Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multisystem disease accumulation in routine claims reveals dynamic multimorbidity and mortality risk in older adults: a multicentre retrospective cohort study

Chisato Okamoto
Osaka University Graduate School of Medicine/Frontier Biosciences
Miyuki Oka
ASCLEPIUS Inc
Akiko Kagawa
Canon ITS Medical Inc
Takeshi Aketa
ASCLEPIUS Inc
Yohei Miyashita
Hanwa Memorial Hospital
Masami Yabumoto
Hanwa Memorial Hospital
Shusaku Yano
Independent Researcher
Masafumi Kitakaze
kitakaze@zf6.so-net.ne.jp

Hanwa Memorial Hospital

Research Article

Keywords: multimorbidity, older adults, administrative claims, disease trajectories, mortality, geriatric epidemiology

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11226533/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. MO and TA are employees of ASCLEPIUS Inc., which is contracted by the Kinshukai Medical Corporation to support research data extraction, data management, and research-dataset construction. AK is an employee of Canon ITS Medical Inc., which provided technical advice on electronic-medical-record data structure and extraction specifications. These companies had no role in study design, analysis, interpretation, or the decision to submit. CO, YM, MY, SY, and MK declare no competing interests.

Abstract

Background

Ageing research treats late-life vulnerability as accumulation of heterogeneous deficits, but multimorbidity is commonly measured as a static count. We asked whether short-term change in claims-documented disease burden captures a dynamic dimension of multimorbidity in older adults.

Methods

We conducted a multicentre retrospective cohort study using longitudinal claims and clinical records from 6,642 adults aged 65 years or older across five Japanese hospitals. Among 5,018 who reached a 6-month landmark, we separated Month 0 burden from six-month multisystem net accumulation and care-utilisation intensity. Associations with post-landmark mortality were assessed using Cox models and C-statistics, with facility-held-out validation of trajectory-only and utilisation-enriched models.

Results

The crude proportion with observed death within 12 months after the landmark ranged from 1.9% to 24.6% across low-burden/low-accumulation and high-burden/high-accumulation phenotypes in acute care, and from 7.0% to 26.4% in chronic care. Net accumulation improved discrimination beyond Charlson Comorbidity Index plus baseline burden (Delta C + 0.042 acute; +0.019 chronic). Trajectory-only models preserved risk ordering in held-out hospitals (C = 0.660 acute; 0.746 chronic).

Conclusions

Claims-derived multisystem accumulation captures a dynamic dimension of late-life multimorbidity that static counts miss and may provide a scalable longitudinal phenotype for ageing research when repeated frailty or functional assessments are unavailable.

Background

Ageing is heterogeneous: people of the same chronological age and with similar disease counts can differ substantially in vulnerability and subsequent decline.1–10 The deficit-accumulation framework formalised one aspect of this heterogeneity by treating late-life vulnerability as the cumulative expression of multiple deficits rather than any single disease.4–7 Multimorbidity research likewise shows that disease burden is distributed across organ systems and changes over time.11–18 Yet most clinical and administrative measures still reduce this process to a cross-sectional comorbidity score or disease count.11–21

Static burden therefore conflates state with trajectory. Two older adults can begin with the same documented burden, but one may remain stable while another accumulates new problems across several systems over the following months.11–18,22 The latter may represent a distinct dynamic multimorbidity state even when the starting count is identical. A longitudinal phenotype should therefore separate where disease burden starts from the extent and distribution of subsequent expansion, and test whether that change carries information beyond established comorbidity measures.

Claims data offer an imperfect but scalable window on this process. They do not measure frailty, intrinsic capacity, cognition, activities of daily living, quality of life, or biological ageing,1–7,18 and diagnosis codes reflect care contact, coding, and billing.23 But because claims are generated repeatedly during clinical encounters, they can trace the accumulation and recession of documented morbidity over time. We therefore treat change in claims-documented burden as an administrative phenotype of dynamic multimorbidity, not as a frailty index, biological-age measure, or adjudicated count of incident disease.

We operationalised this idea in five Japanese hospitals by separating Month 0 active disease burden (D0), six-month net accumulation, the breadth and depth of multisystem expansion, and care-utilisation intensity. We tested three linked hypotheses: that subsequent accumulation would stratify mortality among patients with similar starting burden; that the signal would be multisystem rather than organ-specific and would add information beyond the Charlson Comorbidity Index (CCI) and D0; and that a minimal diagnosis-based phenotype would preserve risk ordering when transported without refitting to held-out hospitals.

Methods

Data source and care-setting classification

We used routinely collected longitudinal claims and clinical records from five hospitals in the Kinshukai group in Osaka, Japan (two acute-care, three chronic-care). For each facility, records were available from the first electronically available month; index admissions occurred between February 1989 and January 2026 (71.3% in 2022 or later), and follow-up data were available through June 2026. Data were extracted in July 2026, after ethics approval and institutional permission on 1 July 2026, and analyses were completed in September 2026; under the approved protocol, data collection was permitted through 31 March 2027. Reporting followed STROBE and RECORD principles.23,27 Facilities were classified using the hospital-code master. Acute-care and chronic-care arms were analysed as setting-specific index-admission strata, not as a single tracked transfer pathway. Facility-stratified, facility-adjusted, and facility-held-out analyses examined whether the trajectory signal persisted across materially different case-mix settings.

Cohort construction and landmark design

The index admission was the first observed age-eligible admission at age 65 years or older. Acute-care and chronic-care arms were analysed as setting-specific index-admission strata. A 6-month landmark cohort was defined among patients who remained observable to the landmark with positive post-landmark follow-up.28 This avoided using post-index dynamic predictors before they could be observed, while changing the estimand: inference applies to patients who reached the landmark rather than to all patients at admission. The 6-month window balanced sufficient time for a trajectory to become observable against sufficient remaining follow-up for 12-month post-landmark death. Of 6,642 index-admission patients, 1,624 did not enter the landmark cohort: 872 had death at or before the landmark (132 acute-care, 740 chronic-care) and 752 lacked positive post-landmark follow-up (210 acute-care, 542 chronic-care). Landmark attrition and the descriptive comparison of landmark and non-landmark patients are reported in Supplementary Table S0.

Disease burden and net accumulation

Active disease burden was calculated monthly from claims-documented diagnosis categories, which may include suspected or provisional labels and were not re-adjudicated. D0 burden and net accumulation should therefore be read as claims-documented diagnosis-category burden and change, not confirmed incident disease or severity. D0 burden was the Month 0 active disease-category count and was not the 10-system count. Monthly counts M1 through M6 were used to form the final active disease count. The primary final count was M6 where available; otherwise, the latest available count in the sequence M5, M4, M3, M2, M1 was used. Six-month net active disease accumulation equalled the final active disease count minus D0. Negative values were retained because documented active burden can decrease when active documentation subsides. Among 5,018 landmark patients, a final count was available for 4,771 (95.1%; M6 in 4,170 [83.1%], an earlier month in 601 [12.0%]); 247 (4.9%) lacked a final count and were excluded from net-based complete-case models. Among non-missing patients, 653 (13.7%) had negative, 893 (18.7%) zero, and 3,225 (67.6%) positive net. For descriptive Fig. 3 analyses, all landmark patients were retained, with missing-final-count patients classified as no observed positive net and zero breadth/depth; regression and discrimination analyses used complete cases.

Disease-system involvement

Monthly disease-system counts were taken from the source monthly claims-derived disease-system variables. For manuscript-facing analyses, candidate source columns were restricted to names beginning with diseases_. The 10 reporting domains—cardiovascular, gastrointestinal, metabolic/endocrine, musculoskeletal, neurological, respiratory, renal/urologic, infectious, haematological, and oncological—were assigned by case-insensitive substring matching to prespecified source-label keywords (Supplementary Table S14); the first available matching source column was used for each domain. This step harmonised preprocessed source labels and was not a new patient-level ICD-10 regrouping. System-specific D0, final, and net values used the same final-available-month rule as total disease burden. Newly documented system involvement was defined as no involvement at D0 with

positive final-period involvement; increased involvement was defined as positive system-specific net change.

Charlson Comorbidity Index

The Charlson Comorbidity Index (CCI)19–21 was derived from ICD-10 diagnosis codes over the index-to-landmark window using the Quan algorithm19 and original Charlson weights.21 Seventeen comorbidity categories were evaluated: myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatic disease, peptic ulcer disease, mild liver disease, diabetes without complications, diabetes with complications, hemiplegia or paraplegia, renal disease, any malignancy, moderate or severe liver disease, metastatic solid tumour, and AIDS/HIV. Original weights of 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 3, 6, and 6 were applied, respectively. Hierarchical suppression prevented double counting: diabetes with complications suppressed diabetes without complications; moderate or severe liver disease suppressed mild liver disease; and metastatic solid tumour suppressed malignancy. ICD-10 codes were normalised by removing non-alphanumeric characters and matched to condition-specific prefixes using the month-year field as the event date. Patients with no qualifying ICD-10 codes in the index-to-landmark window were assigned CCI = 0.

Care-utilisation intensity

Utilisation intensity was constructed separately from disease burden. For each patient, M0-M6 components were summed, log1p-transformed, z-standardised, and averaged across available components. The primary score included hospitalisation-related signals, surgeries/procedures, rehabilitation units, imaging, medications, injections, and billing-line intensity; the resulting average was z-standardised again for modelling. Cost/points variables were excluded from the primary score and examined only in sensitivity analyses. The score was interpreted as an enriched care-activity layer, not disease burden or a causal treatment effect.

Outcomes

The primary endpoint was an observed death recorded within 12 months after the landmark. For descriptive tables and heatmaps, this endpoint is presented as a crude landmark-cohort proportion: patients last observed alive before 12 months remain in the denominator and are not counted as deaths, so these percentages are not Kaplan-Meier estimates of 12-month mortality risk. Deaths were ascertained from death-at-discharge records of the participating hospitals; deaths occurring after a patient left the hospital group were not captured, and such patients were censored at their last observed care-system contact. Time-to-event analyses used observed post-landmark time from the landmark to terminal death or last observed care-system contact, with survival displays truncated at 12 months. Supportive endpoints were all-cause mortality over observed post-landmark follow-up and early care-system exit (absence of prolonged observed care-system presence), defined as death at any time after

the landmark or observed post-landmark care-system time shorter than 18 months; patients alive with at least 18 months of observed care-system time were classified as non-events.

Statistics and reproducibility

Analyses were conducted separately by care-setting arm and followed the figure sequence. Model hierarchies included demographics, CCI, D0 burden, cardiovascular-only net, respiratory-only net, total multisystem net, utilisation intensity, and combined claims-trajectory models. We evaluated a minimal diagnosis-based model (age, sex, CCI, D0, net) and an enriched model adding utilisation. Spearman rank correlation was used for monotone associations between continuous trajectory measures, and global log-rank tests compared the displayed Kaplan-Meier strata; these non-parametric procedures do not require normally distributed data. Cox proportional-hazards models were used for censored time-to-event modelling and were fitted with a ridge (L2) penalty of 0.01 to stabilise coefficient estimation, except for the single-component models that added one disease-system-specific net variable or one utilisation component to the reference model (cardiovascular-only and respiratory-only rows in Table 3 Section B; Tables S3 and S4; Figures S3–S5), which were fitted with a minimal ridge penalty of 0.00001. The principal Cox-model claims concern discrimination and risk ordering rather than inference on individual coefficients; formal Schoenfeld-residual proportional-hazards diagnostics were not part of the prespecified workflow. Incremental discrimination was summarised with common-N C-statistics and paired Delta C contrasts based on 1,000 bootstrap iterations. All statistical tests were two-sided with a nominal alpha level of 0.05. No multiplicity adjustment was applied; P values from organ-specific, utilisation-component, subgroup, and sensitivity analyses should therefore be interpreted as descriptive/exploratory rather than as a family of confirmatory tests. Model- and comparison-specific N and event counts are reported in the corresponding tables or figure panels. Reported P values for main-table C-statistics and paired Delta C contrasts are descriptive two-sided bootstrap values with plus-one smoothing; with 1,000 replicates, the smallest attainable value is 0.002, so P = 0.002 denotes P ≤ 0.002. Supplementary component analyses in Tables S3 and S4 use unsmoothed bootstrap proportions and can therefore fall below this bound. Except for the facility-held-out validation estimates, reported C-statistics are apparent estimates obtained in the sample in which the model was fitted; optimism-corrected internal-validation estimates were not computed. Incremental apparent Delta C values should therefore be read as upper bounds on added information. Reporting was informed by STROBE, RECORD, TRIPOD + AI, and PROGRESS.23,27,29,30

Breadth and depth analyses

Breadth was defined as the number of newly involved disease systems during M0-M6, and depth as the maximum disease-name net gain within any single system. Figure 3 displays breadth and depth as approximately balanced tertile Kaplan-Meier groups. Supplementary Table S11 and Supplementary Figure S9 compare total net accumulation with breadth alone, depth alone, and breadth plus depth in

simplified CCI- and D0-adjusted models that intentionally omit age and sex to isolate the relative information content of the morbidity-expansion components.

Facility-based split validation

Facility-based split validation tested whether risk ordering was preserved across participating hospitals. For acute care, anonymised facility A was the derivation facility and B the held-out validation facility; for chronic care, facilities C and D were derivation facilities and E the held-out validation facility. The trajectory-only model included age at index, male sex, CCI, D0 burden, and six-month net accumulation; the utilisation-enriched model added the M0-M6 utilisation-intensity score. Cox beta coefficients and derivation-sample linear-predictor tertile cutoffs were applied to validation patients without refitting. Kaplan-Meier curves were plotted only in validation facilities for transported risk groups. Validation C-statistics were bootstrapped 1,000 times within the held-out cohort with transported predictors and cutoffs fixed, and percentile 95% CIs were calculated. Continuous predictors were standardised once across the pooled landmark cohort rather than within derivation facilities; consequently, the standardisation constants, but not Cox coefficients or derivation tertile cutoffs, were informed by held-out facilities. The analysis therefore provides facility-held-out transport validation within one hospital group rather than independent external validation or proof of absolute-risk calibration.

Sensitivity analyses

Prespecified sensitivity analyses examined alternative net definitions, facility-adjusted models, alternative utilisation-score definitions, supportive outcomes, calibration deciles, breadth/depth component models, 3-, 6-, and 12-month landmark windows, and the contextual full-sample apparent risk-ordering display. Alternative trajectory windows required observed time beyond the corresponding landmark and recalculated dynamic features and 12-month post-landmark death. The 6-month window remained the prespecified primary specification; 3- and 12-month analyses were descriptive robustness checks and were not used to optimise the primary window. The 12-month window may yield higher apparent discrimination because it selects a smaller survivor/observable subset. The full-sample apparent risk-ordering display is shown in Supplementary Figure S10 for descriptive context only; Main Fig. 6 and Supplementary Table S12 provide the primary facility-held-out evidence.

Use of generative AI assisted tools

OpenAI ChatGPT (GPT-5.5 and GPT-5.6; accessed July–September 2026) was used to assist with drafting and debugging analysis code and with English-language editing. AI-assisted code was reviewed line by line by the authors, executed only in a local computing environment, and checked against expected data structures and statistical outputs. The tools made no autonomous analytical or interpretive decisions, and no patient-level or directly identifying data were entered into external AI systems.

Results

Cohort and landmark population

Among 6,642 patients, 5,018 reached the 6-month landmark (896 acute-care, 4,122 chronic-care). The 1,624 non-landmark patients (872 deaths at or before the landmark; 752 without positive follow-up) fell outside the primary estimand and were older, more often male, and higher in D0 burden and CCI, consistent with the landmark selection. Observed death within 12 months after the landmark was recorded in 132 acute-care (14.7%) and 811 chronic-care patients (19.7%); these percentages are crude landmark-cohort proportions rather than Kaplan-Meier 12-month risk estimates (Table 1; Fig. 1A).

Table 1
Baseline and landmark cohort characteristics by care setting.

Characteristic	Total (N = 5,018)	Acute care (N = 896)	Chronic care (N = 4,122)
Demographics and comorbidity			
Age at index, years, median [IQR]	83.7 [77.5–89.0]	81.4 [74.5–87.6]	84.1 [78.2–89.3]
Male sex, n (%)	1,687 (33.6)	355 (39.6)	1,332 (32.3)
Charlson Comorbidity Index, median [IQR]	2.0 [1.0–4.0]	2.0 [0.0–3.0]	2.0 [1.0–4.0]
Six-month dynamic multimorbidity features			
D0 burden, median [IQR]	8.0 [5.0–11.0]	6.0 [2.0–10.0]	8.0 [5.0–11.0]
Six-month net accumulation, median [IQR]	2.0 [0.0–5.0]	3.0 [0.0–7.0]	2.0 [0.0–4.0]
Net < 0 (negative), n (%)	653 (13.0)	60 (6.7)	593 (14.4)
Net = 0 (zero), n (%)	893 (17.8)	187 (20.9)	706 (17.1)
Net > 0 (positive), n (%)	3,225 (64.3)	582 (65.0)	2,643 (64.1)
Net missing, n (%)	247 (4.9)	67 (7.5)	180 (4.4)
M0–M6 utilisation intensity score, median [IQR]	0.0 [− 0.4–0.3]	0.2 [− 0.5–0.5]	0.0 [− 0.4–0.3]
Primary endpoint and follow-up			
Death within 12 months after landmark, n (%)	943 (18.8)	132 (14.7)	811 (19.7)
Observed post-landmark care-system time, months, median [IQR]	19.0 [6.0–50.0]	21.0 [6.0–68.0]	19.0 [6.0–47.0]

D0 = Month 0 active disease burden. Net accumulation = six-month net active disease accumulation (final available M1–M6 count minus D0). Net distribution rows (Net < 0 / = 0 / > 0 / missing) are expressed as n (%) of the arm landmark-cohort N shown in the column header; the common-N analysis set (Acute N = 829, Chronic N = 3,942) excludes patients with missing net accumulation. Utilisation intensity score = standardised composite of seven M0–M6 care-activity components. IQR = interquartile range. Death within 12 months after landmark is the crude proportion with an observed terminal death recorded within 12 months; patients last observed alive before 12 months remain in the denominator, so this percentage is not a Kaplan-Meier 12-month risk estimate. Deaths were ascertained from death-at-discharge records of the participating hospitals; deaths after leaving the hospital group were not captured. Net-distribution percentages differ from those in the Methods and Supplementary Materials, which express the same category counts as a proportion of patients with a non-missing final count (N = 4,771; e.g., negative net 653/5,018 = 13.0% vs 653/4,771 = 13.7%).

Claims captured short-term change in documented disease burden

The representative trajectory in Fig. 1B shows the phenotype logic: D0 read at Month 0, net accumulation as final active count minus D0, and utilisation as a parallel signal. The measure is shown before modelling; the question is not whether a black-box score predicts death but whether routine claims reveal a legible longitudinal process in the first 6 months.

New disease-system involvement was multisystem, not cardiovascular-dominant

Newly documented involvement during M0–M6 was multisystem, not cardiovascular-dominant (Fig. 2A). In acute care it was most frequent for gastrointestinal (21.8%), infectious (15.3%), respiratory (14.7%), renal/urologic (14.6%), musculoskeletal (14.0%), metabolic/endocrine (11.9%), neurological (11.4%), haematological (8.3%), cardiovascular (6.8%), and oncological (4.1%) systems; chronic-care values were 12.8%, 10.9%, 10.4%, 10.4%, 12.2%, 8.2%, 7.1%, 4.8%, 4.7%, and 1.9%.

Net-accumulation strata showed graded expansion across systems

Net-accumulation strata showed graded multisystem expansion (Fig. 2B, 2C). In acute care the ≥ 3 stratum (n = 433) showed high increased involvement across gastrointestinal (65%), cardiovascular (57%), and metabolic/endocrine (40%) systems; in chronic care the ≥ 3 stratum (n = 1,599) spanned gastrointestinal (52%), musculoskeletal (41%), cardiovascular (39%), metabolic/endocrine (36%), respiratory (32%), and renal/urologic (30%). The gradient indicates net accumulation reflects broader system involvement rather than a numerical artifact.

D0 burden and net accumulation defined mortality phenotypes

Cross-classification by D0 and net showed clear mortality gradients. In acute care, 12-month death rose from 1.9% (low-D0/low-net) to 15.8% (low-D0/high-net), 8.6% (high-D0/low-net), and 24.6% (high-D0/high-net); chronic-care values were 7.0%, 21.8%, 19.0%, and 26.4%. Net accumulation added a dynamic dimension beyond starting burden (Table 2; Fig. 3A, B).

Table 2
D0 burden and six-month net accumulation phenotypes.

D0 stratum	Net accumulation stratum	Acute-care arm (N = 896)			Chronic-care arm (N = 4,122)		
		N	12-month death, %	Observed time, months, median	N	12-month death, %	Observed time, months, median
Low D0	Low net	209	1.9	78.0	869	7.0	51.0
	High net	273	15.8	14.0	1,381	21.8	17.0
High D0	Low net	105	8.6	32.0	610	19.0	24.0
	High net	*309*	*24.6*	*10.0*	*1,262*	*26.4*	*10.0*

D0 burden dichotomised at the arm-specific median (acute < = 6 = low D0; chronic < = 8 = low D0). Net accumulation dichotomised at 0 (< = 0 = low net; >0 = high net); patients without an available M1-M6 final disease count were retained in the low-net group for this descriptive table. The 12-month death percentages are crude observed-death proportions as defined in Table 1 *and are not Kaplan-Meier 12-month risk estimates. Shading intensity reflects the approximate observed-death gradient. Observed time = median months from 6-month landmark to terminal death or last observed care-system contact.*

Six-month net accumulation reflects both breadth and depth of morbidity expansion

Net accumulation was decomposed into breadth (number of newly involved systems) and depth (maximum net gain within a single system). Both correlated with total net (Spearman ρ = 0.67 breadth, 0.76 depth; Fig. 3C, 3D), and breadth/depth tertiles showed graded post-landmark survival separation over 12 months in both arms (global two-sided log-rank $P < 0.001$ for each comparison; Fig. 3E–H). Net accumulation thus integrates lateral spread across systems and intensification within them; subsequent analyses use total net as the aggregate measure.

Table 3
Added prognostic information of multisystem net accumulation.

Model	Acute care (N = 829, events = 132)			Chronic care (N = 3,942, events = 807)		
	C-statistic	ΔC (95% CI)	p	C-statistic	ΔC (95% CI)	p
A. Net accumulation compared with D0 burden alone						
D0 alone	0.651	—	—	0.664	—	—
D0 + cardiovascular net	0.656	+ 0.005 (− 0.001, + 0.032)	0.27	0.667	+ 0.002 (0.000, + 0.006)	0.05
D0 + respiratory net	0.690	+ 0.040 (+ 0.010, + 0.077)	0.006*	0.690	+ 0.026 (+ 0.014, + 0.037)	0.002*
D0 + total multisystem net	0.718	+ 0.067 (+ 0.030, + 0.114)	0.002*	0.691	+ 0.027 (+ 0.018, + 0.038)	0.002*
B. Net accumulation compared with CCI + D0						
CCI alone	0.667 (0.626–0.715)	—	—	0.675 (0.656–0.693)	—	—
CCI + D0	0.678 (0.637–0.721)	+ 0.011	0.002*	0.677 (0.662–0.697)	+ 0.003	0.002*
CCI + D0 + cardiovascular net	0.677	−0.000 (− 0.001, + 0.001)	0.61	0.678	+ 0.001 (− 0.001, + 0.002)	0.48
CCI + D0 + respiratory net	0.702	+ 0.025 (− 0.004, + 0.055)	0.10	0.698	+ 0.021 (+ 0.012, + 0.031)	0.002*
CCI + D0 + total multisystem net (primary)	0.720 (0.688–0.762)	+ 0.042 (+ 0.014, + 0.084)	0.002*	0.696 (0.681–0.714)	+ 0.019 (+ 0.011, + 0.029)	0.002*

Model	Acute care (N = 829, events = 132)			Chronic care (N = 3,942, events = 807)		
	C-statistic	ΔC (95% CI)	p	C-statistic	ΔC (95% CI)	p
A. Net accumulation compared with D0 burden alone						
*All models in Sections A and B include age at index and sex. Section A: reference model = age + sex + D0 burden (labelled "D0 alone"); N = 829 acute, N = 3,942 chronic. Section B: reference model = age + sex + CCI + D0. In Section B, the C-statistics and ΔC values for the cardiovascular-only and respiratory-only net rows come from the disease-system component analysis (Table S3), which used the same patients and events but fitted all models with a minimal ridge penalty of 0.00001 (reference CCI + D0 C = 0.677 in both arms); they are therefore not exactly equal to differences from the CCI + D0 C-statistic shown above. ΔC = paired change in C-statistic; 95% CI from 1,000 bootstrap iterations. Apart from these two rows, all Cox models were fitted with a ridge (L2) penalty of 0.01, and C-statistics are apparent estimates obtained in the sample in which the model was fitted. P values are descriptive two-sided bootstrap values with plus-one smoothing; the smallest attainable value with 1,000 replicates is 0.002, so a reported P of 0.002 denotes P ≤ 0.002. A nominal alpha level of 0.05 was used for descriptive significance marking; no multiplicity adjustment was applied, and secondary/component comparisons are exploratory. * P < 0.05. Complete-case analyses exclude the 247 landmark patients (67 acute-care, 180 chronic-care) without a non-missing final active-disease count; 12-month death counts in these complete-case sets (132 acute-care, 807 chronic-care) may differ from full landmark-cohort counts (132 acute-care, 811 chronic-care).*						

Multisystem net accumulation added information beyond static and organ-specific comparators

In acute care, D0 + total net reached C = 0.718 and CCI + D0 + total net C = 0.720 (+ 0.042 beyond CCI + D0). Cardiovascular-only net added little beyond D0 (ΔC + 0.005), respiratory net more (+ 0.040), and total net most (+ 0.067). In chronic care, CCI + D0 + total net reached C = 0.696 (+ 0.019 beyond CCI + D0); respiratory and total net were similar beyond D0 (+ 0.026 and + 0.027) and cardiovascular-only net minimal (+ 0.002). The dynamic signal is multisystem, not a cardiovascular proxy (Table 3; Fig. 4).

Net accumulation and utilisation intensity were related, supporting a minimal/enriched interpretation

Net accumulation correlated moderately with utilisation intensity (Spearman ρ 0.55 acute, 0.45 chronic). Adding net to CCI + D0 improved discrimination by + 0.042 (acute) and + 0.019 (chronic); adding utilisation beyond CCI + D0 + net improved it further by + 0.034 and + 0.021. The combined CCI + D0 + net + utilisation model reached C = 0.755 (acute) and C = 0.718 (chronic). The incremental value of net after utilisation was small, especially in acute care, indicating overlap—but overlap that is clinically informative, since accumulation partly captures the same compressed trajectory that appears as intensified utilisation. Utilisation is thus an enriched layer rather than a substitute (Table 4; Fig. 5).

Table 4

Net accumulation, utilisation, and internal discrimination of claims-trajectory models.

Model / contrast	Acute care (N = 829, events = 132)			Chronic care (N = 3,942, events = 807)		
	C-statistic (95% CI)	ΔC (95% CI)	p	C-statistic (95% CI)	ΔC (95% CI)	p
A. Correlation between net accumulation and utilisation intensity						
Spearman ρ (net vs utilisation)	ρ = 0.55, p < 0.001			ρ = 0.45, p < 0.001		
B. Sequential model hierarchy (common-N analysis)						
CCI + D0	0.678 (0.637–0.721)	—	—	0.677 (0.662–0.697)	—	—
CCI + D0 + net accumulation	0.720 (0.688–0.762)	+ 0.042 (+ 0.014, + 0.084)	0.002*	0.696 (0.681–0.714)	+ 0.019 (+ 0.011, + 0.029)	0.002*
CCI + D0 + utilisation	0.750 (0.718–0.789)	+ 0.072 (+ 0.039, + 0.106)	0.002*	0.715 (0.701–0.732)	+ 0.037 (+ 0.028, + 0.048)	0.002*
CCI + D0 + net + utilisation (combined model)	0.755 (0.723–0.793)	—	—	0.718 (0.703–0.736)	—	—
C. Incremental discrimination contrasts						
+ net beyond CCI + D0	—	+ 0.042 (+ 0.014, + 0.084)	0.002*	—	+ 0.019 (+ 0.011, + 0.029)	0.002*
+ utilisation beyond CCI + D0	—	+ 0.072 (+ 0.039, + 0.106)	0.002*	—	+ 0.037 (+ 0.028, + 0.048)	0.002*
+ utilisation beyond CCI + D0 + net	—	+ 0.034 (+ 0.012, + 0.060)	0.002*	—	+ 0.021 (+ 0.014, + 0.031)	0.002*

*Common-N analyses: Acute N = 829, Chronic N = 3,942. All models include sex and age at index. ΔC = paired C-statistic change; 95% CI from 1,000 bootstrap iterations. All Cox models were fitted with a ridge (L2) penalty of 0.01, and C-statistics are apparent estimates obtained in the sample in which the model was fitted. P values are descriptive two-sided bootstrap values with plus-one smoothing; the smallest attainable value with 1,000 replicates is 0.002, so a reported P of 0.002 denotes P ≤ 0.002. A nominal alpha level of 0.05 was used for descriptive significance marking; no multiplicity adjustment was applied, and secondary/component comparisons are exploratory. * P < 0.05. Complete-case models exclude the 247 patients (67 acute-care, 180 chronic-care) lacking a non-missing final active-disease count (complete-case 12-month deaths 132 acute-care and 807 chronic-care, vs 132 and 811 in the full landmark cohort).*

Model / contrast	Acute care (N = 829, events = 132)			Chronic care (N = 3,942, events = 807)		
	C-statistic (95% CI)	ΔC (95% CI)	p	C-statistic (95% CI)	ΔC (95% CI)	p
A. Correlation between net accumulation and utilisation intensity						
+ net beyond CCI + D0 + utilisation	—	+ 0.005 (− 0.001, + 0.018)	0.17	—	+ 0.003 (+ 0.000, + 0.007)	0.008*

*Common-N analyses: Acute N = 829, Chronic N = 3,942. All models include sex and age at index. ΔC = paired C-statistic change; 95% CI from 1,000 bootstrap iterations. All Cox models were fitted with a ridge (L2) penalty of 0.01, and C-statistics are apparent estimates obtained in the sample in which the model was fitted. P values are descriptive two-sided bootstrap values with plus-one smoothing; the smallest attainable value with 1,000 replicates is 0.002, so a reported P of 0.002 denotes P ≤ 0.002. A nominal alpha level of 0.05 was used for descriptive significance marking; no multiplicity adjustment was applied, and secondary/component comparisons are exploratory. * P < 0.05. Complete-case models exclude the 247 patients (67 acute-care, 180 chronic-care) lacking a non-missing final active-disease count (complete-case 12-month deaths 132 acute-care and 807 chronic-care, vs 132 and 811 in the full landmark cohort).*

Facility-based split validation supported transported risk ordering

Figure 6 and Table 5 show facility-held-out validation with bootstrap 95% CIs. In acute care (derivation A: 540 patients, 88 events; validation B: 289 patients, 44 events), the trajectory-only model reached C = 0.762 in derivation and 0.660 in validation (95% CI 0.579–0.739), and the utilisation-enriched model 0.779 and 0.693 (0.624–0.764). In chronic care (derivation C + D: 3,245 patients, 694 events; validation E: 697 patients, 113 events), the trajectory-only model reached C = 0.692 and 0.746 (0.706–0.786) and the utilisation-enriched model 0.714 and 0.758 (0.719–0.794). Transported tertiles showed graded post-landmark survival over 12 months in both arms.

Table 5

Facility-based split validation of trajectory-only and utilisation-enriched claims-trajectory models.

Care arm	Model	Transport design	Derivation sample	Validation sample	C-statistic	Validation risk groups
Acute care	Trajectory-only	Derivation: A Validation: B	N=540 events = 88	N=289 events = 44	Cder = 0.762 Cval = 0.660 (0.579–0.739)	Low: 2/52 (3.8%) Intermediate: 14/96 (14.6%) High: 28/141 (19.9%)
Acute care	Utilisation-enriched	Derivation: A Validation: B	N=540 events = 88	N=289 events = 44	Cder = 0.779 Cval = 0.693 (0.624–0.764)	Low: 3/81 (3.7%) Intermediate: 13/92 (14.1%) High: 28/116 (24.1%)
Chronic care	Trajectory-only	Derivation: C and D Validation: E	N = 3,245 events = 694	N=697 events = 113	Cder = 0.692 Cval = 0.746 (0.706–0.786)	Low: 5/219 (2.3%) Intermediate: 31/197 (15.7%) High: 77/281 (27.4%)
Chronic care	Utilisation-enriched	Derivation: C and D Validation: E	N = 3,245 events = 694	N=697 events = 113	Cder = 0.714 Cval = 0.758 (0.719–0.794)	Low: 3/240 (1.2%) Intermediate: 39/211 (18.5%) High: 71/246 (28.9%)

Models were fitted only in the derivation facility/facilities; Cox beta coefficients and derivation-sample linear-predictor tertile cutoffs were then applied to validation patients without refitting. Trajectory-only model = age + sex + CCI + D0 burden + six-month net accumulation. Utilisation-enriched model = trajectory-only model + M0-M6 utilisation-intensity score. Derivation Cox models were fitted with a ridge (L2) penalty of 0.01. Validation hospitals contributed no data to Cox-coefficient estimation or tertile-cutoff derivation; continuous-predictor standardisation constants were, however, calculated once across the pooled landmark cohort. Facility labels are anonymised. Validation risk groups show observed deaths within 12 months/N (%) as crude descriptive proportions within transported tertiles; deaths after leaving the participating hospital group were not captured, and the Kaplan-Meier panels in Fig. 6 *censor patients at their last observed care-system contact. Cder = derivation C-statistic; Cval = validation C-statistic with bootstrap 95% CI estimated in the held-out validation cohort. This table supports facility-held-out risk ordering within the Kinshukai five-hospital dataset and should not be interpreted as definitive external absolute-risk calibration. The full-sample apparent risk-ordering display in Supplementary Figure S10 is descriptive only and is not model-validation evidence.*

Facility heterogeneity was visible but did not erase the trajectory signal

Baseline burden, D0, net, utilisation, observed time, and mortality varied across the five facilities (Figures S1, S2; Tables S1, S2); chronic-care 12-month death ranged from 14.9% to 22.2%. In facility-adjusted models, the acute-care C-statistic rose from 0.705 (CCI + D0) to 0.733 (+ net) and 0.759 (+ utilisation); chronic-care values were 0.680, 0.703, and 0.723. Robustness across materially different facilities is a strength, though transportability beyond the Kinshukai group remains untested.

The supplementary sequence also addressed common alternative explanations

Supplementary analyses addressed alternative explanations. D0 prevalence did not equal dynamic contribution: cardiovascular involvement was common at D0 but weak as a net signal, whereas infectious accumulation (acute) and respiratory accumulation (chronic) contributed more (Figures S3, S4; Table S3). Injections, imaging, and billing-line intensity were the leading utilisation contributors in both arms, whereas medications and hospitalisation-related signals contributed more clearly in chronic care, and system-specific net and utilisation components were correlated but not redundant (Figures S5, S6; Tables S4, S5). Incremental value persisted across alternative net definitions, facility adjustment, utilisation-score definitions, and supportive outcomes (Tables S6–S9; Figure S7). Across 3-, 6-, and 12-month windows the incremental discrimination of net and utilisation was directionally preserved, with 6 months retained as the prespecified primary window (Figure S11; Table S13). Calibration by decile preserved broad ordering; observed rates fell below predicted risks in the highest deciles, a pattern that may partly reflect censoring of patients who left the care system (Table S10; Figure S8), reinforcing use as a risk-ordering phenotype rather than an absolute-risk calculator. Figure S9 and Table S11 provide breadth/depth component sensitivity, and Figure S10 provides a descriptive full-sample apparent risk-ordering display.

Discussion

This study identifies short-term multisystem accumulation as a dynamic dimension of late-life multimorbidity that is not captured by static comorbidity burden. Across acute and chronic care, the meaning of baseline disease burden depended on what happened next: low starting burden followed by substantial accumulation was associated with a substantially higher crude observed-death proportion within 12 months, while high starting burden was further separated by continued expansion. The signal involved spread across systems and deepening within systems, added information beyond CCI and D0, and preserved risk ordering in held-out hospitals. These findings position routine claims as a longitudinal phenotype for studying heterogeneity in late-life disease accumulation, not merely as a source of static risk adjustment.

The central result is the separation of accumulated state from ongoing change. D0 identified the documentation state at the start of the observation window, but subsequent net accumulation changed its prognostic meaning: patients with low initial burden and high accumulation had substantially higher mortality than those with low burden and stable documentation, while high-burden patients were further stratified by continued expansion. Static comorbidity counts cannot represent this distinction because they collapse disease history and recent change into a single value.

This temporal distinction connects to the deficit-accumulation view of frailty while remaining conceptually distinct from a frailty index. The deficit framework shows that vulnerability emerges from the cumulative burden of heterogeneous health deficits rather than from one disease alone.5 Our analysis adds a short-horizon longitudinal dimension: the same starting burden can have different prognostic meaning depending on subsequent accumulation. Breadth and depth further suggest that this change reflects both lateral spread across disease systems and intensification within affected systems. Because our inputs are claims-documented diagnoses rather than symptoms, function, cognition, or physiological deficits, the phenotype should be considered an administrative analogue of changing multimorbidity burden, not a measure of frailty or biological ageing.

The dynamic signal was not cardiovascular-dominant. Cardiovascular disease was common at baseline but contributed little as a dynamic signal, whereas infectious accumulation in acute care, respiratory accumulation in chronic care, and several renal, metabolic, haematological, and gastrointestinal signals contributed more strongly. This pattern supports a multisystem interpretation of dynamic multimorbidity: prognostically relevant change arose through distributed expansion rather than one index diagnosis.

The breadth and depth analyses make that interpretation concrete. Breadth captured lateral spread into newly involved systems and depth captured intensification within the most affected system; both correlated with total net accumulation and both separated mortality. Total net accumulation therefore acts as a compact summary of two modes of morbidity expansion. Its value is not only the numerical gain in discrimination but the ability to distinguish stable burden from active multisystem change.

Utilisation intensity provides a complementary view of the same clinical-administrative process. In claims, accumulation and utilisation overlap because both arise through care contact; that overlap is expected rather than evidence that the measures are interchangeable. One channel records what was documented, whereas the other records how intensively the patient was investigated, treated, and billed. The combined model showed the strongest internal discrimination, while the diagnosis-only trajectory remained informative when utilisation variables are difficult to harmonise across settings.

The held-out-hospital analysis is the strongest evidence that the phenotype is more than a within-sample pattern. Apart from continuous-predictor standardisation constants derived from the pooled landmark cohort, validation-hospital data were not used for Cox-coefficient estimation or tertile-cutoff derivation; the derivation coefficients and cutoffs were applied to the validation hospitals without refitting, with validation C-statistics reported with bootstrap 95% CIs. The trajectory-only model preserved discrimination in both settings, and adding utilisation intensity yielded numerically higher validation C-

statistics; paired within-validation contrasts were not estimated, so this difference is descriptive. In the chronic-care arm, validation discrimination equalled or exceeded the derivation estimate; because the held-out chronic facility had a wider spread of D0 and net values (Table S2) and the derivation and validation confidence intervals overlap, this is consistent with case-mix and sampling variation rather than with overfitting; the derivation interval was, however, estimated with the fitted linear predictor held fixed and therefore does not incorporate model-fitting uncertainty.

Existing one-year mortality scores in hospitalised older adults with multimorbidity report moderate discrimination and often require functional, nutritional, or geriatric-assessment data.24–26 The present claims-only models reached apparent C = 0.755 in acute care and 0.718 in chronic care, with held-out validation C = 0.660 and 0.746, respectively. These values should not be interpreted as superiority over published scores because discrimination depends on case mix, outcome definition, and validation design. Instead, they show that routinely collected longitudinal diagnosis and care-activity signals carry meaningful prognostic information even without specialist geriatric assessment.

The primary implication is for ageing research rather than treatment allocation. Claims-derived accumulation provides a reproducible way to distinguish stable high multimorbidity from short-term morbidity expansion in datasets that lack repeated frailty or functional assessments. It can be used to stratify longitudinal studies and to test concordance with frailty progression, physical function, cognition, institutionalisation, intrinsic capacity, patient-reported outcomes, and biological markers of ageing. Clinical or service use would require independent transportability, calibration, and prospective evaluation.

Several limitations bound this claim. The landmark design applies only to patients who survived and remained observable to 6 months, so it does not describe the sickest early deaths. The 6-month window is pragmatic rather than a proven biological interval, and net accumulation is a fixed-window change measure rather than a biological rate. All five hospitals belong to one Japanese group with a shared accounting system, so transportability to independent groups, regions, and coding environments remains to be confirmed. Claims diagnoses are influenced by suspected-diagnosis handling, rule-out coding, surveillance intensity, and billing, and do not measure frailty, intrinsic capacity, cognition, function, quality of life, or laboratory biology. Deaths were captured only within the participating hospitals, so patients who left the care system were censored at their last contact. If their mortality differed from that of patients who remained under observation, this censoring may be informative. Crude 12-month death proportions (Tables 2 and 5) and the observed event rates used for calibration (Table S10) are therefore likely to underestimate mortality, particularly in high-risk strata with short observed follow-up, and the apparent high-risk overprediction may partly reflect censoring rather than model miscalibration. Absolute-risk use is premature, and the strata should be read as ordered risk groups.

The next step is to link the phenotype to direct measures of ageing and vulnerability - including frailty indices, physical function, cognition, intrinsic capacity, biomarkers, and patient-reported outcomes - and to test transport across fully independent health systems.

Conclusions

In summary, short-term claims-derived multisystem accumulation captured a dynamic component of multimorbidity that static burden missed. Starting burden and subsequent accumulation jointly stratified mortality, and the signal reflected both breadth and depth of disease expansion while retaining risk ordering across held-out hospitals. Claims cannot substitute for frailty or biological-ageing measures, but their longitudinal structure may provide a scalable bridge between deficit-accumulation concepts and real-world population data.

Abbreviations

CCI
Charlson Comorbidity Index
D0
Month 0 active disease burden
ICD-10
International Classification of Diseases, Tenth Revision
M0-M6
index month through Month 6.

Declarations

Ethics approval and consent to participate

This retrospective observational study was approved by the Kinshukai Group Joint Research Ethics Review Committee on 1 July 2026 (approval 2026-03); institutional permission was granted by Kinshukai Medical Corporation on 1 July 2026 (permission 75). The study used pre-existing claims and clinical information from five hospitals, with no intervention, patient contact, or biological samples. The requirement for written informed consent was waived under the approved protocol and relevant Japanese ethical guidelines; information was disclosed through an opt-out procedure allowing patients or their representatives to refuse use of their information. Direct identifiers were removed and replaced with study IDs, with the correspondence table held separately by Kinshukai Medical Corporation. The study followed the Declaration of Helsinki and the Japanese Ethical Guidelines for Medical and Biological Research Involving Human Subjects. The study was not prospectively registered. Patients and the public were not involved in the design, conduct, or reporting of this study.

Consent for publication

Not applicable.

Competing Interests

MO and TA are employees of ASCLEPIUS Inc., which is contracted by the Kinshukai Medical Corporation to support research data extraction, data management, and research-dataset construction. AK is an employee of Canon ITS Medical Inc., which provided technical advice on electronic-medical-record data structure and extraction specifications. These companies had no role in study design, analysis, interpretation, or the decision to submit. CO, YM, MY, SY, and MK declare no competing interests.

Funding

The authors received no specific funding for this study.

Author Contribution

CO: conceptualisation, methodology, formal analysis, visualisation, writing - original draft, writing - review and editing. MO: data curation, software, validation, writing - review and editing. AK: resources, software, technical advice on electronic-medical-record data structure and extraction specifications, writing - review and editing. TA: project administration, resources, data-curation support, writing - review and editing. YM and MY: clinical interpretation, investigation, writing - review and editing. SY: conceptualisation, methodology, supervision, writing - review and editing. MK: conceptualisation, supervision, project administration, writing - review and editing. CO and MO directly accessed and verified the underlying individual-participant data. All authors reviewed the analysis outputs and manuscript, accept responsibility for the integrity of the work, and approved the final manuscript and submission.

Acknowledgements

Not applicable.

Data Availability

Individual-level claims and clinical data and the associated participant-level data dictionary are not publicly available because they contain sensitive patient information and unrestricted external sharing is not permitted by the approved protocol or institutional data-use arrangements. Beginning with publication and for 5 years, qualified researchers may request aggregated analysis outputs and the operational phenotype specification from the corresponding author (kitakaze@zf6.so-net.ne.jp) for methodologically sound, non-commercial research. Access is subject to approval by Kinshukai Medical Corporation and, where applicable, the ethics committee, together with a signed data-use agreement; no

individual-participant data will be transferred. The approved Japanese protocol, an English-language protocol synopsis, and the opt-out information document may be made available on reasonable request.

Code availability:

The statistical code used to construct the longitudinal claims phenotype and reproduce the manuscript analyses is supplied as a cleaned source-code archive with the submission for editorial and peer-review use. The archive contains no patient-level data or direct identifiers; execution requires the protected source data, which cannot be distributed externally. Beginning with publication and for 5 years, the code may also be provided to qualified researchers on reasonable request to the corresponding author, subject to the institutional approval and data-use conditions described above.

References

1. World Health Organization. World report on ageing and health. Geneva: World Health Organization; 2015.
2. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet. 2016;387:2145–54.
3. World Health Organization. Decade of healthy ageing: baseline report. Geneva: World Health Organization; 2020.
4. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet. 2019;394:1365–75.
5. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol Biol Sci Med Sci. 2007;62:722–7.
6. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet. 2013;381:752–62.
7. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol Biol Sci Med Sci. 2001;56:M146–56.
8. Covinsky KE, Pierluissi E, Johnston CB. Hospitalization-associated disability: 'She was probably able to ambulate, but I'm not sure'. JAMA. 2011;306:1782–93.
9. Gill TM, Gahbauer EA, Han L, Allore HG. Trajectories of disability in the last year of life. N Engl J Med. 2010;362:1173–80.
10. Lunney JR, Lynn J, Foley DJ, Lipson S, Guralnik JM. Patterns of functional decline at the end of life. JAMA. 2003;289:2387–92.
11. Nicholson K, Makovski TT, Griffith LE, Raina P, Stranges S, van den Akker M. Multimorbidity and comorbidity revisited: refining the concepts for international health research. J Clin Epidemiol. 2019;105:142–6.

12. 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;380:37–43.
13. Marengoni A, Angleman S, Melis R, et al. Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev. 2011;10:430–9.
14. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA. 2005;294:716–24.
15. Onder G, Palmer K, Navickas R, et al. Time to face the challenge of multimorbidity. A European perspective from the joint action on chronic diseases and promoting healthy ageing across the life cycle (JA-CHRODIS). Eur J Intern Med. 2015;26:157–9.
16. Kingston A, Robinson L, Booth H, Knapp M, Jagger C. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. Age Ageing. 2018;47:374–80.
17. Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. Ann Fam Med. 2012;10:142–51.
18. Arakelyan S, Mikula-Noble N, Ho L, et al. Effectiveness of holistic assessment-based interventions for adults with multiple long-term conditions and frailty: an umbrella review of systematic reviews. Lancet Healthy Longev. 2023;4:e629–44. 10.1016/S2666-7568(23)00190-3.
19. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43:1130–9.
20. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992;45:613–9.
21. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40:373–83.
22. Calderón-Larrañaga A, Fabbri E, González AI, et al. Understanding changes in complex care needs over time: key research insights into multimorbidity trajectories. Lancet Healthy Longev. 2025;6(11):100790. 10.1016/j.lanhl.2025.100790.
23. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12:e1001885.
24. Schneider C, Aubert CE, Del Giovane C, et al. Comparison of 6 mortality risk scores for prediction of 1-year mortality risk in older adults with multimorbidity. JAMA Netw Open. 2022;5:e2223911. 10.1001/jamanetworkopen.2022.23911.
25. Aubert CE, Rodondi N, Terman SW, et al. HOSPITAL Score and LACE Index to Predict Mortality in Multimorbid Older Patients. Drugs Aging. 2022;39:223–34. 10.1007/s40266-022-00927-0.
26. Li Y, Liu X, Kang L, Li J. Validation and comparison of four mortality prediction models in a geriatric ward in China. Clin Interv Aging. 2023;18:2009–19. 10.2147/CIA.S429769.

27. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4:e296.
28. van Houwelingen HC. Dynamic prediction by landmarking in event history analysis. Scand J Stat. 2007;34:70–85.
29. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. 10.1136/bmj-2023-078378.
30. Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10:e1001381. 10.1371/journal.pmed.1001381.

Figures

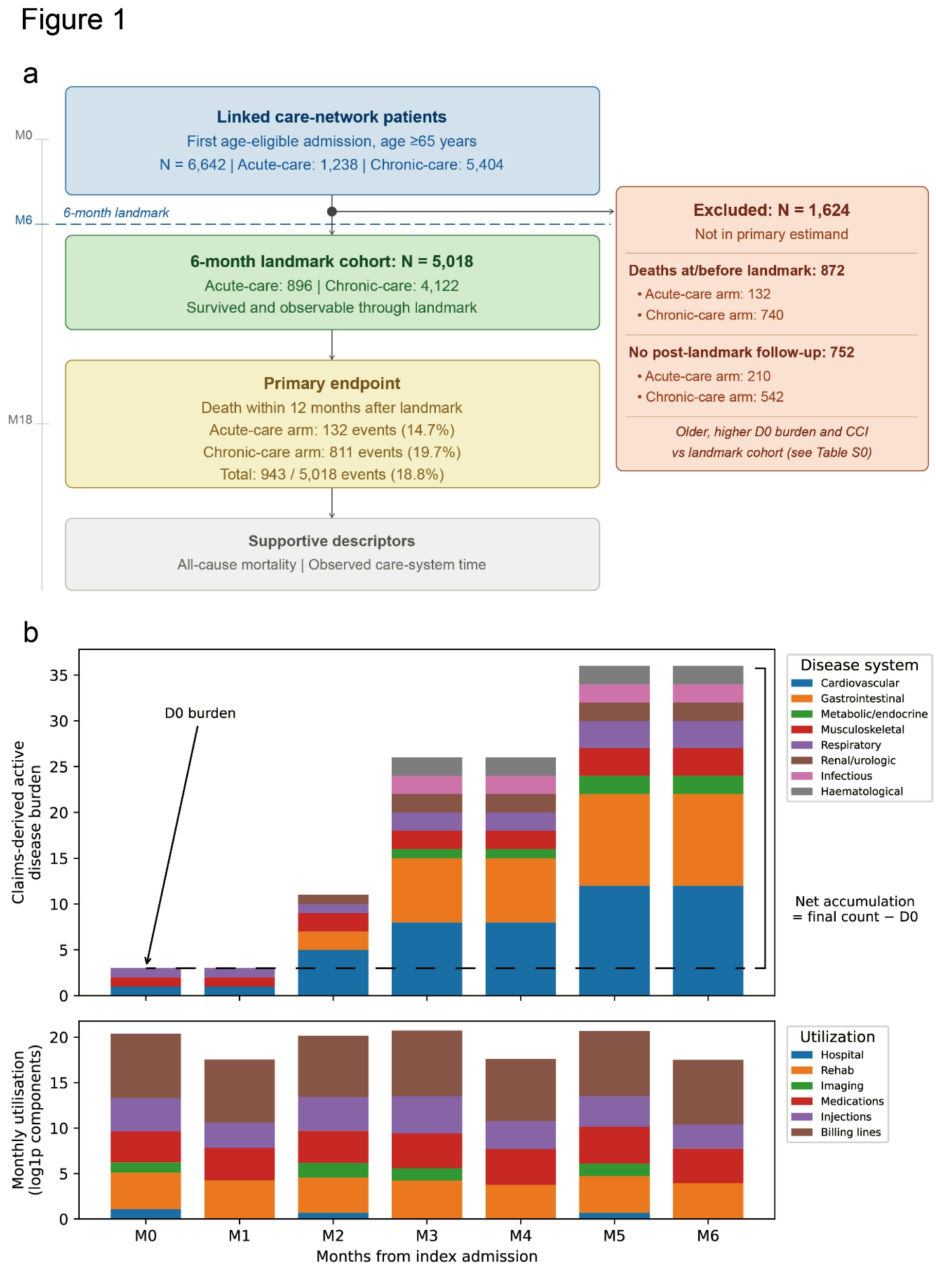


Figure 1

Cohort and 6-month landmark design. Cohort assembly, landmark design, and illustrative longitudinal claims measurement. (A) Patients with first observed admission at age 65 years or older across five Kinshukai hospitals, attrition before the 6-month landmark, acute-care and chronic-care landmark strata, and the primary endpoint of death within 12 months after the landmark. Inference applies to patients who remained observable through the landmark and had post-landmark follow-up. Patients who died at

or before the landmark or lacked positive post-landmark follow-up were not included in the primary landmark estimand. (B) Representative anonymised patient-level monthly trajectory showing active disease burden and utilisation components during M0-M6. Net accumulation was defined as final active disease count minus D0 burden; utilisation was analysed as a parallel care signal and was not part of the disease-burden definition. Note: D0=Month 0 active disease burden; M0=Month 0/index admission; M0-M6=index month through Month 6. Panel B shows an anonymised illustrative trajectory selected to demonstrate feature construction, not a distributional summary.

Figure 2

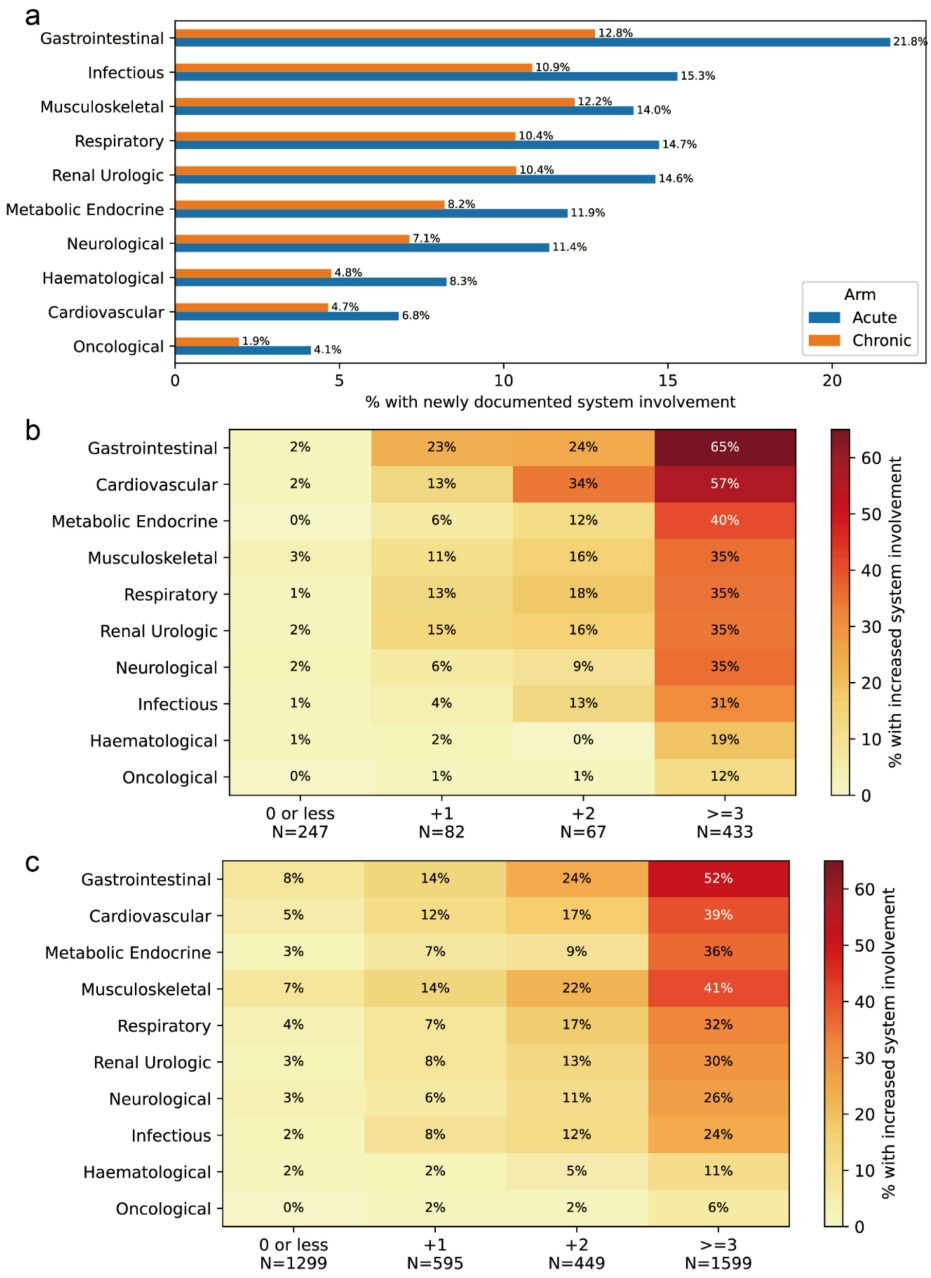


Figure 2

Multisystem disease emergence and expansion during M0-M6. (A) Newly documented disease-system involvement during M0-M6 by care setting. (B and C) Increased disease-system involvement by net-accumulation stratum in the acute-care (B) and chronic-care (C) arms, shown using the same disease-system order for comparability; panels B and C include only patients with an available final M1-M6 disease count (acute N=829; chronic N=3,942). The figure demonstrates that six-month net accumulation reflects multisystem expansion rather than a cardiovascular-only trajectory. Note: D0=Month 0 active disease burden; M0-M6=index month through Month 6. The final count used the M6 active disease count where available, or the last available M1-M6 active disease count.

Figure 3

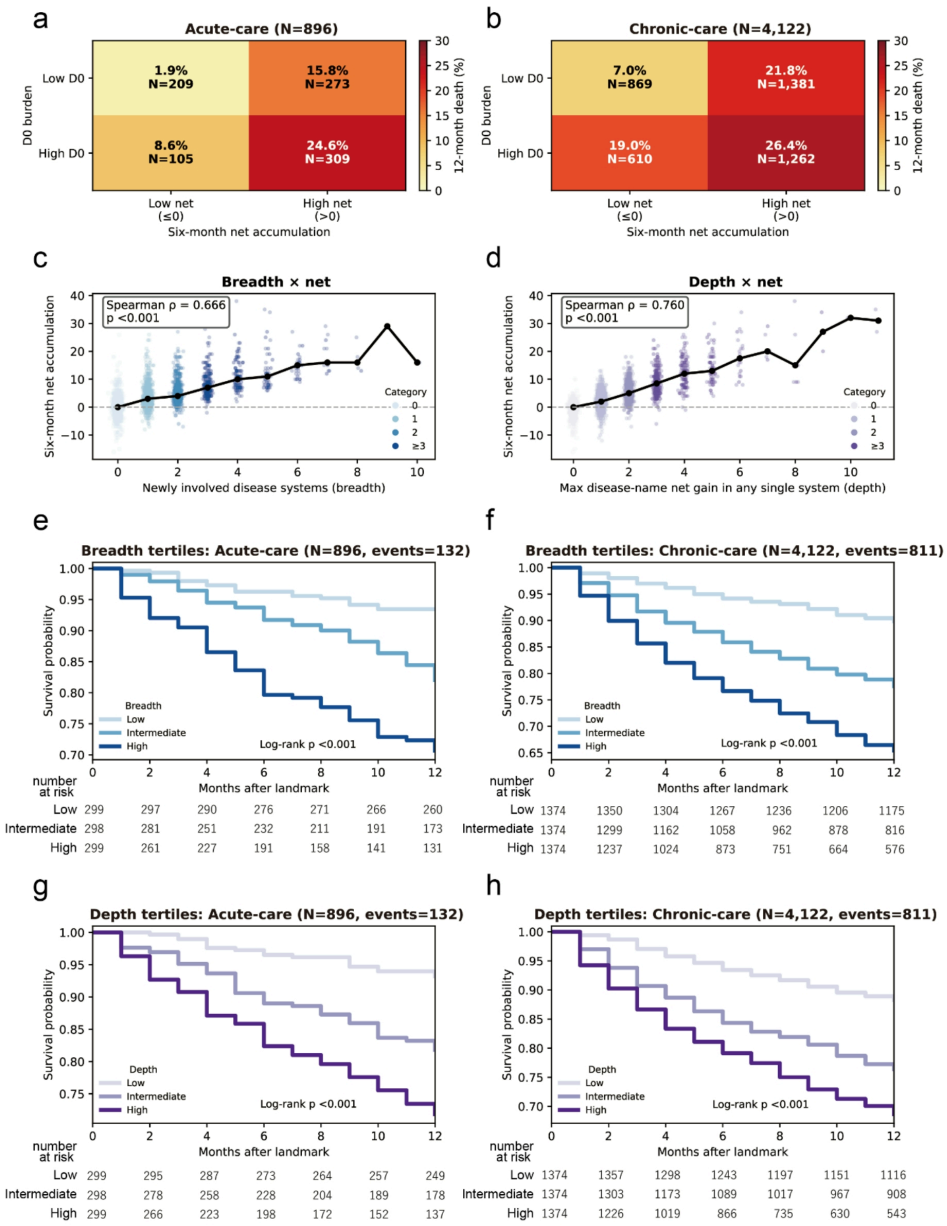


Figure 3

D0 burden, net accumulation, and breadth/depth patterns of morbidity expansion. Panels A and B show the crude proportion with an observed death within 12 months after the landmark by low/high D0 burden and low/high six-month net accumulation in acute-care and chronic-care arms; these percentages are not Kaplan-Meier 12-month risk estimates. Panels C and D show the relationship of total net accumulation with breadth, defined as the number of newly involved disease systems, and depth,

defined as the maximum disease-name net gain within any single system. Panels E-H show Kaplan-Meier post-landmark survival through 12 months by breadth and depth tertiles in acute-care and chronic-care arms; global two-sided log-rank tests compare the three displayed strata in each panel. Note: D0=Month 0 active disease burden; M0-M6=index month through Month 6. Descriptive Figure 3 panels retain all landmark patients, whereas net-based regression and discrimination analyses use complete cases with available final M1-M6 counts.

Figure 4

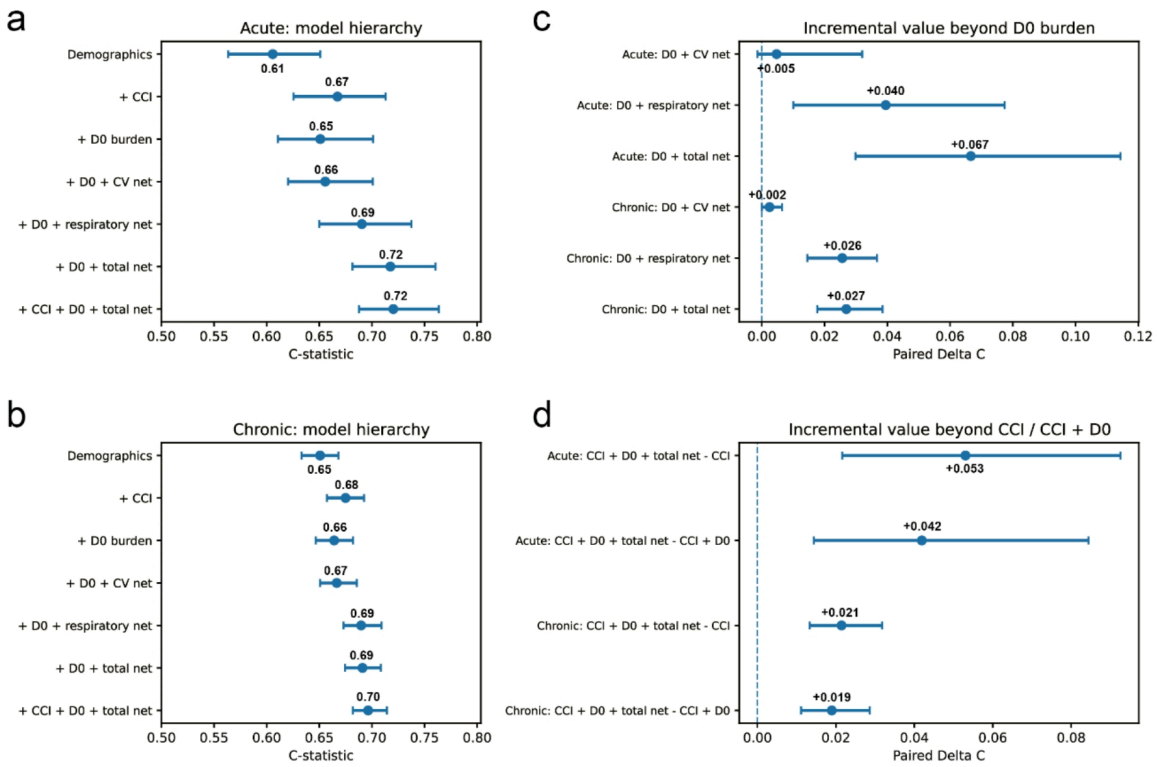


Figure 4

Added prognostic information of multisystem net accumulation. Panels A and B show C-statistics for alternative models in acute-care and chronic-care arms; every model includes age and sex, and each row adds only the listed terms to age and sex (rows are not cumulative). Panel C shows paired Delta-C contrasts for cardiovascular-only, respiratory-only, and total multisystem net accumulation added to D0 burden (reference: age, sex, and D0 burden). Panel D shows paired Delta-C contrasts for D0 burden plus total net accumulation added to CCI (reference: age, sex, and CCI; +0.053 acute, +0.021 chronic) and for total net accumulation added to CCI + D0 (reference: age, sex, CCI, and D0 burden; +0.042 acute, +0.019 chronic). The figure tests whether dynamic multisystem information extends beyond static and organ-specific comparators. Note: CCI=Charlson Comorbidity Index; CV=cardiovascular; Delta C=paired change

in C-statistic. Horizontal bars indicate interval estimates for the corresponding C-statistic or paired Delta-C contrast.

Figure 5

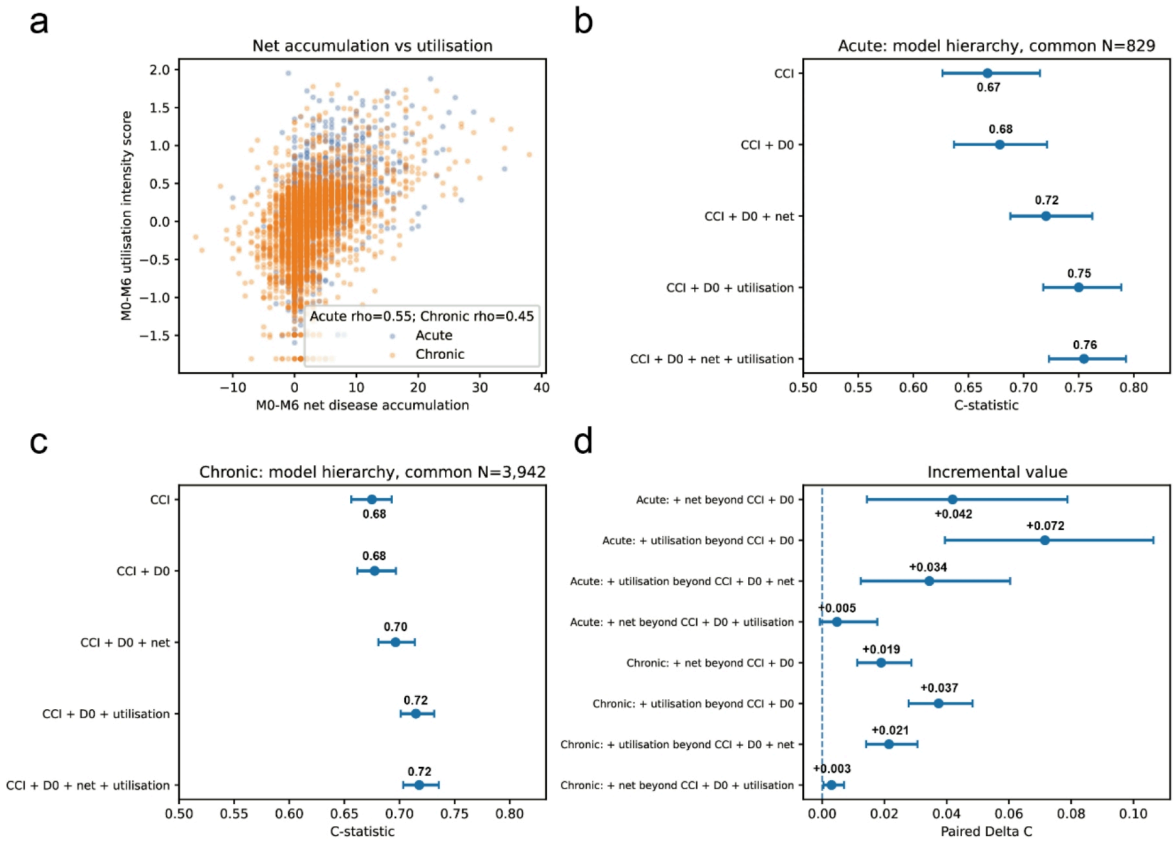


Figure 5

Net accumulation, utilisation, and internal discrimination of claims-trajectory models. (A) Scatterplot of six-month net accumulation versus utilisation intensity, with Spearman correlations by care setting. (B and C) Model hierarchies comparing CCI, CCI + D0, CCI + D0 + net, CCI + D0 + utilisation, and the combined claims-trajectory model. (D) Incremental value of net and utilisation when added to the corresponding common-N reference models. Note: CCI=Charlson Comorbidity Index; D0=Month 0 active disease burden; rho=Spearman rank correlation. Utilisation denotes the standardised M0-M6 utilisation-intensity score; common-N analyses use the same patients within each displayed comparison.

Figure 6

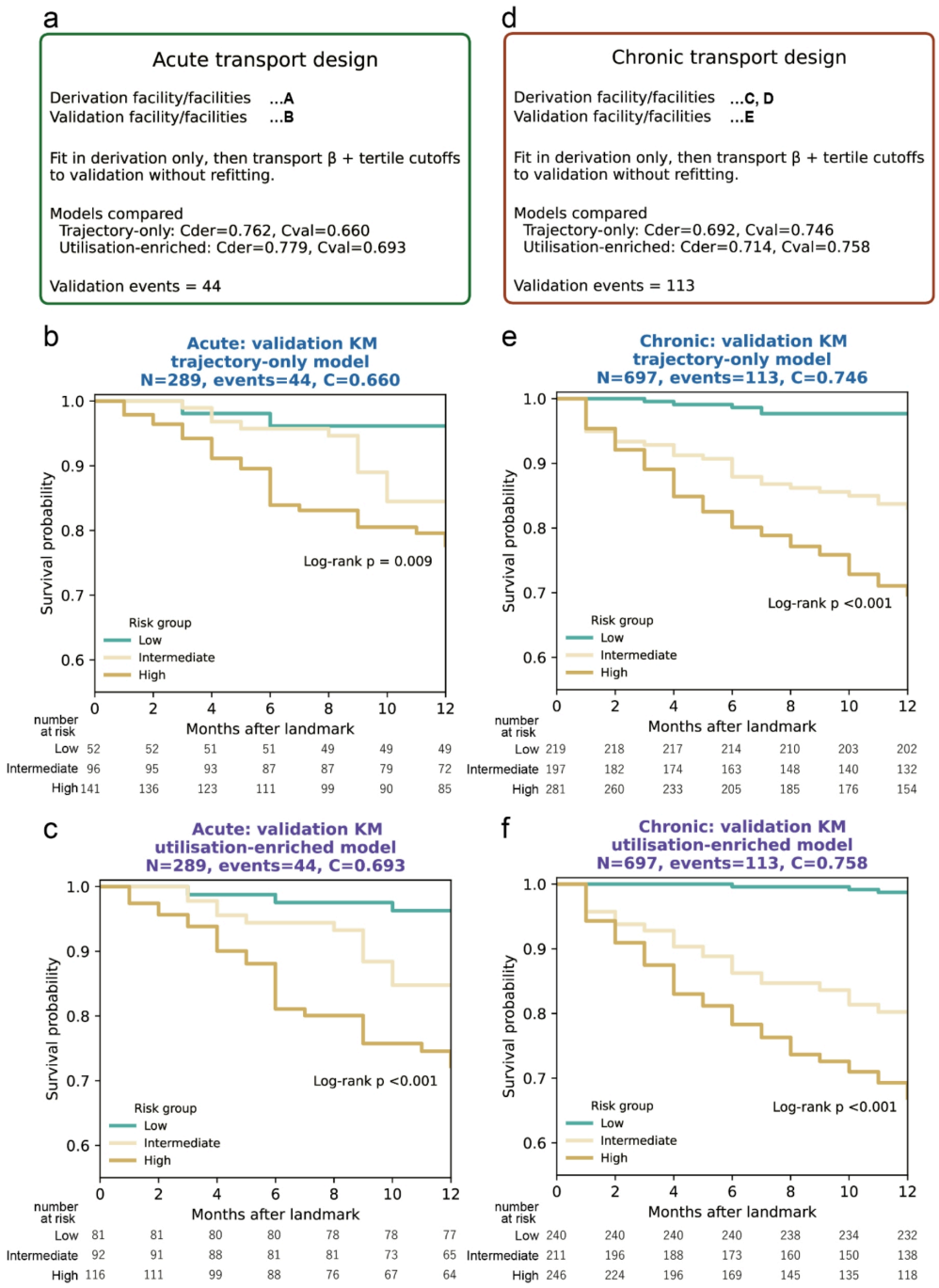


Figure 6

Facility-based split validation of trajectory-only and utilisation-enriched claims-trajectory models. (A and D) Validation design for acute-care and chronic-care analyses: Cox coefficients and derivation-sample linear-predictor tertile cutoffs were fitted in derivation facilities and applied to held-out validation facilities without refitting. (B and E) Validation Kaplan-Meier curves showing transported risk groups from trajectory-only models. (C and F) Validation Kaplan-Meier curves showing transported risk groups

from utilisation-enriched models. Global two-sided log-rank tests compare the three transported risk groups in each Kaplan-Meier panel; P values are shown in the panels. The trajectory-only model included age, sex, CCI, D0 burden, and six-month net accumulation; the utilisation-enriched model added the M0-M6 utilisation-intensity score. Note: CCI=Charlson Comorbidity Index; D0=Month 0 active disease burden; M0-M6=index month through Month 6. Panel titles report validation N, events, and validation C-statistic; bootstrap 95% CIs are reported in Table 5 and Table S12. The figure supports facility-held-out risk ordering within a single hospital group, not external absolute-risk calibration.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BMCGeriatricsSupplementaryInformationOkamotoetal.pdf