Double trouble: a propensity-matched cohort study evaluating the associations between duplicate medical records and patient outcomes

Gavriel Roda [1], Angela Keniston [2], Nicholas Wood,[2] Hillary Western[2]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjqs-2025-019112).

[1]Department of Internal Medicine, University of Colorado Anschutz Medical Campus School of Medicine, Aurora, Colorado, USA
[2]University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA

Correspondence to
Dr Hillary Western;
hillary.western@cuanschutz.edu

Received 31 May 2025
Accepted 11 December 2025

Check for updates

To cite: Roda G, Keniston A, Wood N, *et al*. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2025-019112

ABSTRACT

Introduction Duplicate medical records occur when a single patient is assigned multiple medical record numbers within an instance of an electronic health record, potentially associated with fragmented care and adverse outcomes. Despite these concerns, limited research has evaluated the correlation between duplicate charts and patient outcomes.

Objective To examine the association between duplicate charts and patient outcomes, including hospital length of stay, 30-day readmission, rapid response events, intensive care unit (ICU) level of care, and in-hospital mortality.

Methods This retrospective cohort study analysed hospitalised patients aged 18–89 across 12 hospitals within a large multi-region health system from July 1, 2022 to June 30, 2023. Propensity score matching balanced covariates between patients with and without duplicate charts. Primary outcomes included in-hospital mortality, rapid response events, ICU level of care, hospital length of stay, 30-day readmission and 30-day emergency department visits. Standardised mean differences assessed group balance, and multivariable logistic or linear regression models, adjusted for discharge service and disposition, examined the relationship between patients with and without duplicate records and the selected outcomes.

Results After matching, 1698 patients with duplicate charts were compared with 4388 without. Patients with duplicate records had significantly higher odds of adverse outcomes, including 30-day readmission (OR=1.3, 95% CI 1.1 to 1.5, p=0.0122), ICU level of care (OR=3.5, 95% CI 3.1 to 4.0, p<0.0001), and in-hospital mortality (OR=4.7, 95% CI 3.7 to 6.0, p<0.0001). Additionally, hospital length of stay was 32% longer (p<0.0001) for patients with duplicate charts.

Conclusion Patients with duplicate medical records demonstrated higher odds of adverse patient outcomes compared with those without, including increased mortality, ICU level of care, and prolonged hospitalisation. These findings highlight the need for research to understand the impacts of duplicate charts.

WHAT IS ALREADY KNOWN ABOUT THIS TOPIC

⇒ Duplicate medical records are a source of fragmented care and information gaps. However, the relationship between duplicate charts and patient outcomes remains poorly understood.

WHAT THIS STUDY ADDS

⇒ This propensity-matched cohort study suggests that duplicate records are associated with increased odds of adverse inpatient outcomes, including in-hospital mortality, intensity care unit level of care, and 30-day readmission. To our knowledge, this is the first study to evaluate duplicate records' association with patient outcomes.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY

⇒ These findings highlight the association between duplicate medical records and adverse patient outcomes, emphasising the need for research to understand the impacts of duplicate charts as well as targeted interventions to improve data integrity, enhance patient safety, and inform policy changes in health information management.

INTRODUCTION

Duplicate medical records arise when a single patient is assigned multiple medical record numbers within an instance of an electronic health record (EHR).[1–3] These duplicates may pose challenges to health-care efficiency, contribute to monetary costs and redundant medical workups, increase cognitive burden on providers, and delay care delivery.

Most healthcare systems estimate their duplicate record rate to be between 5% and 10%, although rates vary across institutions.[1 3–5] The American Health Information Management Association recommends a system-wide duplicate rate of 1% or less, yet only 22% of healthcare organisations meet this benchmark, and just 61% actively track their duplicate rate.[6] One study evaluating multiple healthcare systems found that between 16.49% and 40.66% of records shared a matching first and last name with another individual. When birth dates were included as an additional identifier, this overlap decreased to 0.16% to 15.47%. However, these findings suggest that duplicate records may be more prevalent than commonly estimated, as a substantial proportion of patient identifiers appear in multiple charts.[5]

Key patient identifiers—such as legal name, date of birth, and address—are common sources of mismatches, leading to erroneous duplicate creation.[4 7] In a previous study conducted through our institution, approximately 20% of duplicate charts were created through unintentional error during emergency department (ED) registration, with another 20% created owing to a lack of patient-identifying information or confirmation. The remaining duplicate charts were strategically created through a standardised pathway to expedite care for high-risk medical conditions, including myocardial infarctions, strokes, and traumas. These duplicate charts are created despite the patients often having known identities.[3]

Despite the prevalence and perceived risks of duplicate charts, their association with patient outcomes remains poorly understood.[2 8] Using propensity score matching, this study aims to address this gap by evaluating the relationship between duplicate charts and key inpatient outcomes, including hospital length of stay, 30-day readmission, rapid response events, intensive care unit (ICU) level of care, and in-hospital mortality. We hypothesise that patients with duplicate medical records experience longer lengths of stay and higher odds of adverse outcomes compared with those with single, accurate records.

METHODS

Study design and participants

This retrospective cohort study included patients aged 18–89 who were hospitalised at one of 12 partner hospitals within a large multi-region system between July 1, 2022, and June 30, 2023. Figure 1 summarises the cohort selection. Eligible patients were those discharged from medical, surgical, or orthopaedic care teams. Encounters related to obstetrics, maternal and child health services, and discharges performed directly by ED personnel were excluded, given the expected differences in clinical pathways and implications of duplicate records. For patients who had previously been assigned a duplicate chart, only the first encounter with a duplicate chart during the study period was included. For patients without any history of duplication, the first eligible encounter within the time frame was retained.

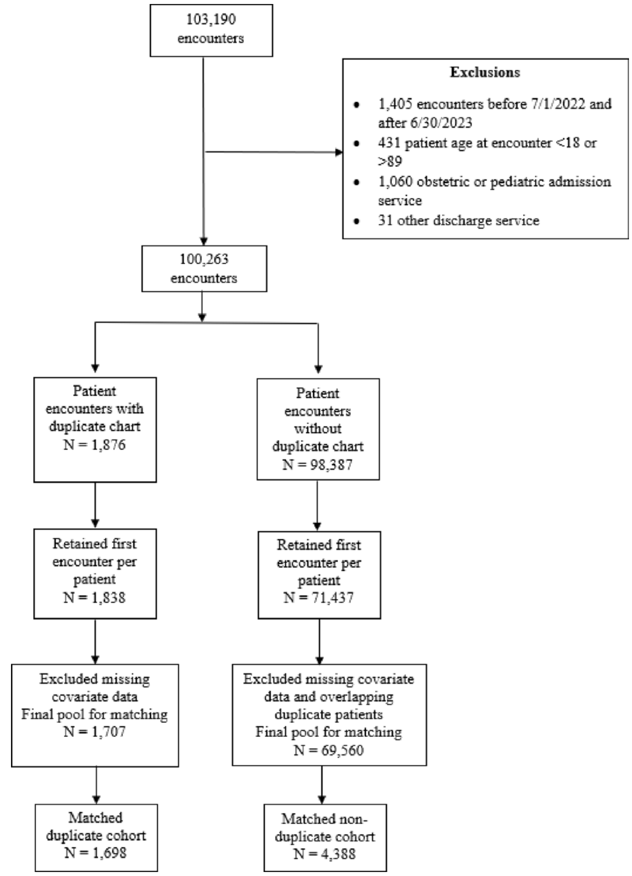

Figure 1 Cohort selection.

Clinical data were extracted from the Epic data warehouse by Health Data Compass—a multi-institutional data warehouse supported by the system and its affiliates. Partial funding for data acquisition through Health Data Compass was provided by the Clinical Effectiveness and Patient Safety Grant Programme, supported by the Institute for Healthcare Quality, Safety, and Efficiency. The Colorado Multiple Institutional Review Board determined that the study qualified as exempt.

Outcome measures

The primary outcome of this analysis was all-cause, in-hospital mortality. Secondary outcomes included ICU admission, occurrence of rapid response events, and hospital length of stay. Additionally, patients were followed for 30 days post discharge to assess rates of hospital readmission and return visits to the ED.

Primary predictor

The primary explanatory variable was the patient having a duplicate medical record during their hospitalisation. Duplicate medical records were detected using

Table 1 Patient characteristics before and after matching

Characteristics	Before matching: With duplicate chart (n=1838)	Before matching: Without duplicate chart (n=71 437)	Before matching: Total (n=73 275)	Before matching: Absolute SMD*	After matching: With duplicate chart (n=1698)	After matching: Without duplicate chart (n=4388)	After matching: Total (n=6086)	After matching: Absolute SMD*
Duplicate chart, n (%)				N/A				N/A
Yes	1838 (100)	0 (0)	1838 (2.5)		1698 (100)	0 (0)	1698 (27.9)	
No	0 (0)	71 437 (100)	71 437 (97.5)		0 (0)	4388 (100)	4388 (72.1)	
Age (years), mean+SD	56+19	60+18	60+18	0.2051	56+19	57+19	57+19	0.0166
Sex, n (%)				0.1199				0.0388
Female	790 (43.0)	35 113 (49.1)	35 903 (49.0)		735 (43.3)	1924 (43.9)	2659 (43.7)	
Male	1048 (57.0)	36 323 (50.9)	37 371 (51.0)		963 (56.7)	2464 (56.1)	3427 (56.3)	
Missing	0	1	1		0 (0)	0 (0)	0 (0)	
Race, n (%)								
American Indian or Alaska Native	23 (1.3)	624 (0.9)	647 (0.9)	0.0275	20 (1.2)	42 (1.0)	62 (1.0)	0.0192
Asian	5 (0.3)	243 (0.3)	248 (0.3)	0.0091	5 (0.3)	11 (0.2)	16 (0.3)	0.0018
Black or African American	258 (14.0)	4719 (6.6)	4977 (6.8)	0.2117	235 (13.8)	647 (14.7)	882 (14.5)	0.0147
Native Hawaiian and Other Pacific Islander	7 (0.4)	122 (0.2)	129 (0.2)	0.0236	5 (0.3)	12 (0.3)	17 (0.3)	0.0036
White or Caucasian	1207 (65.7)	55 845 (78.2)	57 052 (77.9)	0.2541	1147 (67.6)	2949 (67.2)	4096 (67.3)	0.0052
More than one race	15 (0.8)	479 (0.7)	494 (0.7)	0.0167	14 (0.8)	25 (0.6)	39 (0.6)	0.0326
Other	291 (15.8)	8508 (11.9)	8799 (12.0)	0.1098	272 (16.0)	702 (16.0)	974 (16.0)	0.0072
Missing	32 (1.7)	897 (1.3)	929 (1.3)		0 (0)	0 (0)	0 (0)	
Ethnicity, n (%)				0.1111				0.0055
Hispanic, Latino/a, or Spanish origin	353 (19.2)	10 496 (14.7)	10 849 (14.8)		327 (19.3)	797 (18.2)	1124 (18.5)	
Non-Hispanic, Latino/a, or Spanish Origin	1449 (78.8)	59 925 (83.9)	61 374 (83.8)		1371 (80.7)	3591 (81.8)	4962 (81.5)	
Missing	36 (2.0)	1016 (1.4)	1052 (1.4)		0 (0)	0 (0)	0 (0)	
Primary language, n (%)								
English	1678 (91.3)	67 430 (94.4)	69 108 (94.3)	0.1017	1556 (91.6)	4022 (91.7)	5578 (91.7)	0.0145
Spanish	123 (6.7)	2867 (4.0)	2990 (4.1)	0.1020	111 (6.5)	268 (6.1)	379 (6.2)	0.0071
Other	33 (1.8)	1081 (1.5)	1114 (1.5)	0.0219	31 (1.8)	98 (2.2)	129 (2.1)	0.0169
Missing	4 (0.2)	59 (0.1)	63 (0.1)		0 (0)	0 (0)	0 (0)	
Payor, n (%)								
Medicare	756 (41.1)	34 553 (48.4)	35 309 (48.2)	0.1535	694 (40.9)	1852 (42.2)	2546 (41.8)	0.0214
Medicaid	551 (30.0)	13 056 (18.3)	13 607 (18.6)	0.2527	506 (29.8)	1296 (29.5)	1802 (29.6)	0.0277
Commercial	355 (19.3)	17 273 (24.2)	17 628 (24.1)	0.1198	331 (19.5)	837 (19.1)	1168 (19.2)	0.0047
Indigent care	2 (0.1)	73 (0.1)	75 (0.1)	0.0040	2 (0.1)	3 (0.1)	5 (0.1)	0.0115
Self-pay	77 (4.2)	2688 (3.8)	2765 (3.8)	0.0288	72 (4.2)	172 (3.9)	244 (4.0)	0.0107

Continued

Table 1 Continued

	Before matching				After matching			
Characteristics	With duplicate chart (n=1838)	Without duplicate chart (n=71 437)	Total (n=73 275)	Absolute SMD*	With duplicate chart (n=1698)	Without duplicate chart (n=4388)	Total (n=6086)	Absolute SMD*
Other	97 (5.3)	3794 (5.3)	3891 (5.3)	0.0061	93 (5.5)	228 (5.2)	321 (5.3)	0.0099
Middle name, n (%)				N/A				N/A
Yes	1083 (58.9)	42 610 (59.7)	43 693 (59.6)		1010 (59.5)	2637 (60.1)	3647 (59.9)	
No	755 (41.1)	28 827 (40.3)	29 582 (40.4)		688 (40.5)	1751 (39.9)	2439 (40.1)	
Valid social security name, n (%)				0.1614				0.0072
Yes	1691 (92.0)	62 448 (87.4)	64 139 (87.5)		1562 (92.0)	4019 (91.6)	5581 (91.7)	
No	147 (8.0)	8989 (12.6)	9136 (12.5)		136 (8.0)	369 (8.4)	505 (8.3)	
Hyphenated or spaced last name, n (%)				0.0939				0.0092
Yes	121 (6.6)	2954 (4.1)	3075 (4.2)		108 (6.4)	243 (5.5)	351 (5.8)	
No	1717 (93.4)	68 483 (95.9)	70 200 (95.8)		1590 (93.6)	4145 (94.5)	5735 (94.2)	
Admission service, n (%)				1.7190				0.0181
Emergency Medicine	1703 (92.7)	47 324 (66.2)	49 027 (66.9)		1645 (96.9)	4213 (96.0)	5858 (96.2)	
Not Emergency Medicine	54 (2.9)	23 573 (33.0)	23 627 (32.2)		53 (3.1)	175 (4.0)	228 (3.8)	
Missing	81 (4.4)	540 (0.8)	621 (0.9)		0 (0)	0 (0)	0 (0)	
Patient class, n (%)				N/A				N/A
Elective	4 (0.2)	10 708 (15.0)	10 712 (14.6)		0 (0)	60 (1.4)	60 (1.0)	
Emergency	1325 (72.1)	57 487 (80.5)	58 812 (80.3)		1214 (71.5)	3504 (79.9)	4718 (77.5)	
Trauma centre	491 (26.7)	1180 (1.7)	1671 (2.3)		467 (27.5)	767 (17.5)	1234 (20.3)	
Urgent	18 (1.0)	2061 (2.9)	2079 (2.8)		17 (1.0)	57 (1.3)	74 (1.2)	
Missing	0	1	1		0 (0)	0 (0)	0 (0)	
Patient status, n (%)				0.3859				0.0098
Inpatient	1596 (86.8)	52 690 (73.8)	54 286 (74.1)		1472 (86.7)	3707 (84.5)	5179 (85.1)	
Observation patient	242 (13.2)	18 747 (26.2)	18 989 (25.9)		226 (13.3)	681 (15.5)	907 (14.9)	
Stroke alert, n (%)				0.1189				0.0225
Yes	45 (2.4)	485 (0.7)	530 (0.7)		40 (2.4)	94 (2.1)	134 (2.2)	
No	1793 (97.6)	70 952 (99.3)	72 745 (99.3)		1658 (97.6)	4294 (97.9)	5952 (97.8)	
Trauma alert, n (%)				0.6793				0.0070
Yes	635 (34.6)	1480 (2.1)	2115 (2.9)		575 (33.9)	1036 (23.6)	1611 (26.5)	
No	1203 (65.4)	69 957 (97.9)	71 160 (97.1)		1123 (66.1)	3352 (76.4)	4475 (73.5)	
Elixhauser Comorbidity Index, most acuteean+SD	postacute							
Comorbidity index for risk of 30-day, all-cause readmission	0.70+1.8	0.87+2.3	0.87+2.3	N/A	0.70+1.8	0.92+2.3	0.86+2.1	N/A

Continued

Table 1 Continued

	Before matching				After matching			
Characteristics	**With duplicate chart (n=1838)**	**Without duplicate chart (n=71 437)**	**Total (n=73 275)**	**Absolute SMD***	**With duplicate chart (n=1698)**	**Without duplicate chart (n=4388)**	**Total (n=6086)**	**Absolute SMD***
Comorbidity index for risk of in-hospital mortality	2.4+6.5	1.5+5.3	1.5+5.4	0.1258	2.3+6.4	2.3+6.1	2.3+6.2	0.0232
Discharge service, n (%)				N/A				N/A
Medicine	1208 (65.7)	49 345 (69.1)	50 553 (69.0)		1120 (66.0)	3032 (69.1)	4152 (68.2)	
Surgery	582 (31.7)	18 510 (25.9)	19 092 (26.0)		532 (31.3)	1190 (27.1)	1722 (28.3)	
Orthopaedics	48 (2.6)	3582 (5.0)	3630 (5.0)		46 (2.7)	166 (3.8)	212 (3.5)	
Discharge disposition, n (%)				N/A				N/A
Home	1106 (60.2)	59 143 (82.8)	60 249 (82.2)		1040 (61.2)	3414 (77.8)	4454 (73.2)	
Postacute care	426 (23.2)	8713 (12.2)	9139 (12.5)		383 (22.6)	680 (15.5)	1063 (17.5)	
Left against medical advice	40 (2.2)	970 (1.4)	1010 (1.4)		37 (2.2)	100 (2.3)	137 (2.2)	
Hospice	51 (2.8)	953 (1.3)	1004 (1.4)		44 (2.6)	67 (1.5)	111 (1.8)	
Died before discharge	206 (11.2)	1344 (1.9)	1550 (2.1)		185 (10.9)	109 (2.5)	294 (4.8)	
Missing	9 (0.5)	314 (0.4)	323 (0.4)		9 (0.5)	18 (0.4)	27 (0.4)	

*SMD, absolute standardised mean differences.

Table 2 Outcomes after propensity matching

	After matching		
	With duplicate chart (n=1698)	Without duplicate chart (n=4388)	Total (n=6086)
Rapid response event			
Yes	103 (6.1)	219 (5.0)	322 (5.3)
No	1595 (93.9)	4169 (95.0)	5764 (94.7)
Care received in the ICU, n (%)			
Yes	777 (45.8)	850 (19.4)	1627 (26.7)
No	921 (54.2)	3538 (80.6)	4459 (73.3)
Died during hospitalisation, n (%)			
Yes	185 (10.9)	109 (2.5)	294 (4.8)
No	1513 (89.1)	4279 (97.5)	5792 (95.2)
Hospital length of stay (hours), median (IQR)	101 (51, 219)	74 (45, 138)	79 (46, 159)
Readmission within 30 days, n (%)			
Yes	203 (12.0)	472 (10.8)	675 (11.1)
No	1495 (88.0)	3916 (89.2)	5411 (88.9)
Return to ED within 30 days, n (%)			
Yes	207 (12.2)	558 (12.7)	765 (12.6)
No	1491 (87.8)	3830 (87.3)	5321 (87.4)

ED, emergency department; ICU, intensive care unit.

a multistep approach. Epic's 'Duplicate Patient Detection and Management' process is applied to search for matched pairs within the master patient index.[9] In addition, a manual process may be employed, where a clinical staff member, registrar, or other employee identifies that a patient has an additional chart in the EHR. All potential duplicates identified via either method underwent a manual review by a trained data integrity professional to verify that the records belonged to the same individual.

Statistical analysis

Propensity score matching

We employed propensity score matching to create comparable groups based on observed covariates, aiming for matched groups with similar distributions of baseline characteristics.[10–12] A logistic regression model was used to estimate the probability of a patient having a duplicate medical chart, as a function of variables measured at the time of patient admission. Variables hypothesised to be associated with the primary outcome, patient death before discharge, were included in the model, as were variables hypothesised to be associated with the presence of a duplicate record also found to be associated with the outcome of interest. Covariates included in the propensity score model were selected based on a combination of literature review and the clinical and operational experience of the study team. The following baseline covariates were included: patient age, alert category (stroke alert, trauma alert, or neither stroke nor trauma alert), sex assigned at birth (male or female), race (Black or African American, White or Caucasian, American Indian or Alaskan Native, Native Hawaiian or other Pacific Islander, more than one race, or other), ethnicity (non-Hispanic, Latino/a, or Spanish origin or Hispanic, Latino/a, or Spanish origin), preferred language (English, Spanish, or other), having a valid social security number, having a hyphenated or spaced last name, health insurance type (commercial, uninsured care, Medicaid, Medicare, self-pay, or other), admission service (emergency medicine, not emergency medicine), patient status (inpatient or observation), and Elixhauser Comorbidity Index assessing the risk of in-hospital mortality derived using the Elixhauser Comorbidity Software Refined (v2022.1).

We explored multiple matching ratios (1:1, 2:1, and 3:1) and calliper widths (0.1, 0.2, and 0.25) using nearest neighbour matching and assessed covariate balance. After comparing absolute standard mean differences (SMDs) across these matching approaches, 3:1 nearest neighbour matching, with a calliper width equal to 0.2 times the SD of the logit of the propensity score, without replacement, was used to match patients with a duplicate chart to patients without a duplicate chart (maximum SMD – 0.0388; online supplemental Appendix Figure 1). Propensity score distributions showed strong overlap between exposed (those with a duplicate chart) and unexposed (those without a duplicate chart) groups after matching, indicating adequate common support (online supplemental Appendix Figure 2).

Outcome analysis

Means and SD are reported for continuous variables, while frequencies and proportions are reported for categorical variables. We used multivariable regression modelling to examine the relationship between patients with and without duplicate records and the selected outcomes. Linear regression was used for hospital length of stay, while logistic regression was used for in-hospital mortality, rapid response events, ICU level of care, 30-day readmission, and 30-day return to the ED.

A natural log transformation was applied to length of stay to conform to a normal distribution. The associated change in hospital length of stay was calculated by exponentiating the coefficient, subtracting one, and expressing the result as a percentage. We employed Firth's penalised likelihood regression for the in-hospital mortality outcome, which occurred in 2.5% of patients without a duplicate chart during the study period, to mitigate rare event bias. When analysing hospital length of stay, readmission within 30 days, and return to the ED within 30 days, patients who died before discharge were excluded.

In addition, to account for other potential sources of variability, we further adjusted for variables that were not measurable at baseline but were hypothesised to

Table 3 Unadjusted and adjusted odds between duplicate medical records and hospital outcomes

	Unadjusted		Adjusted	
	OR (95% CI)	**P value**	**OR (95% CI)**	**P value**
Rapid response event*	1.2 (0.96 to 1.6)	0.0929	1.2 (1.0 to 1.6)	0.0810
Received care in the ICU*	3.5 (3.1 to 4.0)	<0.0001	3.5 (3.1 to 4.0)	<0.0001
Died during hospitalisation*	4.8 (3.8 to 6.1)	<0.0001	4.7 (3.7 to 6.0)	<0.0001
Readmission within 30 days*†	1.2 (1.0 to 1.5)	0.0129	1.3 (1.1 to 1.5)	0.0122
Return to ED within 30 days*†	1.1 (0.9 to 1.25)	0.5268	1.1 (0.9 to 1.3)	0.2552
	Percentage change (95% CI)	**P value**	**Percentage change (95% CI)**	**P value**
Hospital Length of Stay*†	46% (38 to 54)	<0.0001	32% (25 to 39)	<0.0001

*Adjusted for discharge service (medicine, surgery, or orthopaedics).
†Adjusted for discharge service (medicine, surgery, or orthopaedics) and discharge disposition (against medical advice (AMA), postacute care, home, hospice, or in-hospital mortality).

be associated with the outcomes (discharge disposition and discharge service). All models were also adjusted for discharge service (medicine, surgery, or orthopaedics), and models for hospital length of stay, 30-day readmission, and 30-day return to the ED were also adjusted for discharge disposition (against medical advice, postacute care, home, or hospice).

We assessed missing values across the total cohort and found that the proportion of values missing was less than 2% for each variable. Patients with missing values were excluded from the analysis to perform a complete-case analysis.

Statistical programmes

Propensity score estimation and matching were conducted in R statistical software (version 4.2.3; R Foundation). All other analyses were performed using SAS Enterprise Guide 8.3 (SAS Institute, Inc., Cary, North Carolina, USA).

RESULTS

Between July 1, 2022 and June 30, 2023, a total of 103 190 patient charts from our institution were reviewed and 73 275 patients were eligible for inclusion in the study. From this group, 6086 matched patients were identified with 1698 having duplicate charts and 4388 without duplicate charts. Table 1 summarises patient characteristics before and after propensity matching.

Differences between the two groups were substantially reduced through propensity matching (table 1). The mean age was 60±18 years for patients without duplicate charts and 56±19 years for those with duplication. Similarly, the sex distribution differed, with 49% female patients in the non-duplicate group compared with 43% in the duplicate group. Post-matching, the mean age was 57±19 years for patients without duplicate charts and 56±19 years for those with a duplicate chart. The sex distribution was also balanced, with 44% female patients in the non-duplicate group compared with 43% in the duplicate group. Additionally, baseline health characteristics, such as the Elixhauser Comorbidity Indices for the risk of in-hospital mortality and admission service, were balanced between the two groups after matching.

These results demonstrate that the matching process significantly improved balance between the groups, reducing the absolute SMD for most characteristics to near zero and ensuring comparability for the subsequent analysis of patient outcomes (online supplemental Appendix Figure 1).

Patient outcomes after propensity matching are summarised in table 2. After matching, in-hospital death occurred in 11% of patients with duplicate charts compared with 2.5% in patients without duplicate charts. The median length of stay was also longer in the duplicate chart group (101 hours) compared with those without duplicate charts (74 hours). Additionally, patients with duplicate charts had more frequent rapid response events (6% vs 5%) and were more likely to receive care in the ICU (46% vs 19%). The 30-day readmission rate was higher in the duplicate chart group (12% vs 11%), although the rate of return to the ED within 30 days of discharge was similar between the two groups (12% vs 13%).

Analysis of binary outcomes for matched patients after further adjustment for discharge service and discharge disposition revealed significant differences in receiving care in the ICU, in-hospital mortality, readmission within 30 days, and hospital length of stay. Table 3 presents the unadjusted and adjusted odds ratios and their 95% CIs for these outcomes. After adjustment, patients with duplicate charts had 3.5 times higher odds of receiving care in the ICU (OR=3.5, 95% CI 3.1 to 4.0, p<0.0001), 4.7 times higher odds of dying during hospitalisation (OR=4.7, 95% CI 3.7 to 6.0, p<0.0001), and 1.3 times higher odds of being readmitted within 30 days (OR=1.3, 95% CI 1.1 to 1.5, p=0.0122). Additionally, hospital length of stay was 32% longer for those with duplicate charts after adjusting for discharge service and disposition (p<0.0001). Figure 2 depicts these ORs in a Forest plot.

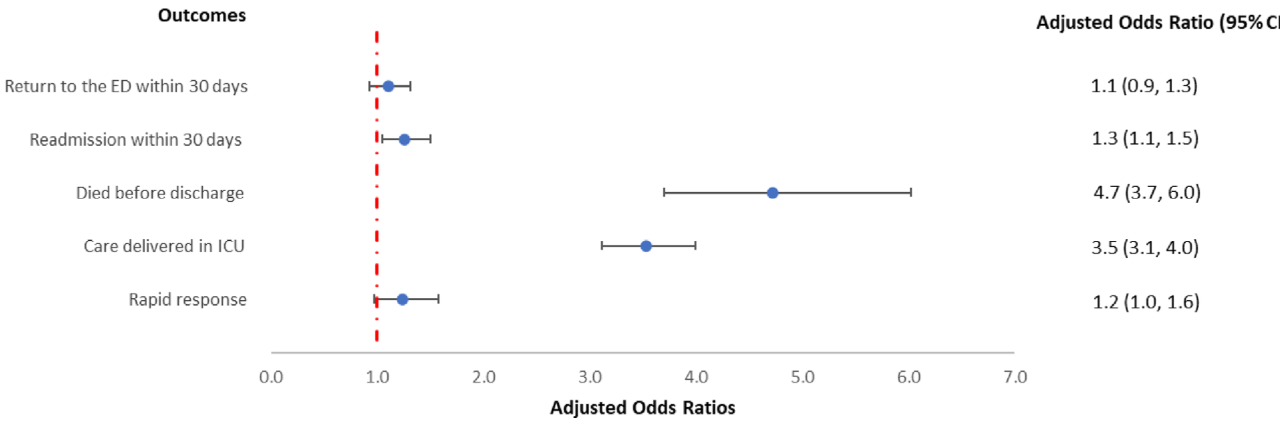

Figure 2 Forest plot of adjusted odds ratios for binary hospital outcomes. ED, emergency department; ICU, intensive care unit.

DISCUSSION

This study identified several adverse patient outcomes that are associated with duplicate medical records, including increased odds of 30-day readmission and receiving care in the ICU. Notably, patients with duplicate medical records demonstrated 4.7-fold higher odds of inpatient mortality compared with those without duplicate records. While our study is unable to establish causality between the evaluated adverse outcomes and chart duplication, it highlights a need for further evaluation into the mechanisms underlying such findings. The study design enhances the significance of these results by accounting for confounding factors that could influence outcome comparisons through propensity matching, creating comparable groups of exposed (patients with duplicate medical records) and unexposed (patients without duplicate medical records) individuals based on measured covariates.

The increased odds of inpatient mortality for patients with duplicate charts are particularly concerning and more profound than expected, given the existing literature citing rates of preventable harm between 7% and 12%.[13 14] A key hypothesis is that the presence of duplicate charts may increase adverse events and thus preventable harms. Mechanisms include preventing providers from accessing critical information, such as allergies or past medical history, information that would alter medical care. An additional hypothesis relates to efficiency: the presence of duplicate charts may contribute to care delays or inaccurate orders as medical teams search for information that is not readily accessible, spend extra effort navigating between multiple charts, or inadvertently overlook key details. We encourage future study of these hypotheses and their relationships between duplicate charts and adverse outcomes.

While duplicate charts may be created in some instances of critical illness, such as stroke alerts and trauma alerts, propensity score matching resulted in similar Elixhauser Comorbidity Indices between the two groups and similar rates of these alert types. Other disease states that may be associated with duplicate chart creation and thus overrepresented in the duplicate chart population, such as myocardial infarctions and burns, were not included in our propensity score model, outside of their general relationship to the Elixhauser Comorbidity Index, and thus may explain some of the increased odds of negative outcomes. Despite this limitation, however, we do not believe this potential difference between groups fully explains the increased odds of adverse events in patients with duplicate charts.

This study has several limitations. First, additional variables, such as the number of diagnoses and number of healthcare encounters, could be factored into a model should this study be replicated, and their absence in our analysis might have affected our results.[15] Next, the use of data from a single health system may limit the generalisability of findings to other regions or healthcare settings. Differences in patient populations, as well as variation in EHR algorithms, documentation practices, and registration workflows, may influence the applicability of these results elsewhere. Despite these limitations, our study highlights a concerning association within our system and underscores the importance for outside systems to investigate their own associations, determine causal pathways, and develop mechanisms to prevent duplicate chart creation and/or conduct data integration expeditiously.

Further limitations include the restriction of our analysis to encounters within the system and did not account for external care; however, examining data on a per-encounter basis may mitigate this limitation to some extent. Also, the retrospective design introduces the potential for bias due to missing or inaccurate documentation. Furthermore, variability in provider or departmental documentation practices may further contribute to inconsistencies in the dataset. Additionally, the study period (July 1, 2022–June 30, 2023) may reflect temporal factors such as updates to clinical guidelines, changes in ICD coding, staffing dynamics, or technological upgrades, all of which could affect documentation and workflow. Despite the use of propensity score matching to reduce bias, residual confounding cannot be ruled out, and causal inferences cannot be made. Finally, there may be unmeasured factors not captured in this analysis that influence the observed associations. Although we selected covariates based on literature review and investigator expertise, we acknowledge that some potentially important variables might not have been included in the analysis, such as the cause for duplicate chart creation as an example. Our data were derived from the EHR, which limited us to variables that were feasibly and consistently extractable.

CONCLUSION

This study provides insights into adverse outcomes in the setting of duplicate medical charts, particularly highlighting an association with inpatient mortality, ICU admissions, and longer hospital stays. Despite the significance of this issue, a causal relationship has not been proved and research on the outcomes associated with duplicate records remains scarce. Future studies are needed to evaluate the mechanisms underlying these findings.

Contributors Study design was completed by HW. GR drafted the abstract, introduction, results, discussion, and conclusion. AK described the methods, performed statistical analyses, and created tables and figures. NW contributed background research for the introduction. GR, AK, and HW collaborated on revisions. All authors reviewed and approved the final manuscript. HW is the guarantor and accepts full responsibility for the integrity of the work. ChatGPT was used to assist with editing. It was not used to generate text but rather to edit for clarity and flow. AI was not used in the study other than in edits.

Funding This study was funded by Clinical Effectiveness and Patient Safety Grant Program (supported by the Institute for Healthcare Quality, Safety and Efficiency) (N/A).

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable.

ORCID iDs
Gavriel Roda https://orcid.org/0000-0001-8594-5078
Angela Keniston https://orcid.org/0000-0003-1399-2881

REFERENCES

1 McClellan MA. Duplicate medical records: a survey of Twin Cities healthcare organizations. *AMIA Annu Symp Proc* 2009;2009:421–5.
2 Joffe E, Bearden CF, Byrne MJ, *et al*. Duplicate patient records--implication for missed laboratory results. *AMIA Annu Symp Proc* 2012;2012:1269–75.
3 Western HL, Lin C-T, Seffrood E, *et al*. SEE ALTERNATE MRN: Duplicate charts in hospitalized patients. *J Hosp Med* 2024;19:852–4.
4 Fox L, Thierry S. EHR preparation: building your mpi game plan. n.d. Available: https://www.elitelearning.com/resource-center/health-information-professionals/ehr-preparation-building-your-mpi-game-plan
5 McCoy AB, Wright A, Kahn MG, *et al*. Matching identifiers in electronic health records: implications for duplicate records and patient safety. *BMJ Qual Saf* 2013;22:219–24.
6 American Health Information Management Association. A realistic approach to achieving a 1% duplicate record error rate. 2020. Available: https://ahima.org/media/m1pldevh/ahima-pim-whitepaper.pdf
7 Khunlertkit A, Dorissaint L, Chen A, *et al*. Reducing and Sustaining Duplicate Medical Record Creation by Usability Testing and System Redesign. *J Patient Saf* 2021;17:e665–71.
8 Stiell A, Forster AJ, Stiell IG, *et al*. Prevalence of information gaps in the emergency department and the effect on patient outcomes. *CMAJ* 2003;169:1023–8.
9 Epic Systems Corporation. *Duplicate Patient Detection and Management Setup and Support Guide*. Verona, WI: Epic Systems Corporation, 2025.
10 Ambroggio L, Cotter JM, Hall M. Methodological progress note: A clinician's guide to propensity scores. *J Hosp Med* 2022;17:283–6.
11 Brookhart MA, Schneeweiss S, Rothman KJ, *et al*. Variable selection for propensity score models. *Am J Epidemiol* 2006;163:1149–56.
12 Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res* 2011;46:399–424.
13 Bates DW, Levine DM, Salmasian H, *et al*. The Safety of Inpatient Health Care. *N Engl J Med* 2023;388:142–53.
14 Panagioti M, Khan K, Keers RN, *et al*. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. *BMJ* 2019;366:l4185.
15 Sahin O, Zhao A, Applegate RJ, *et al*. Epidemiology of Patient Record Duplication. *Appl Clin Inform* 2025;16:24–30.