Combining Charlson and Elixhauser scores with varying lookback predicated mortality better than using individual scores

Emma Pritchard*[1,2], Nicola Fawcett[1,3], T. Phuong Quan[1,2,4], Derrick Crook[1,2,3,4], Tim EA Peto[1,2,3,4], A. Sarah Walker[1,2,4]

1. National Institute for Health Research Health Research Protection Unit in Healthcare Associated Infections and Antimicrobial Resistance, Oxford, UK
2. Nuffield Department of Medicine, University of Oxford, UK
3. Oxford University Hospitals National Health Service Foundation Trust, Oxford, UK
4. National Institute for Health Research Biomedical Research Centre, Oxford, UK

* Corresponding author: Emma Pritchard

Email: emma.pritchard@ndm.ox.ac.uk

Full postal address: Microbiology Level 7, John Radcliffe Hospital, Oxford OX3 9DU, UK

Phone Number: 01865 222196

Email of all other authors:

Nicola Fawcett: nicola.fawcett@spc.ox.ac.uk

T. Phuong Quan: phuong.quan@ndm.ox.ac.uk

Derrick Crook: derrick.crook@ndm.ox.ac.uk

Tim EA Peto: tim.peto@ndm.ox.ac.uk

A. Sarah Walker: sarah.walker@ndm.ox.ac.uk

ABSTRACT (200 words (limit 200))

Objective: To investigate variation in the presence of secondary diagnosis codes in Charlson and Elixhauser comorbidity scores and assess whether including a one-year lookback period improved prognostic adjustment by these scores individually, and combined, for 30-day mortality.

Study Design and Setting: We analysed inpatient admissions from 01-Jan-2007 to 18-May-2018 in Oxfordshire, UK. Comorbidity scores were calculated using secondary diagnostic codes in the diagnostic-dominant episode, and primary and secondary codes from the year before. Associations between scores and 30-day mortality were investigated using Cox models with natural cubic splines for non-linearity, assessing fit using Akaike Information Criteria.

Results: The one-year lookback improved model fit for Charlson and Elixhauser scores vs using diagnostic-dominant methods. Including both, and allowing non-linearity, improved model fit further. The diagnosis-dominant Charlson score and Elixhauser score using a 1-year lookback, and their interaction, provided the best comorbidity adjustment (reduction in AIC: 761 from best single score model).

Conclusion: The Charlson and Elixhauser score calculated using primary and secondary diagnostic codes from 1-year lookback with secondary diagnostic codes from current episode improved individual predictive ability. Ideally, comorbidities should be adjusted for using both the Charlson (diagnostic-dominant) and Elixhauser (one-year lookback) scores, incorporating non-linearity and interactions for optimal confounding control.

Keywords: Charlson; Elixhauser; Comorbidity; Electronic Health Records; ICD-10; Confounding

Running Title: Combining Charlson and Elixhauser scores with varying lookback better predicted mortality

Word count: (3139 words, limit 3,000)

What is new?

Key findings:

- When calculating the Charlson and Elixhauser score, using primary and secondary diagnostic codes from a year prior to current admission in addition to secondary codes in the diagnostic-dominant episode of the current admission provides better adjustment for comorbidity than using the latter alone.
- Using both the Charlson and Elixhauser score in a single model, and allowing for non-linearity in associations and interactions between the two scores, provides better adjustment for comorbidity, compared with using each score separately.

What this adds to what is known?

- Other studies have shown a 1-year lookback to be superior in disease-specific populations, but here we show this in a larger and more general population. Additionally, prognostic ability is improved further when allowing for non-linear effects for both scores, and interactions between them.
- While the Charlson and Elixhauser score are used extensively to adjust for confounding in studies using electronic health records, no study was found which uses them in combination in their widely validated form.

What is the implication, what should change now?

- When adjusting for comorbidities in studies using electronic health records, both the Charlson and Elixhauser score should be included, with the Charlson score calculated using secondary diagnostic codes from the diagnostic-dominant episode in the current admission and the Elixhauser score additionally including primary and secondary diagnostic codes from the year prior to the current admission, and incorporating non-linear associations and their interaction.

1. Introduction

Many observational analyses of electronic health records adjust for patient co-morbid status to reduce confounding bias. Previously published literature(1) state that the Charlson comorbidity score(2) should be calculated using the diagnosis-dominant episode from a hospital spell ("the total continuous stay of a patient [..] on premises controlled by a Health Care Provider"(3)), however this is not consistent with some studies using a one-year lookback(4-6) and no clear guidance. The diagnosis-dominant episode is the first consultant episode in a hospital spell, or the second if the first episode's primary diagnosis (the main condition treated, coded using International Classification of Diseases (ICD) 10) is an "R" code (ICD-10 category containing signs and symptoms) and the second episode's is not(7). Each episode has up to 20 secondary codes for other relevant conditions.

The Charlson score is amongst the most widely used comorbidity scores(8), and was originally derived using a narrow clinical population of 604 female breast cancer patients. It has been adapted for use in administrative data numerous times since(9-15). The Elixhauser score(16) was developed using a large (n=1,779,167), more representative adult hospitalised population from multiple institutions (n=438) for use in administrative data. The resulting modified single score(17) used diagnostic codes from the current admission with a "type" indicating preadmission comorbidity (analogous to those from the diagnosis-dominant episode), and comorbidities from all previous hospital admissions, implying other studies do the same. A systematic review of 54 papers prior to March 2011 found the Elixhauser score to generally be superior to Charlson(18); a conclusion echoed by studies since(6, 19-30). However, none of these studies combined the scores in their validated forms in one model, and this does not seem to have been considered in the literature.

A 1-year lookback period for calculating the Charlson score was introduced in 1992(9), incorporated in score validation(11), and is used in the literature. Studies investigating using a lookback period vs

current admission find arguments both for(9, 31) and against(32); however were all in disease-specific populations. One concern is that coding manuals state that comorbidities must be coded when they have co-existed in conjunction with, and affected the management of, the patient in the current episode(33), potentially excluding e.g. previous stroke in patients admissions with delirium. Secondary diagnostic codes relating to specific comorbidities may therefore be variably present, even if the underlying condition is not. Codes are assigned at discharge, and therefore conditions that develop after admission can also be included.

We therefore aimed to investigate variation in the presence of secondary diagnosis codes contributing to comorbidity scores and assess whether using diagnosis codes over a one-year lookback period improved prognostic adjustment for 30-day mortality for the Charlson and Elixhauser scores individually, and combined in one model, allowing for non-linearity and interaction.

2. Methods

Data came from the Infections in Oxfordshire Research Database (IORD), which contains all admissions to the four hospitals within the Oxford University NHS Foundations Trust from 1st April, 1997(34). These hospitals provide all acute care and pathology services to a population of ~680,000 in the region. Information on out-of-hospital mortality is updated through the National Health Service clinical Spine application(35). IORD has generic Research Ethics Committee, Health Research Authority and Confidentiality Advisory Group approvals (14/SC/1069, ECC5-017(A)/2009).

Electronic coding of secondary codes was introduced 01st Jan, 2006. To reduce coding depth bias, we included all inpatient spells from 01st Jan 2008 to 03rd March 2019, allowing a one-year implementation period and a full one-year lookback for all episodes. Charlson and Elixhauser scores summed 17 and 31 weighted comorbidities, respectively. Charlson weightings were an updated and currently recommended version which changed former weightings to reflect changing mortality(1). Elixhauser calculations used van Walraven weightings(17). For each, we primarily compared two methods: (i) using only secondary ICD-10 codes from the diagnosis-dominant episode in an admission ("DiagDom") (ii) using all primary and secondary ICD-10 codes from the year prior to current admission, plus all secondary ICD-10 codes from the diagnosis-dominant episode ("Lookback_1y"), adding each comorbidity to the score only once if it occurred in several admissions. Sensitivity analyses either (iii) considered the mean number of admissions with an ICD-10 code for that condition in the previous year in (ii) rather than just presence/absence (details in **Supplementary Methods**), to reduce impact of single erroneous codes ("Lookback_weighted") or (iv) included all primary and secondary codes in the year prior to diagnosis, but not secondary ICD-10 codes from the diagnosis-dominant episode ("Lookback_1yonly"), to exclude coded conditions due to presenting disease. The Charlson score was truncated at 0 as recommended(1) (0.3% observations changed from -1 to 0). For consistency, the lowest 0.3% Elixhauser scores were truncated to -5.

2.1 Statistical Analyses

We first estimated variation in the individual codes within the scores by exploring the consistency of their use across patient admissions over time. We then compared the different score calculations by their prognostic impact on time from admission to the earliest of death (in or out of hospital) or 30-days using Cox models (details in **Supplementary Methods**). All models were adjusted for confounders identified in a previous IORD study(36), truncating continuous variables at their 99th percentiles to reduce influence of outliers (details in **Supplementary Methods**). Associations with scores were considered as linear on the log-hazard scale and non-linearly using natural cubic splines with 3 knots placed at the 10th, 50th, and 90th percentile. The Akaike Information Criteria (AIC) was used to assess model performance. Further models included interactions between the "best" (most prognostic) method for calculating the Charlson score and all Elixhauser methods, and the "best" method for calculating the Elixhauser score and all other Charlson methods. Stata 15.1 was used for all analyses.

3. Results

Analysis included 1,004,552 admissions in 454,513 patients; death occurred within 30-days of admission in 35,496 (3.5%) admissions. 18,360 (1.7%) admissions were censored at discharge. The median age at admission was 56.2 (IQR 31.2-74.2) years, 50.4% were females, and 83.4% of white ethnicity (**Table 1**). Most patients were admitted on a weekday (80.7%), as an emergency (70.3%), around half (50.1%) under a medical speciality, and with median 4 (IQR 2-6) secondary diagnostic codes in the diagnosis-dominant episode. 472,517 (47.0%) admissions had one (or more) admissions in the prior year (maximum 156); and hence scores could potentially differ depending on the calculation method. These admissions were in slightly older individuals (median 61.0 years) but were otherwise broadly similar to those without previous admissions in the prior year (**Table 1**).

	Included OUH admissions 2008-2019 (n=1,004,552)	Admissions with another admission in the prior year* (n=472,517)	Admission with no admission in prior year (n=532,035)
Age at last birthday (years)	56.2 (31.2-74.2)	61.0 (35.8-76.7)	51.9 (28.3-71.4)
Female	506,679 (50.4)	237,632 (50.3)	269,047 (50.6)
Ethnicity			
White	837,797 (83.4)	414,266 (87.7)	423,531 (79.6)
Black	12,593 (1.3)	5,759 (1.2)	6,834 (1.3)
Asian	33,104 (3.3)	15,361 (3.3)	17,743 (3.3)
Other	18,547 (1.9)	8,146 (1.7)	10,401 (2.0)
Unknown	102,511 (10.2)	28,985 (6.1)	73,526 (13.8)
Admission method			
Elective	261,547 (26.0)	118,267 (25.0)	143,280 (26.9)
Emergency	706,302 (70.3)	341,969 (72.4)	364,333 (68.5)
Other	36,703 (3.7)	12,281 (2.6)	24,422 (4.6)
Admission Source			
Usual residence	925,136 (92.1)	441,711 (93.5)	483,425 (90.9)

	Included OUH admissions 2008-2019 (n=1,004,552)	Admissions with another admission in the prior year* (n=472,517)	Admission with no admission in prior year (n=532,035)
Temporary residence	5,503 (0.6)	2,101 (0.4)	3,402 (0.6)
NHS general ward	65,814 (6.6)	25,205 (5.3)	40,609 (7.6)
Other	8,099 (0.8)	3,500 (0.7)	4,599 (0.9)
Consultant code			
Surgery	482,215 (48.0)	195,329 (41.3)	286,886 (53.9)
Medicine	503,558 (50.1)	264,535 (56.0)	239,023 (44.9)
Other	18,779 (1.9)	12,653 (2.7)	6,126 (1.2)
Admission day of week			
Weekday	810,884 (80.7)	384,342 (81.3)	426,542 (80.2)
Weekend	193,668 (19.3)	88,175 (18.7)	105,493 (19.8)
Admission year	2013 (2010-2016)	2014 (2010-2016)	2013 (2011-2016)
Any complex admission[†] in last year	87,754 (8.7)	87,754 (18.6)	532,035 (0.0)
Admissions in the last year	0 (0-2)	2 (1-3)	0 (0.0)
Number of diagnostic codes in this admission	4 (2-6)	5 (2-7)	3 (2-6)

	Included OUH admissions 2008-2019 (n=1,004,552)	Admissions with another admission in the prior year* (n=472,517)	Admission with no admission in prior year (n=532,035)
Clinical Classification Software groups (10 most prevalent)			
Low risk	139,380 (13.9)	58,346 (12.4) [1]	81,034 (15.2) [1]
Non-specific chest pain	87,630 (8.7)	37,123 (7.9) [3]	50,507 (9.5) [3]
Cancer	80,512 (8.0)	52,416 (11.1) [2]	28,096 (5.3) [7]
Headache	78,615 (7.8)	30,538 (6.5) [4]	48,077 (9.0) [4]
Other	74,386 (7.4)	27,043 (5.7) [6]	47,343 (8.9) [5]
Superficial injury	69,100 (6.9)	18,084 (3.8) [8]	51,016 (9.6) [2]
Enteritis and ulcerative colitis	61,784 (6.2)	27,882 (5.9) [5]	33,902 (6.4) [6]
Complication of device	36,646 (3.7)	25,930 (5.5) [7]	-
Influenza	30,528 (3.0)	-	17,920 (3.4) [9]
Spondylosis	28,638 (2.9)	-	20,561 (3.9) [8]
Urinary Tract Infections	-	15,504 (3.3) [9]	-
Pneumonia	-	14,283 (3.0) [10]	-

	Included OUH admissions 2008-2019 (n=1,004,552)	Admissions with another admission in the prior year* (n=472,517)	Admission with no admission in prior year (n=532,035)
Skin and superficial tissue infections	-	-	13,154 (2.5) [10]
Deaths within 30 days of admission	35,496 (3.5)	22,765 (4.8)	

136

137 * In these admissions, the scores calculated in the diagnostic-dominant and lookback methods may differ

138 † Two or more consultant episodes within an admission

139 Note: Data are n (%) or median (IQR).

140

141 Table 1: Cohort Characteristics

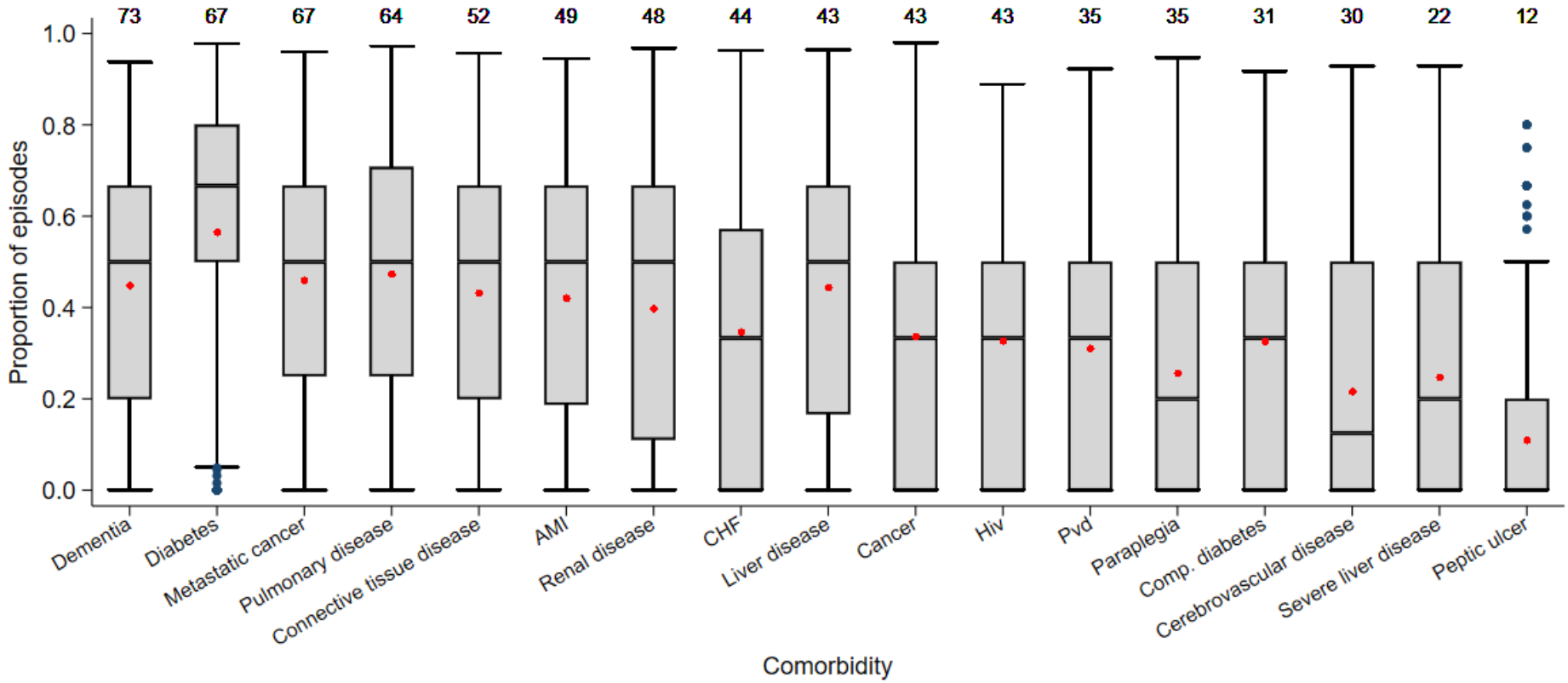
142

3.1 *Variation in codes contributing to comorbidity scores*

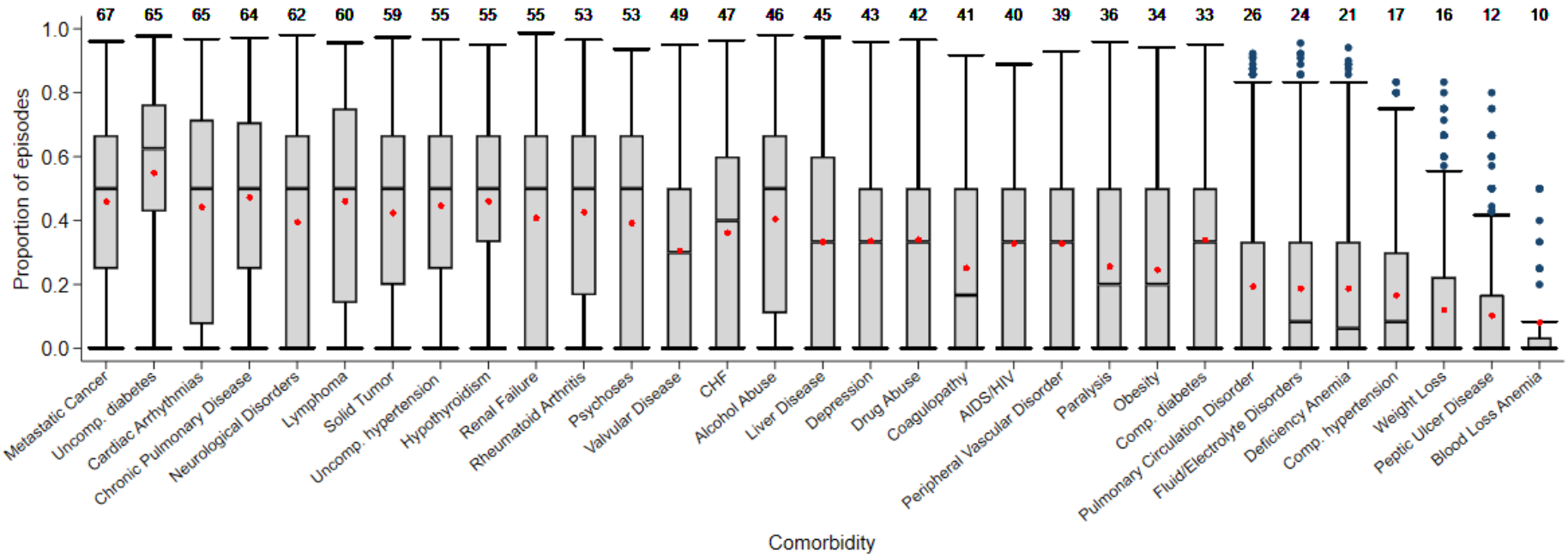
As motivation for this study, we first considered variation in codes over time. New codes representing comorbidities may legitimately arise but codes should not be lost assuming the comorbidity remains present. However, **Figure 1A** shows large variation in the proportion of diagnosis-dominant episodes with a Charlson comorbidity ICD-10 code after the first occurrence of a code for that comorbidity. Peptic ulcer and severe liver disease were the least consistently recorded comorbidities, with only 12% and 22% of individuals having the respective ICD-10 codes recorded in all diagnosis-dominant episodes following first occurrence of the codes, possibly reflecting curability of peptic ulcer disease. Dementia was the most consistently recorded comorbidity, with 73% of individuals having relevant codes in all subsequent diagnosis-dominant episodes. Findings were similar for Elixhauser components (**Figure 1B**); e.g. 67% of individuals had metastatic cancer codes in all subsequent diagnosis-dominant episodes after first occurrence.

The inconsistency in diagnostic code recording explains changes in comorbidity scores when comparing the diagnostic-dominant and 1-year lookback method (examples in **Supplementary Figure 1**). Large proportions of admissions in individuals who had a comorbidity code earlier in the study period but not in the current diagnostic-dominant episode had 1, 2, or 3+ uses of diagnostic codes for that comorbidity in the year prior to the current admission (i.e. would be included in the lookback) (**Supplementary Figure 2**). For example, 24% of diagnostic-dominant admissions without a Charlson metastatic cancer code (but at least one code previously) had ≥3 metastatic cancer-coded episodes in the prior year. Patterns were similar for all comorbidities in Charlson and Elixhauser scores.

A – Charlson score



B: Elixhauser Score



Note: The number on top of the plot shows the percent of individuals with a proportion of 1. The box plots exclude those with a proportion of 1. The red dots show the mean proportion out of those not equal to 1.

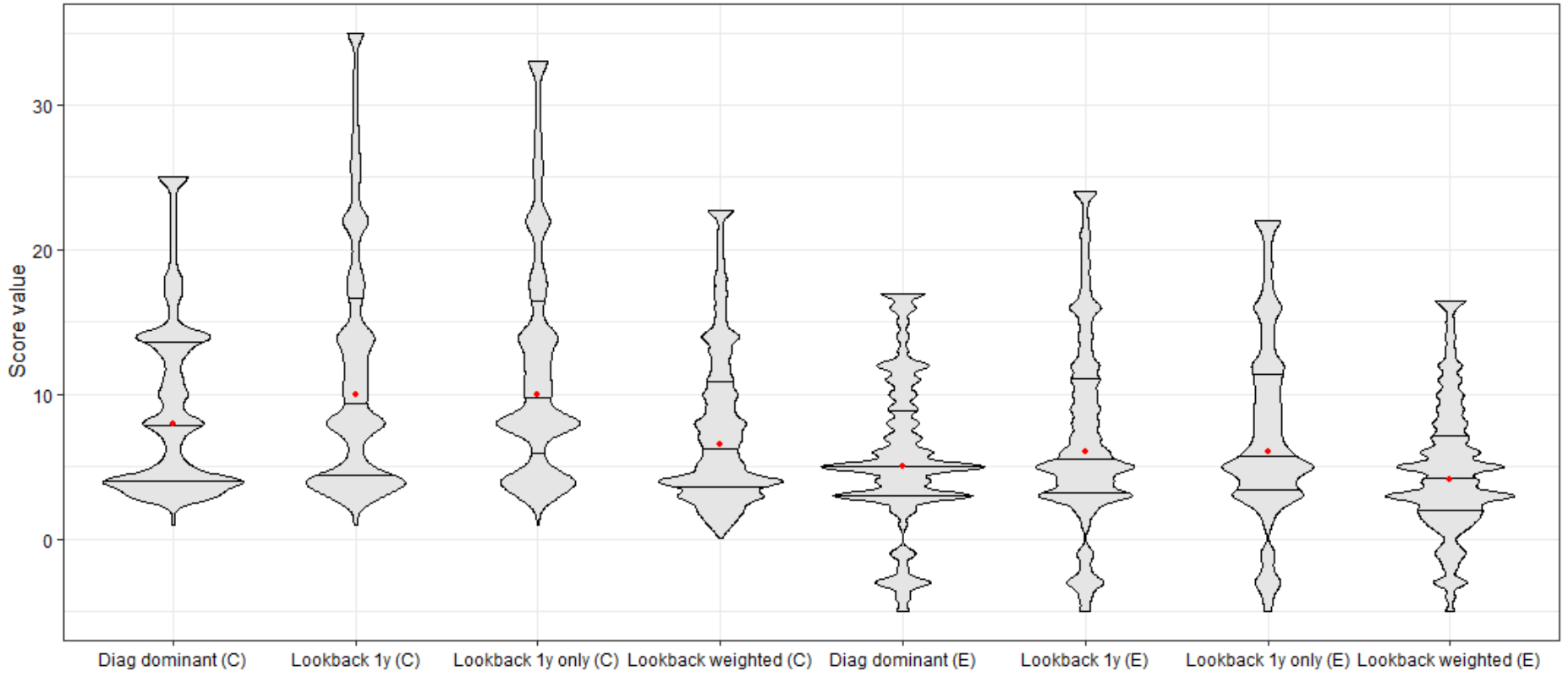
Figure 1: Consistency of ICD-10 recording for the comorbidities in diagnostic-dominant episodes strictly after a first occurrence of a code for that comorbidity.

3.2 *Comorbidity score distributions*

Comorbid condition prevalence ranged from <0.1%: to 8.4% (Charlson) and 16.7% (Elixhauser) of all episodes, and 10.7% (Charlson) and 21.0% (Elixhauser) of diagnosis-dominant episodes (**Supplementary Table 1**). Pulmonary disease was the most common Charlson condition (10.7% of dominant episodes), followed by diabetes (9.5%), compared with uncomplicated hypertension (21.0%), and cardiac arrhythmias (10.7%) for Elixhauser.

Adding all primary and secondary ICD-10 codes in the year prior to current admission to those from the diagnosis-dominant episode (lookback_1y) reduced the proportion of admissions with zero Charlson score from 65.8% to 57.1% (60.3% to 51.0% for Elixhauser) (**Figure 2**). The median Charlson score increased from 8 to 10 and the 99th percentile from 25 to 35; the median Elixhauser score increased from 5 to 6 and the 99th percentile increased from 17 to 24. Including prior codes reduced the density of both scores, particularly at lower values.

Weighting codes in previous admissions by their prevalence (Lookback_weighted) reduced the range of values (0-22.75), giving this method a denser distribution (**Figure 2**). 532,035 (53.0%) admissions had no prior admission in the previous year, meaning scores calculated using only prior codes (lookback_1yonly) had much higher percentages of zeros (73.5% of admissions for Charlson score). Despite this, the distribution of non-zero values was very similar to lookback_1y (**Figure 2**).



Zero score, N (%)	660,936 (65.8)	573,515 (57.1)	733,534 (73.5)	573,610 (57.1)	605,411 (60.3)	512,372 (51.0)	738,003 (73.5)	509,348 (50.7)
Median (IQR)	8 (4,14)	10 (4,17)	10 (6,17)	6.5 (4, 11)	5 (3, 9)	6 (3, 11)	6 (3, 12)	4.2 (2, 7.4)
Range	[0, 25]	[0,35]	[0,33]	[0,22.75]	[-5, 17]	[-5, 24]	[-5, 22]	[-5, 16.5]

Note: The violin plots and the median and IQR are for non-zero values only. The percent zero and range are for all values of all scores. The red dot shows the mean value of the non-zero scores.

Figure 2: Charlson (C) and Elixhauser (E) scores calculated using different methods.

3.3 Impact of different calculation methods on associations with 30-day mortality

There was evidence of significant non-linearity in the effects of all the comorbidity scores, although the actual impact was relatively small for the Elixhauser score (**Supplementary Figure 3A**). All scores had a greater increase in mortality risk per unit higher score at lower values versus higher values. Despite its more limited range, risk increases per unit higher score were greater for Elixhauser than all Charlson scores, even more so at higher score values, with risk increases attenuating markedly for higher Charlson scores. Individually, the lookback_1y method for both the Charlson and the Elixhauser score provided the best fit (lowest AIC) when associations with mortality were modelled non-linearly using splines (**Table 2**). As the weighted Charlson score was highly correlated with the lookback_1y method (correlation coefficient=0.90, **Table 3**), and the lookback_1y method performed better than the weighted method when allowing for non-linear affects (AIC: 888,738 [lookback_1y] vs 888,811 [weighted]), the weighted method was not considered further in combined models.

Consistent with previous comparisons, the Elixhauser score calculated under both DiagDom and Lookback_1y methods had substantially lower AICs than the Charlson score calculated under all four methods. However, including both individual scores in the same model improved model fit further (**Table 2**); increases in DiagDom Charlson score from approximately 15 upwards did not affect mortality, whereas risk continued to increase with higher Lookback_1y Elixhauser scores (**Supplementary Figure 3B**).

Adding interactions between each of the spline terms for the Charlson and Elixhauser scores further improved model fit (**Table 2**), to a much greater degree than including the two scores as main effects. For example, AIC dropped from 885,874 with main non-linear effects only to 885,611 with interactions for the DiagDom Charlson/Lookback_1y Elixhauser model (correlation 0.61), which had the lowest AIC of all models including interactions. Risk was low when both Charlson and Elixhauser scores were low,

and high when either score was high (**Figure 3**); once one score was elevated, there was relatively little change in risk associated with increasing scores of the other, in contrast to continued risk increases in models without the interaction.

Most other factors in the multivariable models had similar effects across all models (**Supplementary Figure 4**). Clinical Classifications Software (CCS) group (reflecting the primary diagnosis) was the only factor with some differences in model estimates, likely due to similarities between primary codes in CCS groups and secondary codes in the Charlson and Elixhauser scores from the diagnostic-dominant episode. However, variation was small in magnitude with no clear pattern.

Model	AIC	Δ AIC from lowest
Single comorbidity variable		
DiagDom C (Linear)	889255	3644
DiagDom C (Spline)	889070	3459
Lookback_1y C (Linear)	889027	3416
Lookback_1y C (Spline)	888738	3126
Lookback_1yonly C (Linear)	890565	4953
Lookback_1yonly C (Spline)	890504	4893
Lookback_weighted C (Linear)	888950	3339
Lookback_weighted C (Spline)	888811	3199
DiagDom E (Linear)	886926	1315
DiagDom E (Spline)	886913	1302
Lookback_1y E (Linear)	886 587	976
Lookback_1y E (Spline)	**886372**	**761**
Additive Models (all non-linear effects)		
DiagDom C + Lookback_1y E	**885874**	**263**
Lookback_1y C + DiagDom E	886222	610
Lookback_1y C + Lookback_1y E	886033	421
Interaction Models (all splines)		
DiagDom C * Lookback_1y E	**885611**	**0**
Lookback_1y C * DiagDom E	885775	164
Lookback_1y C * Lookback_1y E	885759	148

Table 2: AIC of adjusted models for 30-day mortality with different Charlson and Elixhauser scores

Comorbidity score	Charlson DiagDom	Charlson Lookback_1y	Charlson Lookback_1yonly	Charlson Weighted	Elixhauser DiagDom	Elixhauser Lookback_1y
Charlson DiagDom	1					
Charlson Lookback_1y	0.82	1				
Charlson Lookback_1yonly	0.50	0.84	1			
Charlson Weighted	0.87	0.90	0.68	1		
Elixhauser DiagDom	0.69	0.62	0.44	0.63	1	
Elixhauser Lookback_1y	0.61	0.76	0.67	0.67	0.83	1

Table 3: Pearson's correlation coefficient between the different methods of calculation for the Charlson and Elixhauser score for diagnostic dominant episodes.

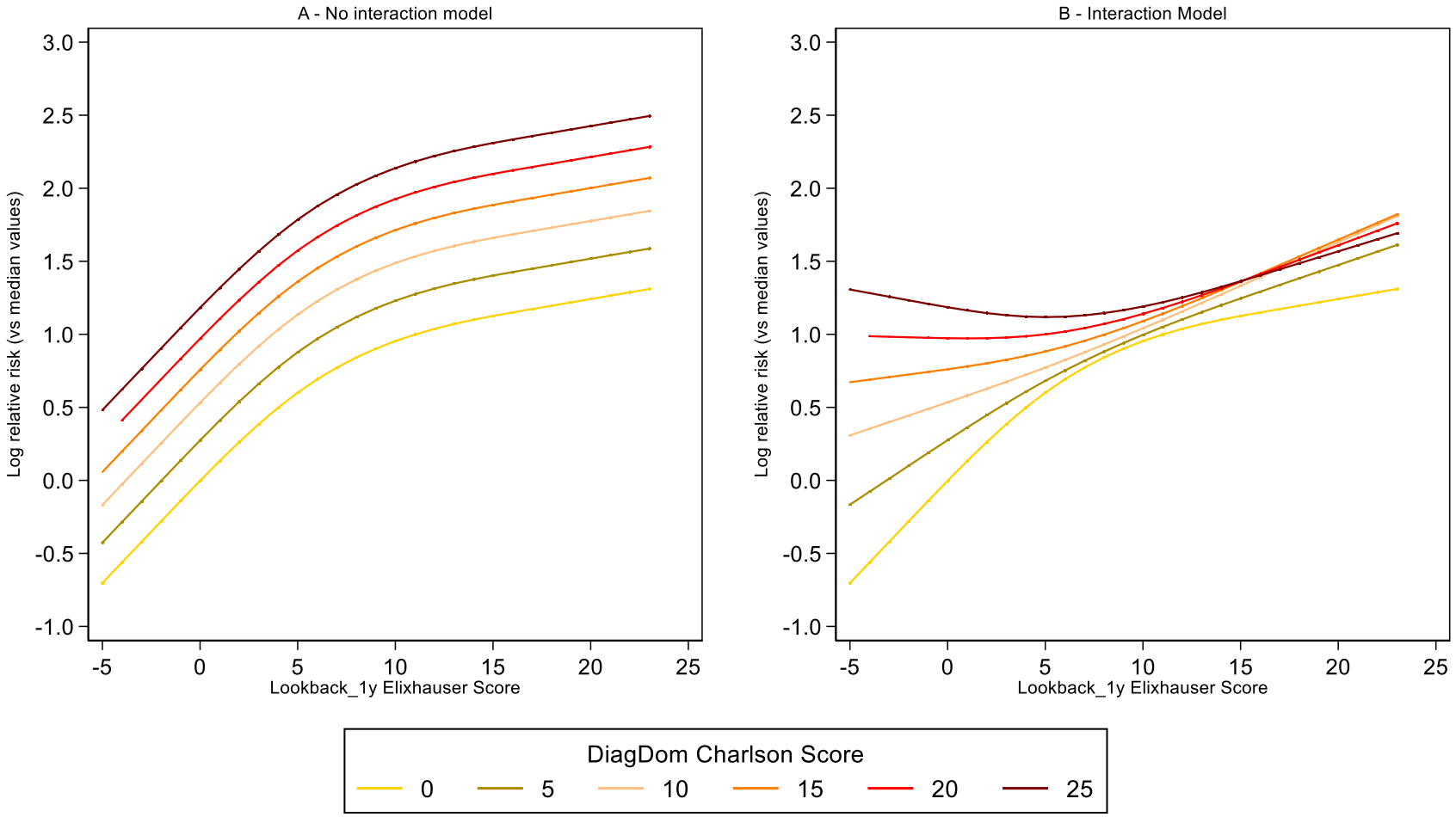


Figure 3: Estimated overall association between the DiagDom Charlson and Lookback_1y Elixhauser scores and 30-day mortality including main effects only (left) and interactions (right).

4. Discussion

We found that including comorbidities from the previous year when calculating the Charlson and Elixhauser score improved model fit when each was considered individually, implying superior control of confounding, compared with only using comorbidities recorded in the diagnostic-dominant episode. The Elixhauser score was more prognostic than the Charlson score when using one score to adjust for risk of 30-day mortality, as shown in previous studies(6, 18-20, 22-29). However, we found that including both comorbidity scores together with their interaction, and incorporating non-linear effects, provided the best model fit.

Most previous studies have concluded that the Elixhauser score provides superior comorbidity risk-adjustment, as here, possibly due to the greater number of included comorbidities achieving better discrimination(20, 37). Other explanations could include differences in coding algorithms; e.g., renal disease occurs in 4.2% of diagnostic-dominant episodes using the Charlson algorithm, versus 4.9% for the Elixhauser algorithm (**Supplementary Table 1**), or variations in scoring systems to define weightings with regression-coefficient based models (e.g.(38, 39)) outperforming mathematically-incorrect risk-ratio derivations (e.g. (2, 11)) for the Charlson score(40). Southern et al.(41) concluded that differences between scores in associations with mortality are due to both additional comorbidities in the Elixhauser score and updated coding for variables present in the Charlson score.

Including comorbidities reflected by primary and secondary diagnostic codes over the year preceding the current admission was superior for both individual scores in terms of impact on 30-day mortality risk. Deyo et al.(9) first showed including lookback improved model fit for the Charlson score, and our results extend this investigation to updated weights on a current and wider population and additionally for the Elixhauser score. A similar study found that using a 1-year lookback for the Charlson score, rather than just the current admission, improved model fit for 1-year mortality, but not in-hospital

mortality, 30-day mortality, 30-day readmission, and length of stay(31).This study was restricted to hip fracture patients aged ≥65 years, so may not be generalisable. Using a one-year lookback plus the admission of interest was recommended for the Charlson score by Armitage et al.(42), but this was not formally tested. While we only considered a maximum of 1-year lookback, studies investigating comorbidity consistency over a 5-year lookback(43) and entire lifetime records(26) showed that, while longer lookback periods captured more comorbidities and did not degrade model fit, using more than 1-year of lookback gave little improvement. These studies did not compare varying lookbacks with diagnostic-dominant episodes or consider the contribution of prior primary diagnostic codes. Our findings show a 1-year lookback to be superior to the diagnostic-dominant episode in a large non-specific population. Where data are not available to calculate a lookback period, our results support using the Elixhauser score and allowing for non-linearity.

Importantly, we found that including both the Charlson and Elixhauser scores in the same model gave better model fit versus either single comorbidity score, despite 9 of the 17 (Charlson)/31 (Elixhauser) comorbidities being in common; specifically diabetes (uncomplicated and complicated), congestive heart failure, HIV, metastatic cancer, renal disease, chronic pulmonary disease, rheumatic disease, and peripheral vascular disease. To our knowledge, other studies have not assessed both scores in a single model, possibly due to concerns about over-fitting. This was not a major concern in our study (**Table 3**, **Supplementary Figure 3**), potentially because of coding algorithm differences noted above. Two studies combined the Charlson and Elixhauser comorbidities into a single index, using 37 unique comorbidities with weightings based on Schneeweiss(38, 44) or 32 conditions with weightings based on three predefined derivation methods (Charlson(2), Schneeweiss(38), Van Walraven(17))(45). Similar to our findings, both concluded that the combined score performed better than either score alone, suggesting each provides additional information regarding 30-day mortality; however, they did not compare their combined score with models including separate scores and their interaction. Using the Charlson and Elixhauser scores as main effects, rather than a combined index, builds on extensive

297 research validating these scores for comorbidity adjustment in administrative data(46-50). Although a
298 combined score allows weightings to be derived from a single population, meaning they should be
299 more consistent, producing a new combined score with unique weightings requires additional quality
300 assessment and validation(45). Importantly, comorbidity scores should use weights derived from
301 regression-coefficient based models rather than risk-ratio based models as the latter are
302 mathematically incorrect(40).

303

304 Including interactions between Charlson and Elixhauser scores improved model fit substantially,
305 suggesting the Charlson score had a different effect on mortality dependent on the value of the
306 Elixhauser score and vice-versa. Interestingly, the best model included the Charlson score calculated
307 from the diagnostic-dominant episode, and the Elixhauser score calculated from the one-year
308 lookback, suggesting these two scores are capturing unique aspects of comorbidity. Using the
309 diagnosis-dominant episode may be important here as it captures the most active or currently
310 problematic diagnosis affecting current management(9). The estimated associations indicated
311 "thresholding" of risk at high values of either score, and low mortality risk only when both were low.
312 Importantly, since the greatest difference in predicted risk between main effects and interaction
313 models was at the highest values of either score (**Figure 3**), not incorporating interactions runs the risk
314 of residual confounding, since these patients may systematically differ on other characteristics.

315

316 We aimed to investigate different methods for calculating Charlson and Elixhauser scores in electronic
317 health records, which rely on ICD-10 coding. Although these codes are assigned by experienced coding
318 teams in hospitals, this data is collected for reimbursement(51) rather than epidemiology, purposes
319 which are not always congruent(52). Importantly, codes assigned are judged relevant to the current
320 admission, and pre-existing comorbidities may not always meet this condition. Other reasons why
321 assigned codes may not reflect clinical conditions include gaming/up-coding, where codes are

distorted to meet targets, or tunnel vision, where aspects of clinical performance which are measured are focused on and unmeasured areas are neglected(53). Coding will also be affected by poor communication between the patient and the clinician, as this reduces clarity of records provided to the coder(54). These limitations are common to all studies using ICD-10 codes.

The main specific limitation of our study is that it was conducted using a single population from Oxfordshire. ICD-10 coding practices may vary between different hospital groups, e.g. some codes may be preferentially used due to differences in on-site training. Further investigation into varying coding practices, and its impact on score calculation, across different hospital groups would be of interest. However, associations between different comorbidity score calculation methods and non-linear associations with outcomes should be less affected by geographical region. Our sample is also large, accounting for around 1% of the UK population, which increases generalisability.

In conclusion, we found that calculating Charlson and Elixhauser scores using primary and secondary ICD-10 codes from a 1-year lookback in addition to secondary ICD-10 codes from the diagnostic-dominant episode in an admission improved their individual predictive ability, but including both improved model fit further. We recommend that, for studies using electronic health records, comorbidities should be adjusted for using both the Charlson (diagnostic-dominant) and Elixhauser (one-year lookback) scores, incorporating non-linearity and including an interaction term between the two scores, to ensure maximum control of confounding.

ACKNOWLEDGEMENTS

We thank all the people of Oxfordshire who contribute to the Infections in Oxfordshire Research Database.

Research Database Team (Oxford): R Alstead, C Bunch, DW Crook, J Davies, J Finney, J Gearing (community), H Jones, L O'Connor, TEA Peto (PI), TP Quan, J Robinson (community), B Shine, AS Walker, D Waller, D Wyllie. Patient and Public Panel: G Blower, C Mancey, P McLoughlin, B Nichols.

Endocarditis database team (Leeds). MW Baig

NIHR Health Protection Research Unit Steering Committee: J Coia, N French, C Marwick, M Sharland.

Funding:

Financial support: The research was funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at the University of Oxford in partnership with Public Health England (PHE) [HPRU-2012-10041 and NIHR200915] and the NIHR Oxford Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or PHE. TEAP and ASW are NIHR Senior Investigators.

REFERENCE LIST

1. Dr Foster Intelligence. Understanding HSMRs. A Toolkit on Hospital Standardised Mortality Ratios. 2012 [Available from: https://pdfs.semanticscholar.org/4d5e/5abde666f78f27563382d40443d4bcf59e24.pdf].

2. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Diseases. 1987.

3. NHS. Hospital Provider Spell. 2019 [Available from: https://www.datadictionary.nhs.uk/data_dictionary/nhs_business_definitions/h/hospital_provider_spell_de.asp?shownav=1].

4. Downer B, Al Snih S, Raji M, Chou L-N, Kuo Y-F, Markides KS, et al. Healthcare utilization of Mexican-American Medicare beneficiaries with and without Alzheimer's disease and related dementias. PLoS One. 2020;15(1):e0227681-e.

5. Bannay A, Chaignot C, Blotière P-O, Basson M, Weill A, Ricordeau P, et al. The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. Medical care. 2016;54(2):188-94.

6. Kim C-Y, Sivasundaram L, LaBelle MW, Trivedi NN, Liu RW, Gillespie RJ. Predicting adverse events, length of stay, and discharge disposition following shoulder arthroplasty: a comparison of the Elixhauser Comorbidity Measure and Charlson Comorbidity Index. Journal of shoulder and elbow surgery. 2018;27(10):1748-55.

7. Aylin P, Bottle A, Hua Jen M, Middleton S. HSMR Mortality Indicator 2009.

8. Yurkovich M, Antonio Avina-Zubieta J, Thomas J, Gorenchtein M, Lacaille D. A systematic review identifies valid comorbidity indices derived from administrative health data. 2015.

9. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992;45(6):613-9.

10. D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. J Clin Epidemiol. 1996;49(12):1429-33.

11. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol. 2011;173(6):676-82.

12. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005;43(11):1130-9.

13. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. J Clin Epidemiol. 1993;46(10):1075-9; discussion 81-90.

14. Romano PS, Roost LL, Jollis JG. Further evidence concerning the use of a clinical comorbidity index with ICD-9-CM administrative data. Journal of Clinical Epidemiology. 1993;46(10):1085-90.

15. Sundararajan V, Henderson T, Perry C, Muggivan A, Quan H, Ghali WA. New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality. Journal of clinical epidemiology. 2004;57(12):1288-94.

16. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity Measures for Use with Administrative Data. 1998.

17. Van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical Care. 2009.

18. Sharabiani MT, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. Med Care. 2012;50(12):1109-18.

19. Chang H-J, Chen P-C, Yang C-C, Su Y-C, Lee C-C. Comparison of Elixhauser and Charlson Methods for Predicting Oral Cancer Survival. Medicine (Baltimore) 2016.

20. Gutacker N, Bloor K, Cookson R. Comparing the performance of the Charlson/Deyo and Elixhauser comorbidity measures across five European countries and three conditions. European Journal of Public Health. 2015.

21. Ladha KS, Zhao K, Quraishi SA, Kurth T, Eikermann M, Kaafarani HMA, et al. The Deyo-Charlson and Elixhauser-van Walraven Comorbidity Indices as predictors of mortality in critically ill patients. Open. 2015;5:8990.

22. Lieffers JR, Baracos VE, Winget M, Fassbender K. A comparison of charlson and elixhauser comorbidity measures to predict colorectal cancer survival using administrative health data. Cancer. 2011.

23. Menendez ME, Ring D. A Comparison of the Charlson and Elixhauser Comorbidity Measures to Predict Inpatient Mortality After Proximal Humerus Fracture. Journal of orthopaedic trauma. 2015;29(11):488-93.

24. Menendez ME, Neuhaus V, van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clinical orthopaedics and related research. 2014;472(9):2878-86.

25. Menendez ME, Ring D, Harris MB, Cha TD. Predicting In-Hospital Mortality in Elderly Patients With Cervical Spine Fractures: A Comparison of the Charlson and Elixhauser Comorbidity Measures. Spine. 2015;40(11):809-15.

26. Metcalfe D, Masters J, Delmestri A, Judge A, Perry D, Zogg C, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. BMC Medical Research Methodology. 2019;19(1).

27. Buhr RG, Jackson NJ, Kominski GF, Dubinett SM, Ong MK, Mangione CM. Comorbidity and thirty-day hospital readmission odds in chronic obstructive pulmonary disease: a comparison of the Charlson and Elixhauser comorbidity indices. BMC Health Serv Res. 2019;19(1):701.

28. Cai M, Liu E, Zhang R, Lin X, Rigdon SE, Qian Z, et al. Comparing the Performance of Charlson and Elixhauser Comorbidity Indices to Predict In-Hospital Mortality Among a Chinese Population. Clin Epidemiol. 2020;12:307-16.

29. Tsai KY, Hsieh KY, Ou SY, Chou FH, Chou YM. Comparison of Elixhauser and Charlson Methods for Discriminative Performance in Mortality Risk in Patients with Schizophrenic Disorders. Int J Environ Res Public Health. 2020;17(7).

30. Yang CC, Fong Y, Lin LC, Que J, Ting WC, Chang CL, et al. The age-adjusted Charlson comorbidity index is a better predictor of survival in operated lung cancer patients than the Charlson and Elixhauser comorbidity indices. Eur J Cardiothorac Surg. 2018;53(1):235-40.

31. Toson B, Harvey LA, Close JCT. The ICD-10 Charlson Comorbidity Index predicted mortality but not resource utilization following hip fracture. Journal of Clinical Epidemiology. 2015;68(1):44-51.

32. Dobbins TA, Creighton N, Currow DC, Young JM. Look back for the Charlson Index did not improve risk adjustment of cancer surgical outcomes. J Clin Epidemiol. 2015;68(4):379-86.

33. NHS Digital. National clinical coding standards: ICD 10 5th edition (2017), accurate data for quality information: Stationery Office; 2017.

34. Finney JM, Walker AS, Peto TEA, Wyllie DH. An efficient record linkage scheme using graphical analysis for identifier error detection. BMC Medical Informatics and Decision Making. 2011;11(1).

35. NHS Digital. Spine. 2019 [Available from: https://digital.nhs.uk/services/spine].

36. Walker AS, Mason A, Quan P, Fawcett NJ, Watkinson P, Llewelyn M, et al. Mortality risks associated with emergency admissions during weekends and public holidays: an analysis of electronic health records. wwwthelancetcom. 2017;390.

37. Stukenborg GJ, Wagner DP, Connors AF. Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. Medical care. 2001;39(7):727-39.

38. Schneeweiss S, Wang PS, Avorn J, Glynn RJ. Improved comorbidity adjustment for predicting mortality in Medicare populations. Health services research. 2003;38(4):1103-20.

39. Sullivan LM, Massaro JM, D'Agostino RB, Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med. 2004;23(10):1631-60.

40. Mehta HB, Mehta V, Girman CJ, Adhikari D, Johnson ML. Regression coefficient-based scoring system should be used to assign weights to the risk index. J Clin Epidemiol. 2016;79:22-8.

41. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Medical care. 2004;42(4):355-60.

42. Armitage JN, van der Meulen JH, Group RCoSC-mC. Identifying co-morbidity in surgical patients using administrative data with the Royal College of Surgeons Charlson Score. The British journal of surgery. 2010;97(5):772-81.

468 43. Preen DB, Holman CDAJ, Spilsbury K, Semmens JB, Brameld KJ. Length of comorbidity
469 lookback period affected regression model performance of administrative health data. Journal of
470 Clinical Epidemiology. 2006;59(9):940-6.

471 44. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted
472 mortality in elderly patients better than existing scores. Journal of Clinical Epidemiology. 2011.

473 45. Simard M, Sirois C, Candas B. Validation of the Combined Comorbidity Index of Charlson and
474 Elixhauser to Predict 30-Day Mortality Across ICD-9 and ICD-10. Medical care. 2018;56(5):441-7.

475 46. Bing, Couris CM, Fushimi K, Graham P, Hider P, Januel JM, et al. Updating and validating the
476 charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data
477 from 6 countries. American Journal of Epidemiology. 2011.

478 47. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index.
479 Journal of Clinical Epidemiology. 1994;47(11):1245-51.

480 48. Cronin-Fenton DP, Nørgaard M, Jacobsen J, Garne JP, Ewertz M, Lash TL, et al. Comorbidity
481 and survival of Danish breast cancer patients from 1995 to 2005. British Journal of Cancer.
482 2007;96:1462-8.

483 49. Frenkel WJ, Jongerius EJ, Mandjes-Van Uitert MJ, Van Munster BC, De Rooij SE. Validation of
484 the Charlson Comorbidity Index in Acutely Hospitalized Elderly Adults: A Prospective Cohort Study. J
485 Am Geriatr Soc. 2014;62:342-6.

486 50. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, et al. A new Elixhauser-
487 based comorbidity summary measure to predict in-hospital mortality. Medical care. 2015;53(4):374-
488 9.

51. Reforming NHS financial flows: introducing payment by results. London: Department of Health; 2002.

52. Fawcett N, Young B, Peto L, Quan TP, Gillott R, Wu J, et al. 'Caveat emptor': the cautionary tale of endocarditis and the potential pitfalls of clinical coding data-an electronic health records study. BMC Med. 2019;17(1):169.

53. Shaw J, Taylor R, Dix K. Uses and Abuses of Performance Data in Healthcare.; 2015.

54. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. Health Services Research2005.

55. NHS Digital. About the Summary Hospital-level Mortality Indicator (SHMI) 2019 [Available from: https://digital.nhs.uk/data-and-information/publications/ci-hub/summary-hospital-level-mortality-indicator-shmi].

Supplementary Methods

From 01st Jan 2007 to 03rd March 2019, there were 2,753,465 episodes from 2,505,488 spells in 750,361 individuals. We first excluded admissions before 01st Jan 2008 (n=200,407), to reduce coding bias. We excluded admissions with missing primary diagnostic codes (n=12,710). We excluded admissions in whom the mortality rate was very low, specifically retaining only ordinary admissions and excluding admissions which were day cases (n=855,243; 19 deaths within 30-days of admission), regular day admissions (n=209,965; 21 deaths), regular night admissions (n=24; 0 deaths), those using mother and baby delivery facilities only (n=0), and those missing this information (n= 2; 0 deaths). This left 1,242,925 episodes from 566,044 individuals. We then also excluded two groups with distinct reasons for hospitalisation, indicated by low comorbidities and low mortality, specifically maternity admissions (admission method code 31 (admitted ante partum) or 32 (admitted post-partum); 144,358 spells; 11 deaths; 95.1% with a Charlson score of 0 calculated using the diagnostic dominant method) and new-born babies (admission method code 82 (the birth of a baby within current health care provider) or 83 (baby born outside of health care provider except when born at home as intended) with an age less than 0.2 years (empirically chosen); 94,002 spells, 681 (0.7%) deaths, 99.98% with a Charlson score of zero). This left 1,004,565 admissions in 454,526 individuals; 13 admissions from 13 individuals with unknown sex were excluded, leaving 1,004,552 admissions from 454,513 individuals for analysis.

Due to the large number of admissions, we included all of the following administrative variables in Cox regression models to adjust for confounding as fully as possible, as previously identified through variable selection in a previous IORD study(36). These were age, sex, ethnicity (White, Black, Asian, Other, Unknown), admission source (usual residence, temporary residence, NHS general ward, other), admission method (elective, emergency, or other), Clinical Classifications Software (CCS) group(55) (35 categories), consultant code (surgery, medical, and other), admission day of week, admission year,

admission day of year, admission hour, any complex admission in last year (defined as admissions with two or more consultant episodes), number of admissions in last year, and number of diagnosis codes. CCS groups containing fewer than 3000 individuals were combined into "other" and CCS groups with <1% mortality were combined into "low risk" to improve model stability and convergence. Admission day of year was modelled using a sin()+cos() function to ensure a smooth transition in risk between years. Natural cubic splines were used for non-linear effects of continuous variables. Pairwise interactions were included based on selection in the previous study; specifically admission hour with admission day of week, number of admissions in previous year with number of complex admissions in the previous year, and age with number of admissions in the last year. For individuals with no vital status checked >30 days after admissions and were not known to have died were censored at discharge date.

Calculating the Lookback_weighted score

For each comorbidity of the Charlson score, the number of occurrences in the last year (including in the diagnostic-dominant episode) were counted and then divided by the total number of admissions in that timeframe. This resulted in a number between 0 (no occurrences of the comorbidity in the prior year) and 1 (comorbidity present in all admissions in the prior year). This was then multiplied by the individual comorbidity weighting e.g. 14 for dementia. This process was repeated for all 17 comorbidities, and then the comorbidity weightings were summed together to produce an overall score. This score was always less than or equal to the lookback_1y method as, for the lookback_1y method, the total weighting of the comorbidity was considered irrespective of the number of times the comorbidity occurred in the prior year.

Supplementary Table 1: Frequency of the conditions constituting the Charlson and the Elixhauser scores in episodes and individuals.

(A) Charlson Comorbidities

Charlson Component	Score weighting	Episodes, n (%)	Individuals, n (%)	Diagnostic-dominant Episodes, n (%)
TOTAL		**2,225,998**	**454,513**	**1,004,552**
Acute myocardial infarction	5	74,791 (3.4)	24,611 (5.4)	43,370 (4.3)
Cerebral vascular disease	11	30,316 (1.4)	14,443 (3.2)	16,876 (1.7)
Congestive heart failure	13	53,980 (2.4)	19,614 (4.3)	30,087 (3.0)
Connective tissue disorder	4	30,253 (1.4)	9,744 (2.1)	17,575 (1.8)
Dementia	14	45,684 (2.1)	15,314 (3.4)	28,482 (2.8)
Diabetes	3	173,001 (7.8)	40,712 (9.0)	95,481 (9.5)
Liver disease	8	11,801 (0.5)	2,986 (0.7)	6,042 (0.6)
Peptic ulcer	9	7,928 (0.4)	4,816 (1.1)	3,418 (0.3)
Peripheral vascular disease	6	33,951 (1.5)	14,619 (3.2)	19,905 (2.0)
Pulmonary disease	4	186,372 (8.4)	60,764 (13.4)	107,102 (10.7)
Cancer	8	63,971 (2.9)	25,486 (5.6)	37,135 (3.7)
Diabetes complications	-1	18,411 (0.8)	5,823 (1.3)	9,199 (0.9)
Paraplegia	1	14,434 (0.7)	5,903 (1.3)	8,179 (0.8)
Renal disease	10	83,180 (3.7)	24,067 (5.3)	42,050 (4.2)

Charlson Component	Score weighting	Episodes, n (%)	Individuals, n (%)	Diagnostic-dominant Episodes, n (%)
Metastatic cancer	14	57,437 (2.6)	16,008 (3.5)	29,903 (3.0)
Severe liver disease	18	5,007 (0.2)	1,898 (0.4)	2,255 (0.2)
HIV	2	1,067 (0.1)	392 (0.1)	536 (0.1)

552

553 **(B) Elixhauser Comorbidities**

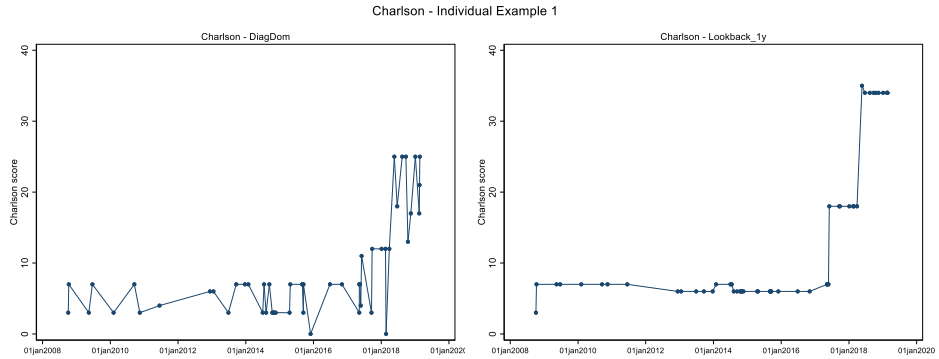
Elixhauser Component	Score weighting	Episodes, n (%)	Individuals, n (%)	Diagnostic-dominant episodes
TOTAL		**2,225,998**	**454,513**	**1,004,552**
Congestive Heart Failure	7	62,412 (2.8)	21,655 (4.8)	34,603 (3.4)
Cardiac Arrhythmias	5	181,676 (8.2)	58,269 (12.8)	107,308 (10.7)
Valvular Disease	-1	60,854 (2.7)	23,794 (5.3)	33,286 (3.3)
Pulmonary circulation disorders	4	14,400 (0.7)	6,685 (1.5)	7,888 (0.8)
Peripheral Vascular Disorders	2	37,432 (1.7)	15,730 (3.5)	21,785 (2.2)
Hypertension, uncomplicated	0	372,486 (16.7)	117,271 (25.8)	210,419 (21.0)
Paralysis	7	17,074 (0.8)	7,129 (1.6)	9,887 (1.0)
Other Neurological Disorders	6	63,958 (2.9)	21,308 (4.7)	39,232 (3.9)
Chronic Pulmonary Disease	3	186,934 (8.4)	60,865 (13.4)	107,449 (10.7)

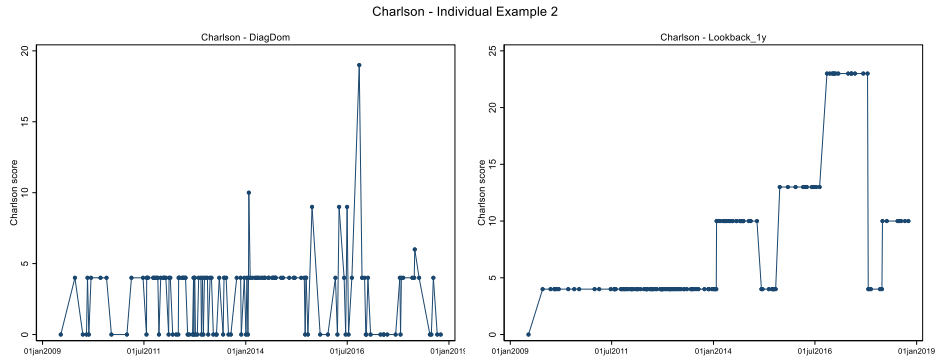
Elixhauser Component	Score weighting	Episodes, n (%)	Individuals, n (%)	Diagnostic-dominant episodes
Diabetes, Uncomplicated	0	168,869 (7.6)	40,348 (8.9)	92,936 (9.3)
Diabetes, Complicated	0	21,694 (1.0)	6,794 (1.5)	11,185 (1.1)
Hypothyroidism	0	58,639 (2.6)	20,110 (4.4)	36,127 (3.6)
Renal Failure	5	290,631 (13.1)	24,725 (5.4)	49,448 (4.9)
Liver Disease	11	31,660 (1.4)	10,187 (2.2)	16,609 (1.7)
Peptic Ulcer Disease	0	6,955 (0.3)	4,335 (1.0)	2,998 (0.3)
AIDS/ HIV	0	972 (0.0)	360 (0.1)	500 (0.1)
Lymphoma	9	9,961 (0.5)	2,980 (0.7)	6,373 (0.6)
Metastatic Cancer	12	57,437 (2.6)	16,008 (3.5)	16,008 (3.5)
Solid Tumour w/o metastasis	4	60,701 (2.7)	23,180 (5.1)	37,656 (3.8)
Rheumatoid Arthritis/ Collagen Vascular	0	37,895 (1.7)	12,171 (2.7)	21,455 (2.1)
Coagulopathy	3	10,515 (0.5)	4,914 (1.1)	6,288 (0.6)
Obesity	-4	27,146 (1.2)	14,625 (3.2)	13,618 (1.4)
Weight Loss	6	14,273 (0.6)	9,578 (2.1)	6,386 (0.6)
Fluid and Electrolyte Disorders	5	56,679 (2.6)	29,807 (6.6)	34,245 (3.4)
Blood Loss Anaemia	-2	465 (0.0)	311 (0.1)	289 (0.0)
Deficiency Anaemia	-2	25,765 (1.2)	13,074 (2.9)	12,549 (1.3)

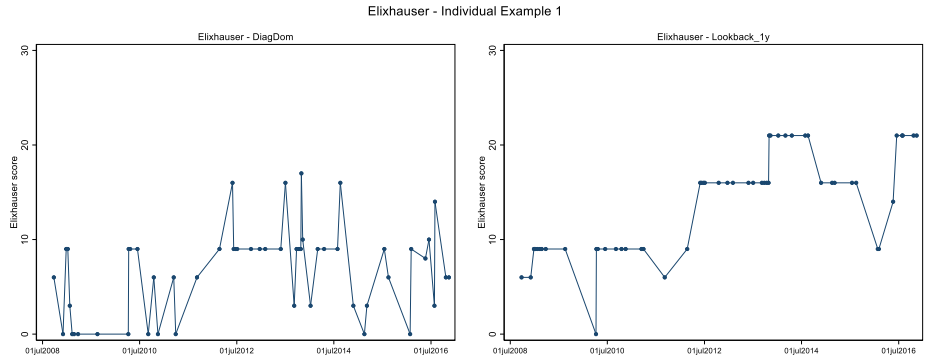
Elixhauser Component	Score weighting	Episodes, n (%)	Individuals, n (%)	Diagnostic-dominant episodes
Alcohol Abuse	0	39,588 (1.8)	16,056 (3.5)	28,107 (2.8)
Drug Abuse	-7	9,043 (0.4)	4,523 (1.0)	6,334 (0.6)
Psychoses	0	7,798 (0.4)	2,620 (0.6)	5,265 (0.5)
Depression	-3	66,447 (3.0)	31,935 (7.0)	45,229 (4.5)
Hypertension, Complicated	0	5,701 (0.3)	3,029 (0.7)	3,216 (0.3)

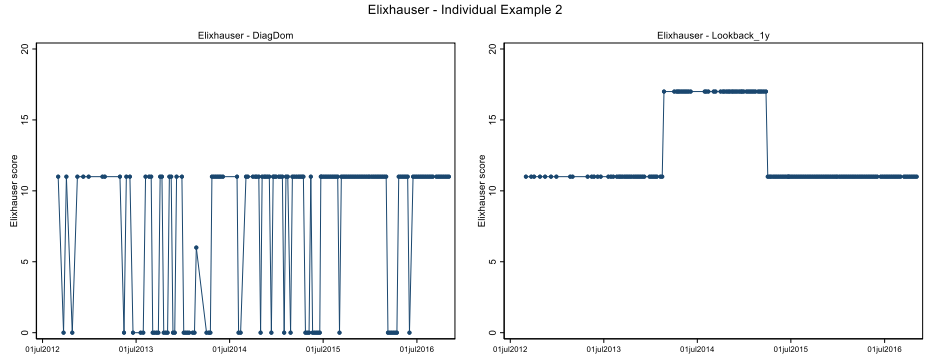
554

Supplementary Figure 1: Examples of individuals' Charlson and Elixhauser scores calculating using the DiagDom and the Lookback_1y method.



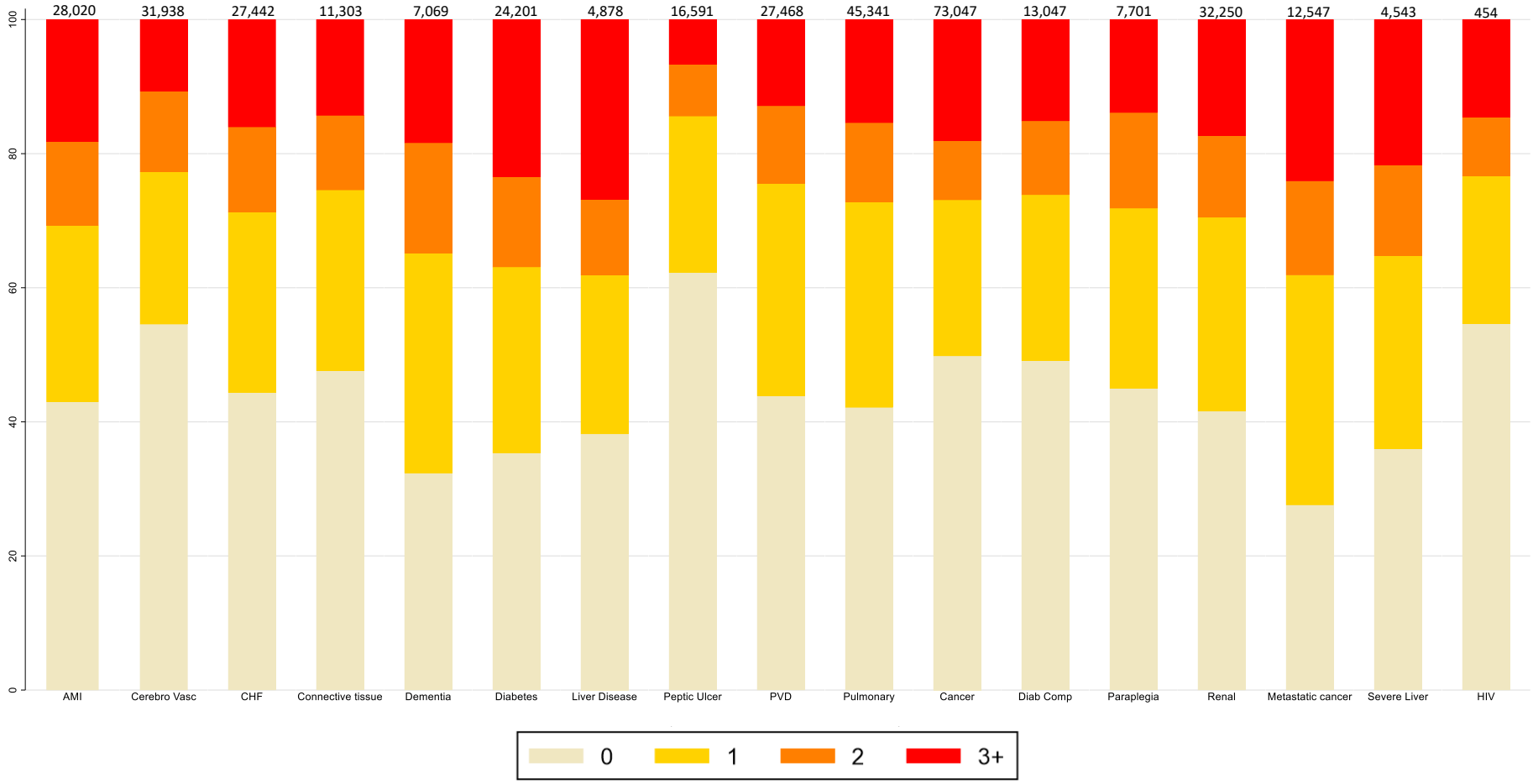




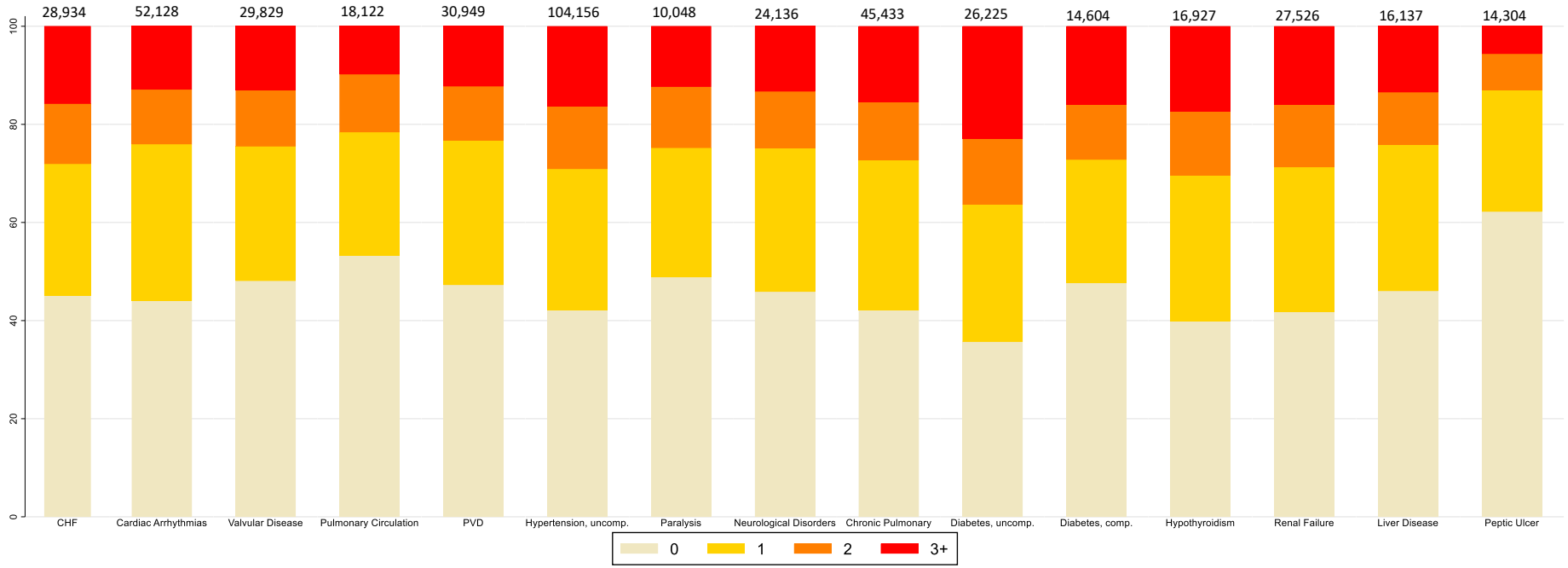


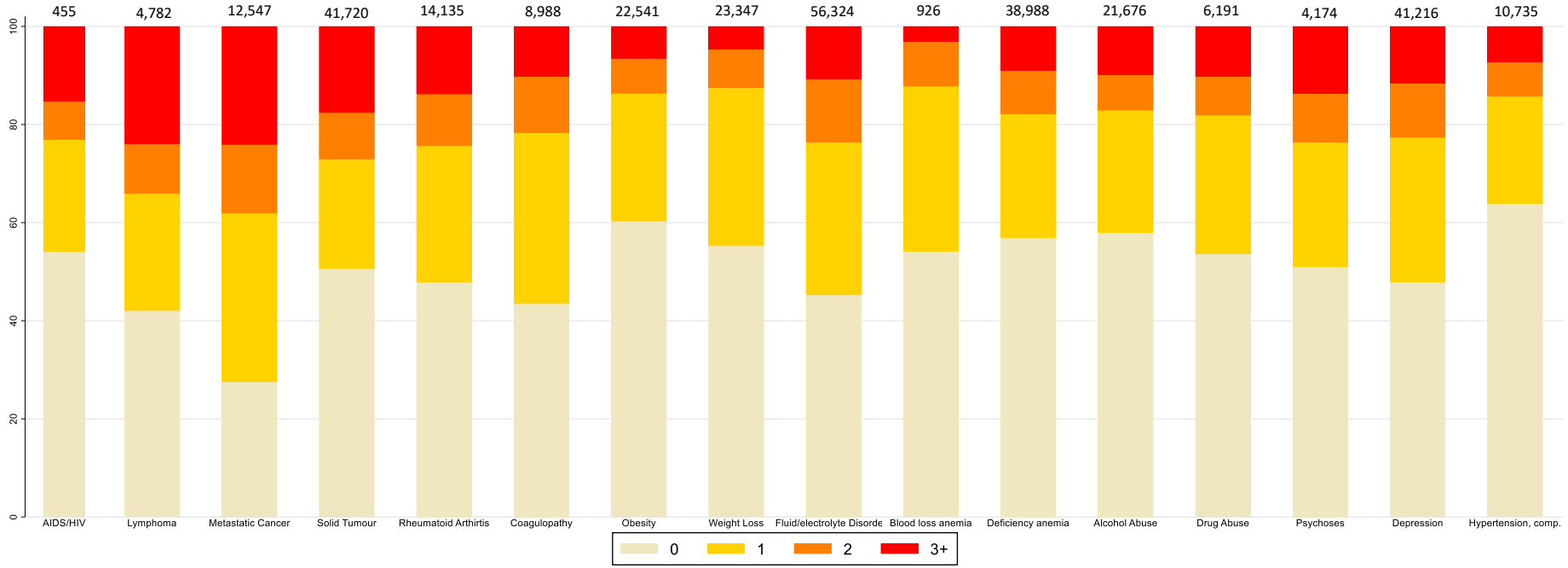
Supplementary Figure 2: Admissions in individuals who did not have a particular comorbiditiy code in the diagnostic-dominant episode but did have at some point earlier in the study period. Colours represent the number of comorbidity codes in the prior year which hence would have been included in the one-year lookback and not in diagnostic-dominant method.

A: Charlson Score



B – Elixhauser Score





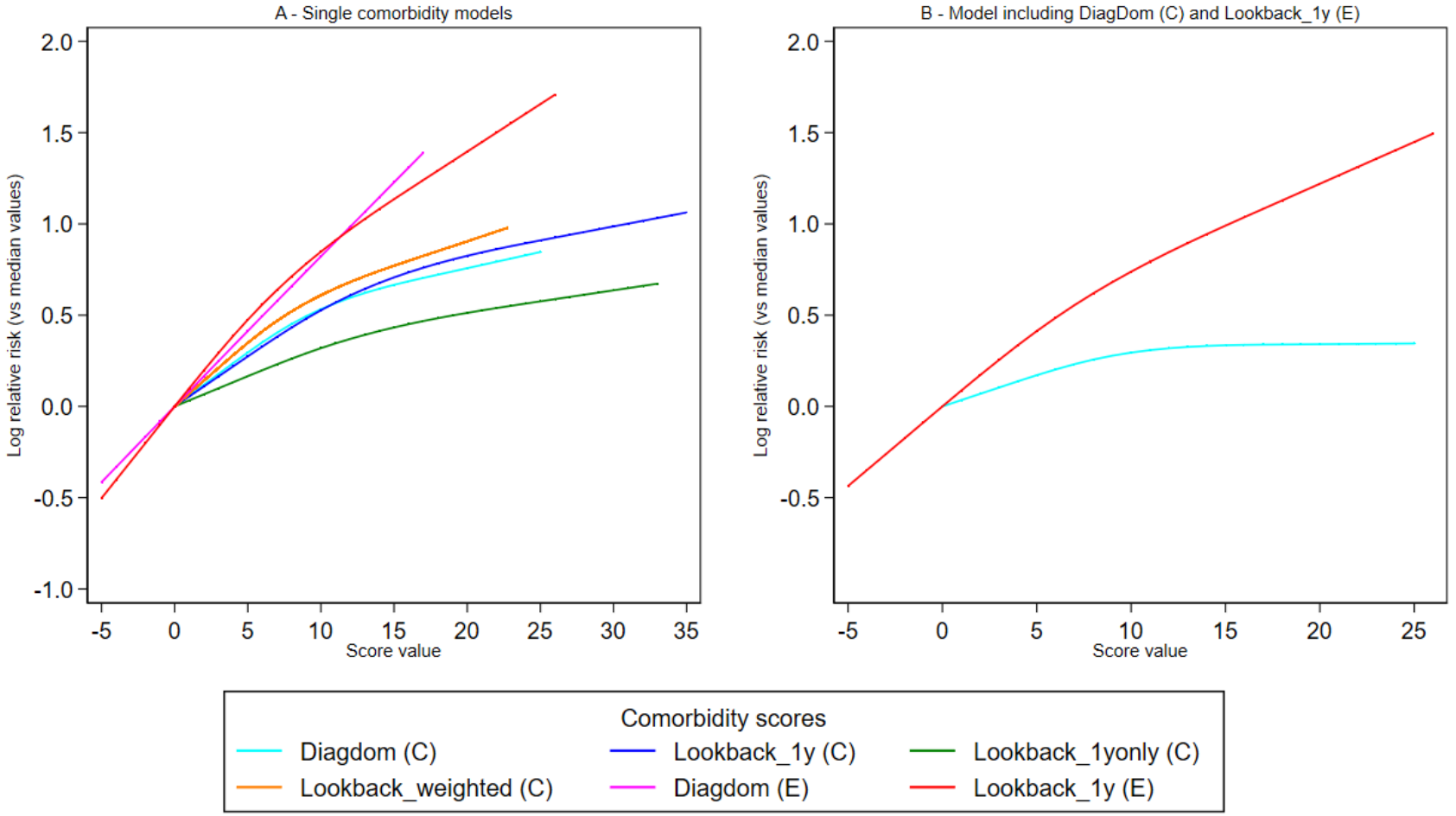
569

570 Note: Numbers on top of bars show the total number of diagnostic dominant admissions with no comorbidity-specific code, but at least one prior use of a

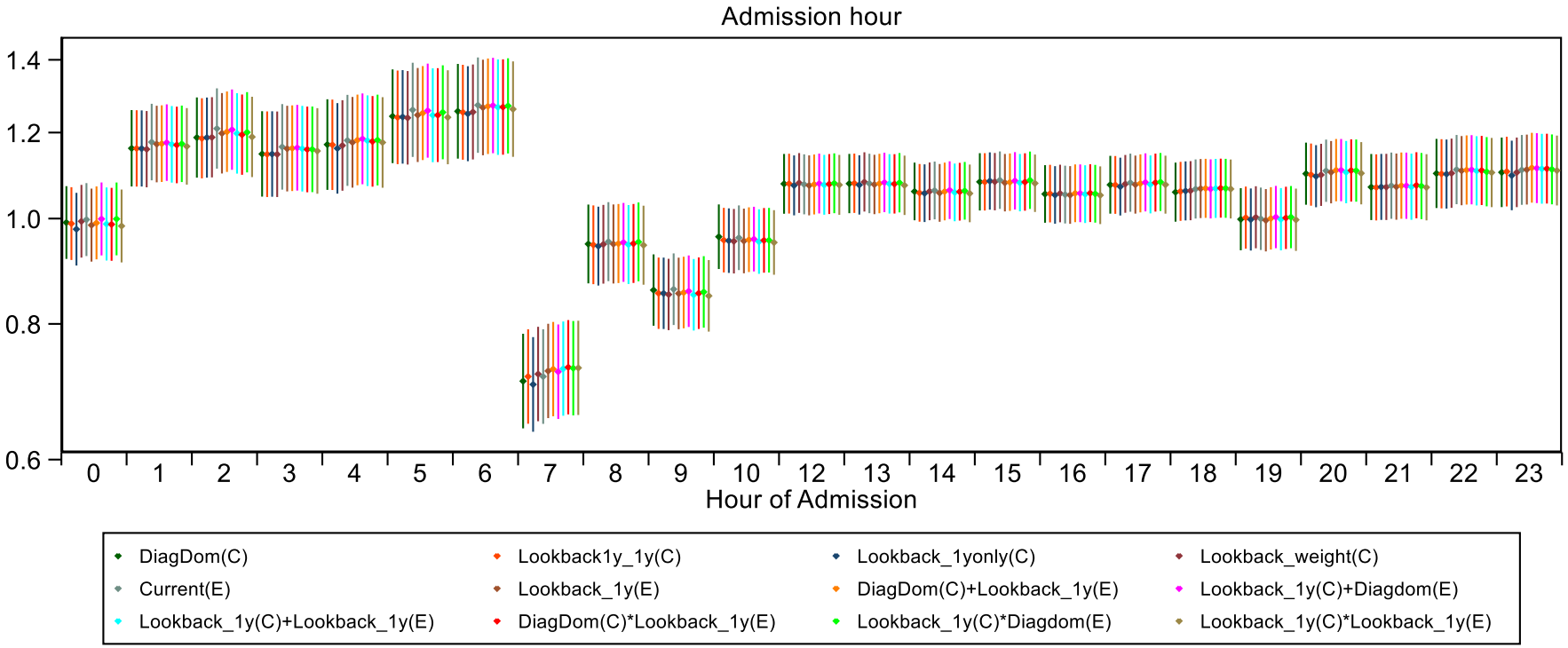
571 relevant code in the study period.

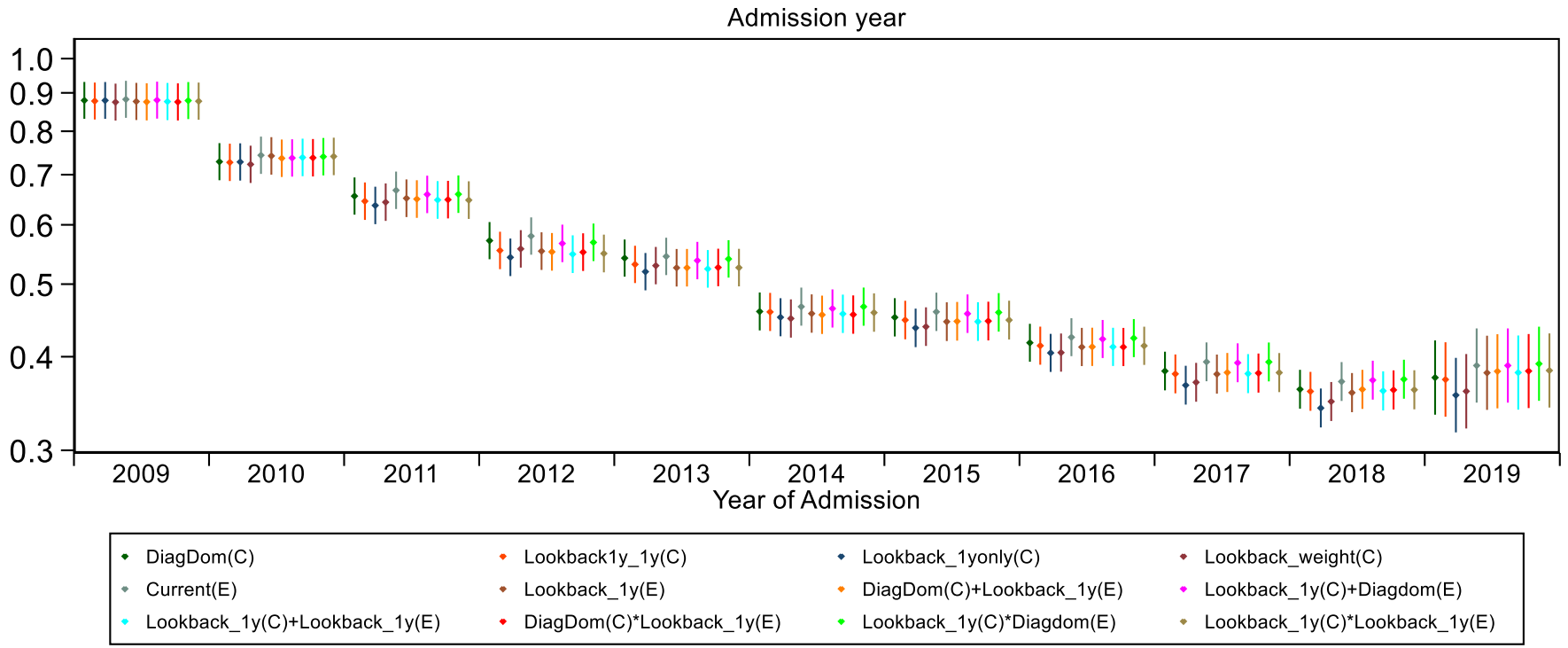
572

Supplementary Figure 3: Adjusted effects of Charlson (C) and Elixhauser (E) scores on 30-day mortality as main effects individually (left), and when the Charlson (DiagDom) and Elixhauser (Lookback_1y) score together as two main effects (right).



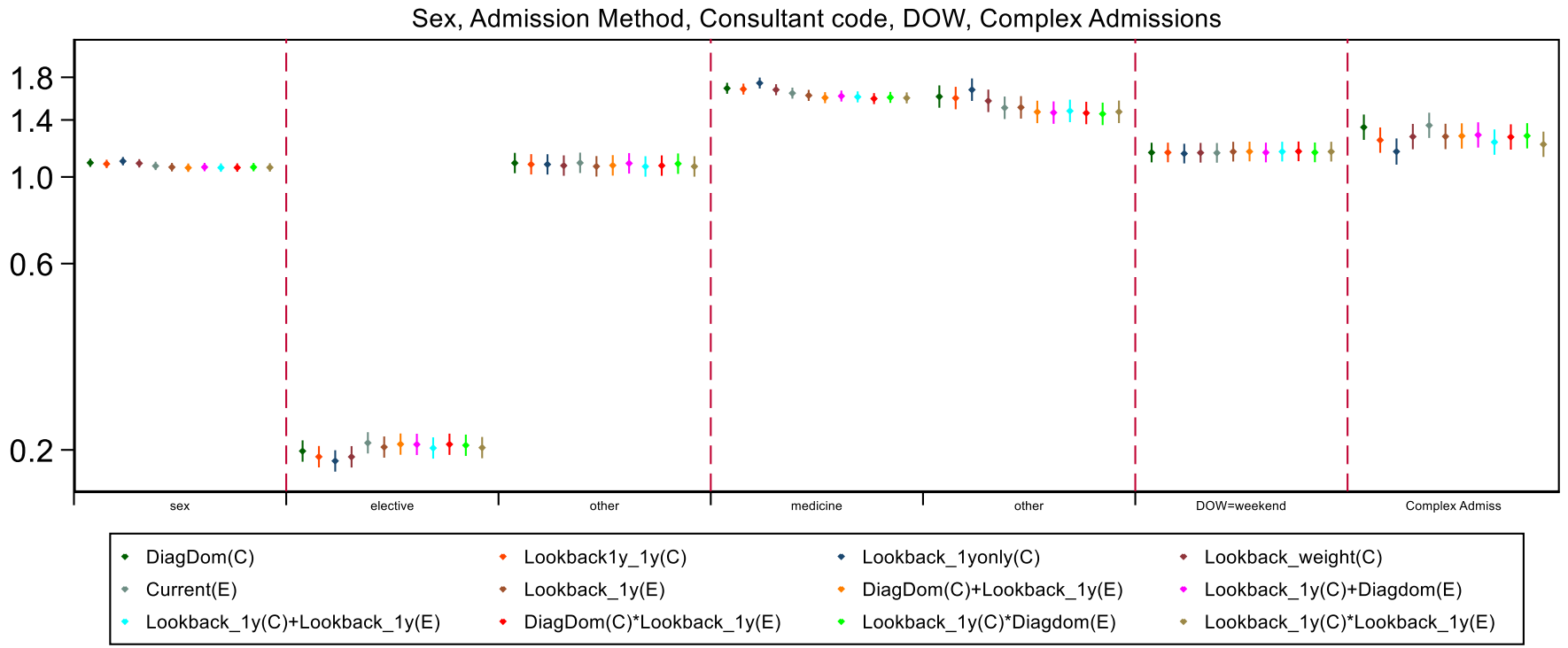
Supplementary Figure 4: Hazard ratios (95% CI) for other variables included in models with splines for comorbidity scores




Admission year
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Year of Admission
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

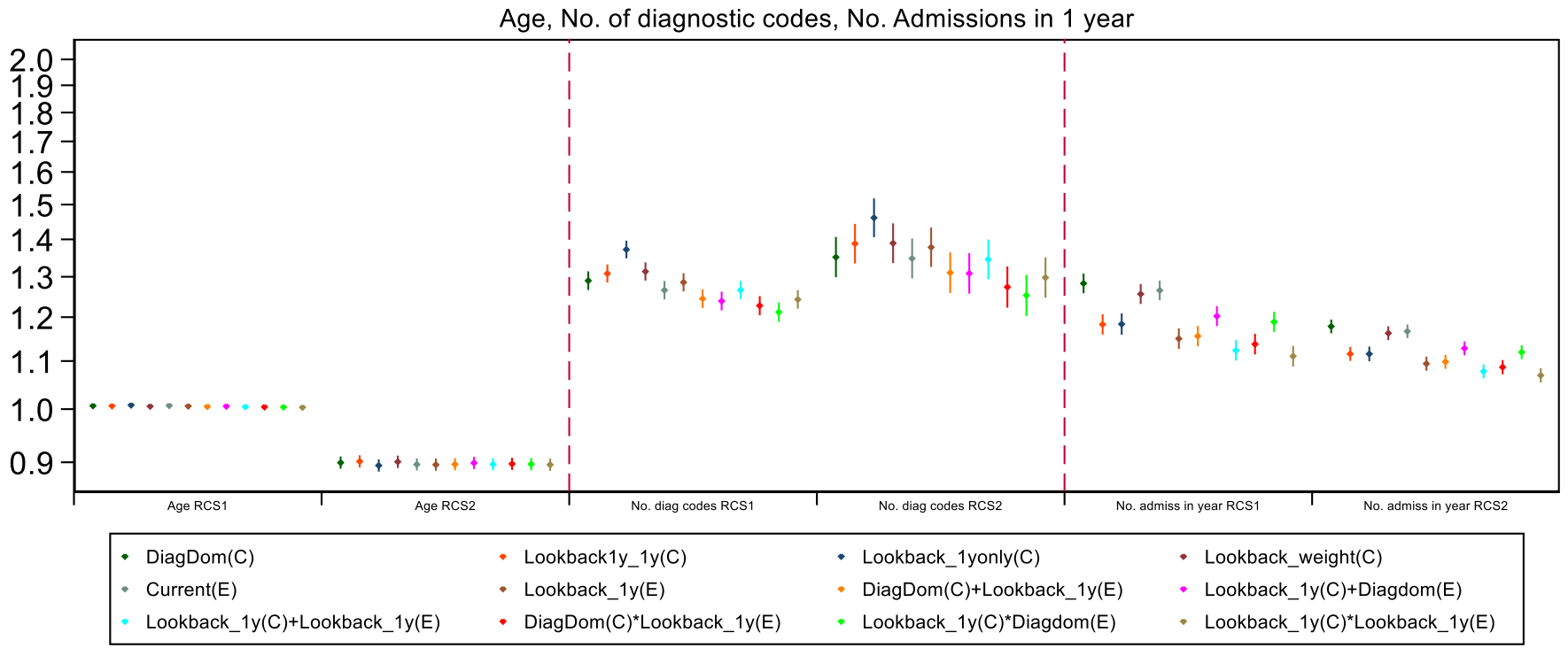
579

580


Sex, Admission Method, Consultant code, DOW, Complex Admissions
1.8
1.4
1.0
0.6
0.2
sex
elective
other
medicine
other
DOW=weekend
Complex Admiss
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

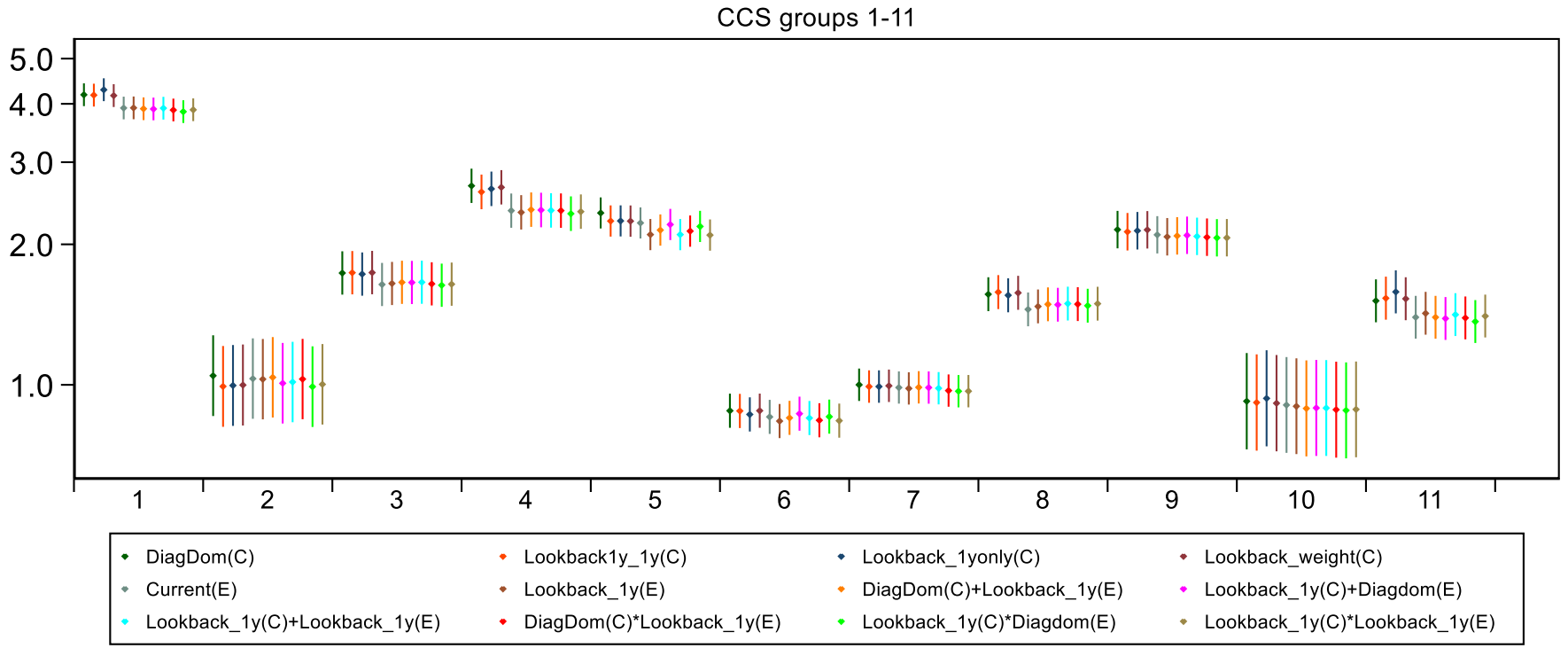
581

582

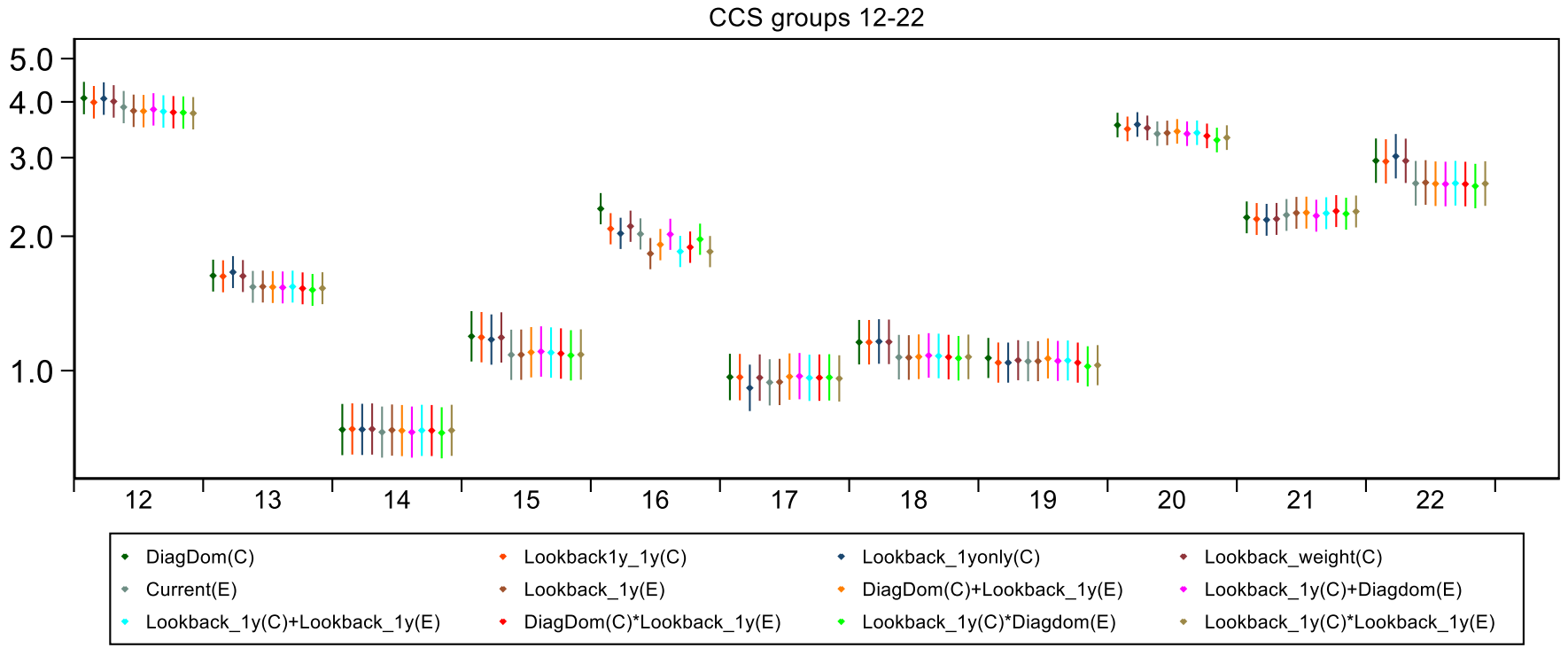

Age, No. of diagnostic codes, No. Admissions in 1 year
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
Age RCS1
Age RCS2
No. diag codes RCS1
No. diag codes RCS2
No. admiss in year RCS1
No. admiss in year RCS2
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

583

584

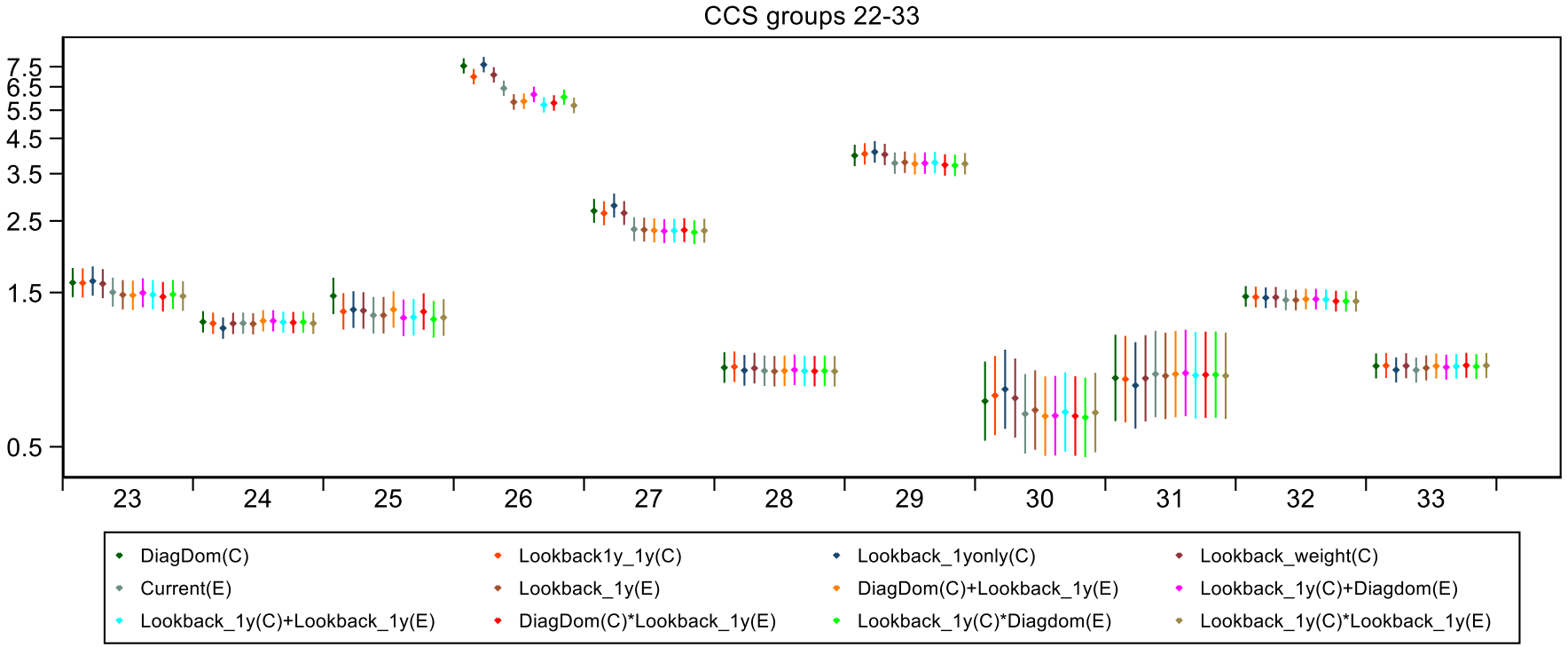

CCS groups 1-11
5.0
4.0
3.0
2.0
1.0
1
2
3
4
5
6
7
8
9
10
11
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

585


CCS groups 12-22
5.0
4.0
3.0
2.0
1.0
12
13
14
15
16
17
18
19
20
21
22
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

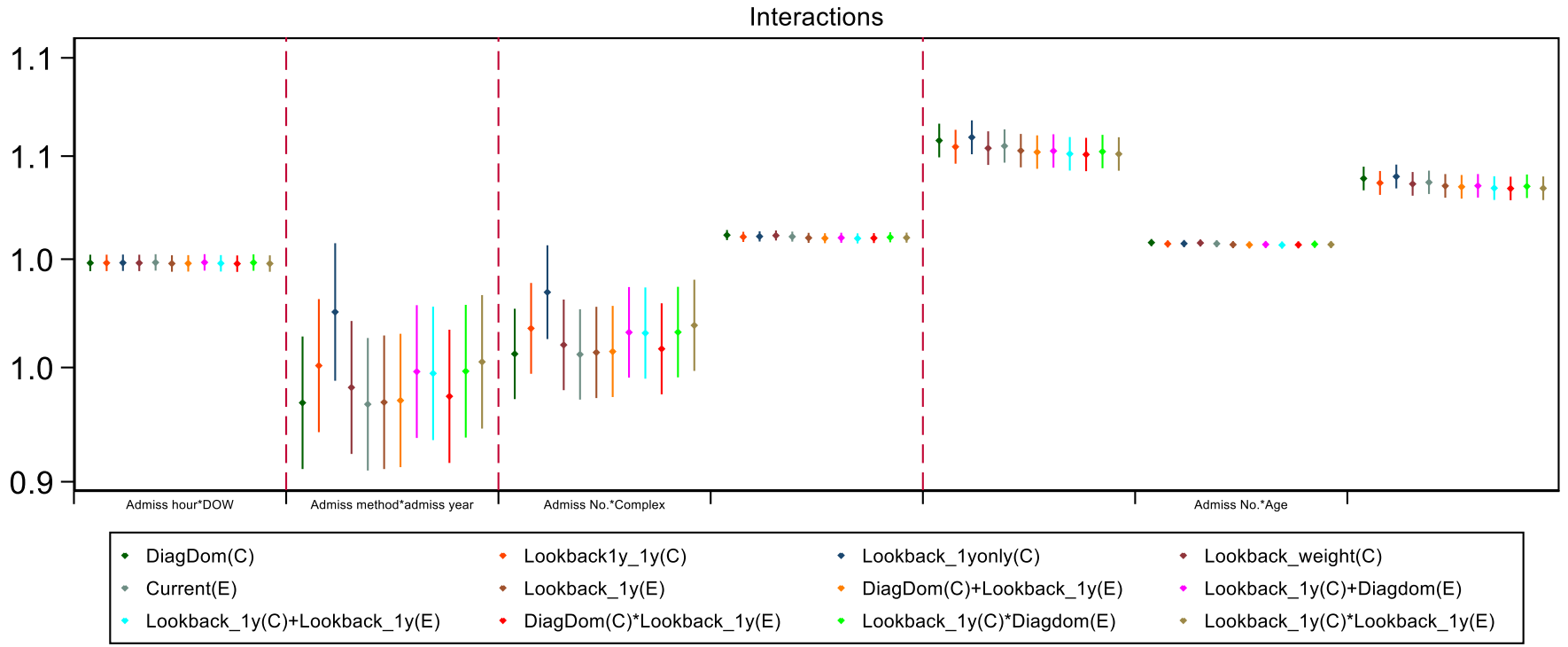
586

587


CCS groups 22-33
7.5
6.5
5.5
4.5
3.5
2.5
1.5
0.5
23
24
25
26
27
28
29
30
31
32
33
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

588

589


Interactions
1.1
1.1
1.0
1.0
0.9
Admiss hour*DOW
Admiss method*admiss year
Admiss No.*Complex
Admiss No.*Age
DiagDom(C)
Lookback1y_1y(C)
Lookback_1yonly(C)
Lookback_weight(C)
Current(E)
Lookback_1y(E)
DiagDom(C)+Lookback_1y(E)
Lookback_1y(C)+Diagdom(E)
Lookback_1y(C)+Lookback_1y(E)
DiagDom(C)*Lookback_1y(E)
Lookback_1y(C)*Diagdom(E)
Lookback_1y(C)*Lookback_1y(E)

590