N.A.; Pyatnitskiy, M.A.; Rudenko, V.A.; Larina, O.V.; Trifonova, O.P.; Kisrieva, J.S.; Samenkova, N.F.; Kuznetsova, G.P.; Karuzina, I.I.; Lisitsa, A.V. Applying of Hierarchical Clustering to Analysis of Protein Patterns in the Human Cancer-Associated Liver. *PLoS ONE* **2014**, *9*, e103950. [CrossRef]
39. Muntaner, C.; Chung, H.; Benach, J.; Ng, E. Hierarchical cluster analysis of labour market regulations and population health: A taxonomy of low- and middle-income countries. *BMC Public Health* **2012**, *12*, 286. [CrossRef] [PubMed]
40. Kimes, P.K.; Liu, Y.; Hayes, D.N.; Marron, J.S. Statistical Significance for Hierarchical Clustering. *Biometrics* **2017**, *73*, 811–821. [CrossRef]
41. Charrad, M.; Ghazzali, N.; Boiteau, V.; Niknafs, A. La Librairie NbClust pour L'estimation du Nombre Optimal de Classes dans un Jeu de Données. In Proceedings of the XXIème Rencontre de la Société Francophone de Classification, Rabat, Morocco, 10–12 September 2014.
42. Estimating the Number of Clusters in a Data Set via the Gap Statistic | Request PDF. Available online: https://www.researchgate.net/publication/4772044_Estimating_the_Number_of_Clusters_in_a_Data_Set_Via_the_Gap_Statistic (accessed on 24 July 2021).
43. Schäfer, I.; Kaduszkiewicz, H.; Wagner, H.-O.; Schön, G.; Scherer, M.; van den Bussche, H. Reducing complexity: A visualisation of multimorbidity by combining disease clusters and triads. *BMC Public Health* **2014**, *14*, 1285. [CrossRef]
44. Putter, H.; Fiocco, M.; Geskus, R.B. Tutorial in biostatistics: Competing risks and multi-state models. *Stat. Med.* **2007**, *26*, 2389–2430. [CrossRef] [PubMed]
45. Higgs, N.T. Practical and Innovative Uses of Correspondence Analysis. *Statistician* **1991**, *40*, 183. [CrossRef]
46. Schäfer, I.; von Leitner, E.-C.; Schön, G.; Koller, D.; Hansen, H.; Kolonko, T.; Kaduszkiewicz, H.; Wegscheider, K.; Glaeske, G.; van den Bussche, H. Multimorbidity Patterns in the Elderly: A New Approach of Disease Clustering Identifies Complex Interrelations between Chronic Conditions. *PLoS ONE* **2010**, *5*, e15941. [CrossRef] [PubMed]
47. Prados-Torres, A.; Poblador-Plou, B.; Calderón-Larrañaga, A.; Gimeno-Feliu, L.A.; González-Rubio, F.; Poncel-Falcó, A.; Sicras-Mainar, A.; Alcalá-Nalvaiz, J.T. Multimorbidity Patterns in Primary Care: Interactions among Chronic Diseases Using Factor Analysis. *PLoS ONE* **2012**, *7*, e32190. [CrossRef]
48. Leon, B.M.; Maddox, T.M. Diabetes and cardiovascular disease: Epidemiology, biological mechanisms, treatment recommendations and future research. *World J. Diabetes* **2015**, *6*, 1246–1258. [CrossRef]
49. Nemeroff, C.B. The State of Our Understanding of the Pathophysiology and Optimal Treatment of Depression: Glass Half Full or Half Empty? *Am. J. Psychiatry* **2020**, *177*, 671–685. [CrossRef] [PubMed]
50. Generalized Anxiety Disorder | NEJM. Available online: https://www.nejm.org/doi/full/10.1056/NEJMcp1502514 (accessed on 19 November 2022).
51. Lotfaliany, M.; Bowe, S.J.; Kowal, P.; Orellana, L.; Berk, M.; Mohebbi, M. Depression and chronic diseases: Co-occurrence and communality of risk factors. *J. Affect. Disord.* **2018**, *241*, 461–468. [CrossRef] [PubMed]
52. Momen, N.C.; Plana-Ripoll, O.; Agerbo, E.; Benros, M.E.; Børglum, A.D.; Christensen, M.K.; Dalsgaard, S.; Degenhardt, L.; de Jonge, P.; Debost, J.-C.P.G.; et al. Association between Mental Disorders and Subsequent Medical Conditions. *N. Engl. J. Med.* **2020**, *382*, 1721–1731. [CrossRef]
53. Ording, A.G.; Garne, J.P.; Nyström, P.M.W.; Cronin-Fenton, D.; Tarp, M.; Sørensen, H.T.; Lash, T.L. Hospital Recorded Morbidity and Breast Cancer Incidence: A Nationwide Population-Based Case-Control Study. *PLoS ONE* **2012**, *7*, e47329. [CrossRef]
54. Gaudet, M.M.; Gierach, G.L.; Carter, B.D.; Luo, J.; Milne, R.L.; Weiderpass, E.; Giles, G.G.; Tamimi, R.M.; Eliassen, A.H.; Rosner, B.; et al. Pooled Analysis of Nine Cohorts Reveals Breast Cancer Risk Factors by Tumor Molecular Subtype. *Cancer Res.* **2018**, *78*, 6011–6021. [CrossRef]
55. Yang, X.R.; Chang-Claude, J.; Goode, E.L.; Couch, F.J.; Nevanlinna, H.; Milne, R.L.; Gaudet, M.; Schmidt, M.K.; Broeks, A.; Cox, A.; et al. Associations of Breast Cancer Risk Factors With Tumor Subtypes: A Pooled Analysis From the Breast Cancer Association Consortium Studies. *JNCI J. Natl. Cancer Inst.* **2011**, *103*, 250–263. [CrossRef] [PubMed]
56. Andaya, A.A.; Enewold, L.; Horner, M.-J.; Jatoi, I.; Shriver, C.D.; Zhu, K. Socioeconomic disparities and breast cancer hormone receptor status. *Cancer Causes Control CCC* **2012**, *23*, 951–958. [CrossRef] [PubMed]
57. Barber, L.E.; Zirpoli, G.R.; Cozier, Y.C.; Rosenberg, L.; Petrick, J.L.; Bertrand, K.A.; Palmer, J.R. Neighborhood disadvantage and individual-level life stressors in relation to breast cancer incidence in US Black women. *Breast Cancer Res. BCR* **2021**, *23*, 108. [CrossRef] [PubMed]
58. Schoemaker, M.J.; Jones, M.E.; Wright, L.B.; Griffin, J.; McFadden, E.; Ashworth, A.; Swerdlow, A.J. Psychological stress, adverse life events and breast cancer incidence: A cohort investigation in 106,000 women in the United Kingdom. *Breast Cancer Res. BCR* **2016**, *18*, 72. [CrossRef]
59. Fry, A.; Littlejohns, T.J.; Sudlow, C.; Doherty, N.; Adamska, L.; Sprosen, T.; Collins, R.; Allen, N.E. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **2017**, *186*, 1026–1034. [CrossRef]

60. Batty, G.D.; Gale, C.R.; Kivimäki, M.; Deary, I.J.; Bell, S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: Prospective cohort study and individual participant meta-analysis. *BMJ* **2020**, *368*, m131. [CrossRef]
61. Murtagh, F.; Legendre, P. Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm. *arXiv* **2011**, arXiv:1111.6285.
62. Zhang, Z.; Murtagh, F.; Van Poucke, S.; Lin, S.; Lan, P. Hierarchical cluster analysis in clinical research with heterogeneous study population: Highlighting its visualization with R. *Ann. Transl. Med.* **2017**, *5*, 75. [CrossRef]