Journal of Clinical Epidemiology 172 (2024) 111400

Journal of Clinical Epidemiology

METHODOLOGICAL CONSIDERATIONS RELATED TO EQUITY, DIVERSITY, AND INCLUSION IN CLINICAL EPIDEMIOLOGY

Avoiding double counting: the effect of bundling hospital events in administrative datasets for the interpretation of rural-urban differences in Aotearoa New Zealand

Rory Miller[a,*], Gabrielle Davie[b], Sue Crengle[c], Jesse Whitehead[d], Brandon De Graaf[b], Garry Nixon[e]

[a]Department of General Practice and Rural Health, University of Otago; Te Whatu Ora – Waikato (Thames Hospital), 55 Hanover Street, Dunedin, New Zealand 9016

[b]Department of Preventative and Social Medicine, University of Otago, Dunedin, New Zealand

[c]Ngāi Tahu Māori Research Unit, University of Otago, Dunedin, New Zealand

[d]Te Ngira Institute for Population Research, University of Waikato, Hamilton, New Zealand

[e]Department of General Practice and Rural Health, Dunstan Hospital, University of Otago, Clyde, New Zealand

Accepted 21 May 2024; Published online xxxx

Abstract

Background and Objectives: All publicly funded hospital discharges in Aotearoa New Zealand are recorded in the National Minimum Dataset (NMDS). Movement of patients between hospitals (and occasionally within the same hospital) results in separate records (discharge events) within the NMDS and if these consecutive health records are not accounted for hospitalization (encounters) rates might be overestimated. The aim of this study was to determine the impact of four different methods to bundle multiple discharge events in the NMDS into encounters on the relative comparison of rural and urban Ambulatory Sensitive Hospitalization (ASH) rates.

Methods: NMDS discharge events with an admission date between July 1, 2015, and December 31, 2019, were bundled into encounters using either using a) no method, b) an "admission flag", c) a "discharge flag", or d) a date-based method. ASH incidence rate ratios (IRRs), the mean total length of stay and the percentage of interhospital transfers were estimated for each bundling method. These outcomes were compared across 4 categories of the Geographic Classification for Health.

Results: Compared with no bundling, using the date-based method resulted in an 8.3% reduction (150 less hospitalizations per 100,000 person years) in the estimated incidence rate for ASH in the most rural (R2-3) regions. There was no difference in the interpretation of the rural-urban IRR for any bundling methodology. Length of stay was longer for all bundling methods used. For patients that live in the most rural regions, using a date-based method identified up to twice as many interhospital transfers (5.7% vs 12.4%) compared to using admission flags.

Conclusion: Consecutive events within hospital discharge datasets should be bundled into encounters to estimate incidence. This reduces the overestimation of incidence rates and the undercounting of interhospital transfers and total length of stay.

Keywords: Rural health; Administrative data; Interhospital transfer; Rural-urban disparities; Linked data; Ambulatory sensitive hospitalisations

1. Introduction

Health data related to publicly funded hospital events are commonly stored within administrative datasets that can be accessed for research purposes. In Aotearoa New Zealand (NZ), these data are stored in the National Minimum Dataset (NMDS) of hospital discharges that is maintained by NZ Manatū Hauora: Ministry of Health (MOH) and Te Whatu Ora - Health NZ (TWO) [1]. The NMDS records a new entry for each admission-discharge event (event hereafter). New events are generated when a patient is admitted to a hospital for 3 hours or more and then discharged (admission-discharge events less than 3 hours are recorded in a separate database, the National Nonadmitted Patient Collection – NNPAC). The discharge may be to home, to a different facility (eg, interhospital transfer) or a different

Funding: This study was supported by funding from the Health Research Council of New Zealand [19/488]. The funders did not have any input into the design of the study.

* Corresponding author. University of Otago; Te Whatu Ora – Waikato (Thames Hospital), Department of General Practice and Rural Health, 55 Hanover Street, Dunedin, New Zealand 9016.

E-mail address: Rory.miller@otago.ac.nz (R. Miller).

https://doi.org/10.1016/j.jclinepi.2024.111400

0895-4356/

What is new?

Key findings

- Using a dates-based method of bundling resulted in a lower and more realistic incidence of Ambulatory Sensitive Hospitalisation (ASH) and identifies longer mean length of stay and more inter-hospital transfers, although did not alter interpretation of any rural-urban differences.
- The dates method limits overcounting and identifies more inter-hospital transfers, which is important for interpreting rural health outcomes.

What this adds to what is known?

- If multiple events are not accounted for an overcount of the incidence of some conditions (ASH in this example) may occur.
- However, this didn't result in a change in interpretation of rural-urban difference for ASH.

What is the implication, what should change now?

- Consecutive events within hospital discharge datasets should be bundled into encounters to estimate incidence.
- Researchers should be explicit in describing how health encounters are formed from administrative datasets.

service or specialty within the same facility (eg, from an intensive care unit to a general medical ward – sometimes called a "statistical discharge"; however, if a change of clinical condition occurred [e.g., a patient developed a hospital-acquired pneumonia] and the patient didn't change service then a new event is not generated) [1,2]. If the discharge destination is to a different (or within the same) health care facility, it is possible that multiple events for the same presentation or health care encounter (encounter hereafter) exist [3,4]. If grouping these continuous events into encounters is not performed, the condition of interest might be overcounted, and therefore its incidence overestimated [5–7]. This is particularly important for conditions or populations where interhospital transfer is common, for example, in serious injury, myocardial infarction or for rural populations [3–5,8–10].

Multiple methods of bundling consecutive discharge events into health care encounters have been described. These include: using admission or discharge flags [2,11,12], using the dates of admission and discharge [3,10,13–15], or considering each event as its own encounter [16]. However, many studies do not articulate how individual events are grouped together into a health-care encounter [17–20].

Nearly 20% of people in NZ live in a rural area and should they need emergency or hospital-level health care, many of these people will access a rural or smaller urban hospital in the first instance [13,14,21,22]. Rural hospitals generally have fewer resources compared with urban hospitals and if higher levels of care are required, the patient is often transferred to an urban hospital [23–25]. These transfers may result in multiple events within hospitalization datasets (e.g., NMDS). However, these interhospital transfers are excluded from some studies [17,18].

The aim of this study was to determine the impact of a) no method, b) an admission flag, c) a discharge flag, or d) a date-based method to bundle events in the NMDS into encounters on the relative comparison of rural and urban Ambulatory Sensitive Hospitalization (ASH) rates.

2. Methods

2.1. Data sources

Publicly funded hospital admission, discharge, and clinical data were sourced from the NMDS [1]. As with all other administrative datasets maintained by TWO and the MOH, hospital admission and discharge events are linked by an encrypted unique patient identifier (the National Health Index – NHI) that is assigned to individuals who access health care in NZ. Approximately 95% of all individuals have an NHI assigned [26].

Data are stored within the NMDS as "events", with each event (row) containing demographic information, event start (admission) and end (discharge) dates, administrative data including an "admission source code" and an "event end type code". These two codes are coded manually by clerical staff, and therefore subject to error or misinterpretation. The admission source code "describes the nature of the admission for a hospital inpatient health event." The discharge event end type code identifies how a health-care event ended. Each event also contains International Classification for Disease 10th edition Australian Modification (ICD10-AM) codes for up to 15 diagnoses and procedures, respectively [1,27].

All NZ publicly funded hospital discharge events between July 1, 2015, and December 31, 2019, were obtained. Consistent with the inclusion criteria for ASHs, only hospital discharges for patients younger than 75 years at the start of the health encounter were considered.

Ethnicity was classified using a prioritization method [28]. Where a patient identified with more than one ethnicity, a single ethnicity was assigned with the following order of priority: Māori (the Indigenous people of NZ), Pacific, Asian, and European/Other. Ethnicity was further categorized into a binary classification: Māori and non-Māori.

Denominator data was sourced from the 2013 and 2018 NZ Census Estimated Resident Population [29]. Census data were grouped into 5-year age groups, Māori and non-Māori ethnicity and by male and female sex at birth. Data for intermediate years were estimated using linear interpolation and extrapolation.

2.2. Terminology

An event is an admission-discharge event or row within the NMDS. An encounter is a group of one or more continuous events.

2.3. Ambulatory Sensitive Hospitalizations

ASH are defined as hospitalizations of people less than 75 years of age "resulting from diseases sensitive to prophylactic or therapeutic interventions that are deliverable in a primary health care setting" and are similar to potentially preventable hospitalizations in Australia [30,31]. ASH were determined to be associated with a health-care encounter if any of the relevant ICD10-AM codes (Supplementary file 1) were recorded in the primary diagnosis field of any event within the health-care encounter [32]. ASH was chosen as it is a composite outcome with a wide but relatively standard set of diagnoses and is often used as a health system performance indicator.

2.4. Geographic Classification for Health (GCH)

The GCH is an index of rurality developed specifically for health research and policy in NZ [21]. The index defines five categories, with two urban categories (U1 (most urban), U2 (least urban)) and three rural categories (R1 (least rural) to R3 (most rural)) [21].

Patients were assigned to one of four GCH categories (U1, U2, R1, and R2-3) by matching the patient's domicile code as recorded in the NMDS to the appropriate GCH category using a concordance file [33]. The R2-3 category was created by combining R2 and R3 due to concerns about the small numbers that may occur in R3 (1% of people in NZ live in R3 areas) and potential inaccuracies in the GCH determination for the R3 category [21]. Domicile codes record the geographical area within which a health-care user's usual address is located at the time of admission and contain approximately 2000 people each [34]. Patients whose GCH classification was not able to be identified were excluded from the analysis.

2.5. Bundling methodologies

Three methods were used to bundle events into encounters and compared to each other and a fourth approach of not bundling. For all bundling methods, data were grouped by encrypted NHI and arranged by date. A unique identifier was assigned to all events within the same encounter. Within an encounter, the length of stay was determined by the difference between the admission date of the first event and the discharge date of the last event. An example of how the different bundling methods are applied is shown in Figure 1.

2.5.1. No bundling method

Each event was considered its own encounter.

2.5.2. Admission source flag

Using the "admission source code" variable, consecutive events with at least one "T" (transfer) were identified and grouped together along with the event (flagged "R" (routine)) that preceded the first identified "T" flag.

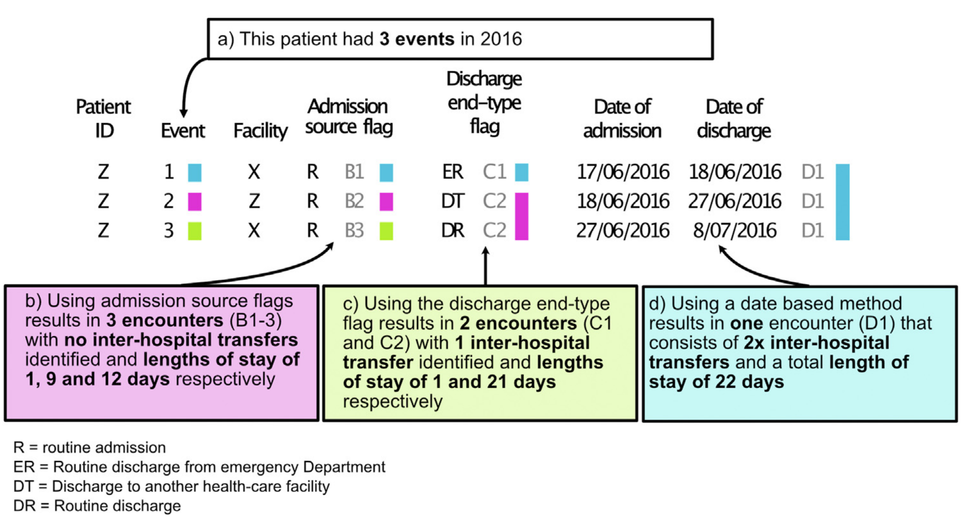

Figure 1. An illustrative example of how the different bundling methods (a) no bundling, b) admission source flag, c) event end-type flag and d) dates based, admission source flag and event end-type flag) were applied to data recorded in the New Zealand National Minimum Dataset for a hypothetical patient admitted for an Ambulatory Sensitive Hospitalization (ASH) condition.

2.5.3. Discharge flag

The following codes were identified in the ‘‘event end type code’’ variable: "DA" (Discharged to an acute facility), "DT" (Discharge of patient to another health-care facility), "DW" (Discharge to other service within same facility between the following types of specialty: AT&R (Assessment treatment and rehabilitation), mental health, personal health and palliative care), "DF" (Statistical discharge for change in funder), "ET" (Discharge from Emergency department acute facility to another health-care facility), "EA" (Discharge from Emergency department acute facility to specialist facility for neonates and burns only). The event (or consecutive events) following a relevant event end type code were bundled together into an encounter.

2.5.4. Dates method

Events that occurred with 1 day or less between the discharge date of one event and the admission date of the following event were considered one encounter.

2.6. Outcomes

The main outcome measures were ASH incidence rate ratios (IRRs), mean total length of hospital stay and the percentage of interhospital transfers identified by the different methods. As interhospital transfers are an explicit exclusion criteria from some studies, a secondary outcome was ASH IRRs after any encounter that included an interhospital transfer were excluded [17]. All outcomes were compared across 4 categories of the GCH for the total population and separately for Māori and non-Māori.

2.7. Statistical analysis

Means and standard deviations (SD) were calculated for continuous variables, as well as frequencies and percentages for categorical variables. Crude and age-sex standardized incidence rates (per 100,000 person years) and 95% confidence intervals (CIs) were calculated for ASH in the four rural-urban categories and separately for total, Māori and non-Māori using the direct standardization method [35]. The 2018 Māori census Estimated Resident Population was used as the standard population. Age-sex standardized IRRs and 95% CIs were calculated for each age-ethnicity strata [35]. This analysis was performed as IRRs are commonly used to assess rural-urban variation in health service utilization [21,36]. The most urban GCH category (U1) was considered the reference group.

All data management and analysis were performed using the R statistical programming language (version 4.2.1) in the RStudio (version 2022.12.0+353, Posit software) Integrated Development Environment [21].

3. Results

There were 470,945 patients with a total of 747,243 ASH discharge events identified in the NMDS dataset between 1 July 2015 and 31 December 2019. Table 1 describes the characteristics of patients hospitalized at least once with ASH. Most patients resided in the U1 (60%) GCH category, with 21% in U2, 12% in R1, and 7% in the R2-3 category. There were 6632 (0.1%) patients excluded who were not able to be assigned to a GCH category, leaving 466,603 for analysis. Around a fifth of patients (21%) was Māori and compared with U1 (17% Māori), a larger percentage of the R2-3 category (36%) was Māori.

3.1. ASH incidence rates

Table 2 and Figure 2 show the incidence rates per 100,000 person years, while Figure 3 shows the rural-urban incidence rate ratios for ASH for each of the four GCH categories. In all GCH categories, there was reduced ASH incidence rates when any of the three bundling methods were used compared to using no bundling methodology. Using the dates-based methodology resulted in the

Table 1. Characteristics (number and column percentage, unless otherwise stated) for 466,603 patients that had at least one Ambulatory Sensitive Hospitalization between July 1, 2015, and December 31, 2019

		Geographic classification for health			
Variable	Overall, 466,603	U1, 287,113 (62%)	U2, 96,345 (21%)	R1, 53,309 (11%)	R2-3, 29,836 (6.4%)
Age (years) - mean (SD)	38 (25)	37 (25)	39 (25)	41 (25)	41 (26)
<5 years	75,608 (16%)	49,465 (17%)	14,621 (15%)	7401 (14%)	4121 (14%)
5-15 years	47,367 (10%)	28,682 (10.0%)	9621 (10.0%)	5601 (11%)	3463 (12%)
65+ years	85,027 (18%)	46,373 (16%)	19,074 (20%)	12,399 (23%)	7181 (24%)
Sex					
Female	234,071 (50%)	144,037 (50%)	49,527 (51%)	26,099 (49%)	14,408 (48%)
Male	232,532 (50%)	143,076 (50%)	46,818 (49%)	27,210 (51%)	15,428 (52%)
Prioritized ethnicity					
Māori	100,224 (21%)	48,797 (17%)	28,525 (30%)	12,210 (23%)	10,692 (36%)
Non-Māori	366,379 (79%)	238,316 (83%)	67,820 (70%)	41,099 (77%)	19,144 (64%)

Table 2. The number, crude (Crude) and age-sex standardized Ambulatory Sensitive Hospitalization incidence rates (Standardized) with 95% confidence intervals in brackets for each of the four bundling methods and Geographic Classification of Health (GCH) categories

	Geographic classification for health											
	U1			U2			R1			R2-3		
Bundling method	*N*	Crude	Standardized	*N*	Crude	Standardized	*N*	Crude	Standardized	*N*	Crude	Standardized
Total												
None	448,866	3310	3062 (3052, 3072)	158,963	4204	3683 (3663, 3703)	85,618	3370	2891 (2870, 2913)	48,888	1757	1530 (1515, 1545)
Admission source flag	442,158	3260	3028 (3018, 3038)	154,449	4085	3614 (3594, 3634)	83,315	3279	2838 (2816, 2859)	46,979	1689	1485 (1471, 1500)
Discharge flag	438,702	3235	3009 (3000, 3019)	153,354	4056	3596 (3576, 3616)	82,148	3234	2808 (2786, 2830)	45,867	1649	1457 (1443, 1472)
Dates	431,629	3183	2961 (2951, 2970)	150,759	3987	3540 (3521, 3560)	80,604	3173	2759 (2738, 2781)	44,703	1607	1424 (1410, 1439)
Māori												
None	83,085	4635	4638 (4606, 4670)	49,010	5306	5318 (5271, 5366)	20,478	4047	4060 (4004, 4116)	18,337	2285	2307 (2273, 2340)
Admission source flag	82,118	4581	4584 (4552, 4615)	48,100	5207	5219 (5172, 5266)	20,004	3954	3965 (3910, 4020)	17,655	2200	2220 (2187, 2253)
Discharge flag	81,533	4548	4551 (4520, 4582)	47,769	5172	5183 (5136, 5229)	19,693	3892	3903 (3849, 3958)	17,241	2149	2167 (2135, 2200)
Dates	80,199	4474	4476 (4445, 4507)	46,956	5084	5094 (5048, 5140)	19,311	3817	3827 (3773, 3881)	16,840	2099	2116 (2084, 2149)
Non-Māori												
None	365,781	3108	2861 (2851, 2871)	109,953	3848	3242 (3220, 3265)	65,140	3202	2679 (2655, 2704)	30,551	1544	1278 (1260, 1295)
Admission source flag	360,082	3060	2829 (2819, 2840)	106,380	3723	3180 (3157, 3203)	63,324	3112	2632 (2607, 2656)	29,331	1482	1244 (1227, 1261)
Discharge flag	357,211	3035	2812 (2802, 2822)	105,623	3697	3166 (3143, 3189)	62,471	3071	2608 (2583, 2632)	28,634	1447	1222 (1206, 1240)
Dates	351,483	2987	2766 (2756, 2777)	103,844	3634	3118 (3095, 3140)	61,314	3014	2563 (2539, 2588)	27,873	1408	1195 (1178, 1212)

biggest reduction in rates (approximately 150 admissions per 100,000 person years). The percentage reduction ranged from 3% in U1 (total population) to 8% in the R2-3 Māori population. Supplementary Table 2 includes the percentage difference between the number of ASH encounters, and the crude and age-standardized rates with each bundling methodology compared to no bundling method.

The choice of bundling method did not result in any differences to the interpretation of the rural-urban

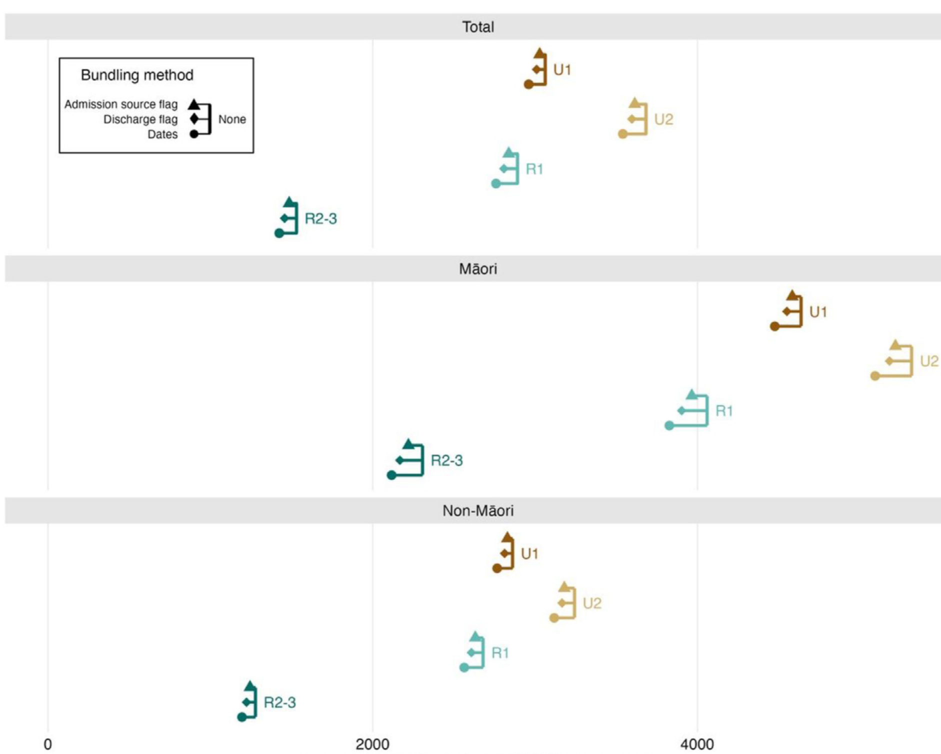

Figure 2. The age-sex standardized rate per 100 00 person years for Ambulatory Sensitive Hospitalizations. Rates for each of the three bundling methods (a) admission source flag — triangle, (b)discharge flag — diamond and (c) dates method — circle) are shown compared to no bundling method (vertical line) for each Geographic Classification for Health category (U1, U2, R1, and R2-3) and by total and Māori and non-Māori separately.

Figure 3. The age-sex standardized incidence rate ratios (and 95% confidence intervals) for Ambulatory Sensitive Hospitalizations in each of the four bundling methods (vertically) comparing the four Geographic Classification for Health categories (U1, U2, R1, and R2-3). U1 is the reference population.

standardized IRRs. The incidence rate for patients living in R2-3 areas was half that of patients that lived in U1 areas, while patients living in R1 had between 7% (non-Māori) and 15% (Māori) lower incidence rates.

3.2. *Length of stay*

Table 3 presents the mean (and 95% CI) total length of stay for patients admitted with ASH. The median and interquartile ranges are presented in Supplementary table 3.

Table 3. The mean length of stay (and 95% confidence intervals) for Ambulatory Sensitive Hospitalization for each of the four bundling methods and Geographic Classification of Health (GCH) categories

	Geographic classification for health			
Bundling method	**U1**	**U2**	**R1**	**R2-3**
Total				
None	2.0 (2.0, 2.0)	1.9 (1.9, 1.9)	2.1 (2.1, 2.1)	2.1 (2.1, 2.1)
Admission source flag	2.4 (2.4, 2.4)	2.2 (2.2, 2.2)	2.6 (2.6, 2.6)	2.8 (2.8, 2.8)
Discharge flag	2.8 (2.8, 2.8)	2.6 (2.6, 2.6)	2.9 (2.9, 2.9)	3.1 (3.1, 3.1)
Dates	2.8 (2.8, 2.8)	2.7 (2.7, 2.7)	3.0 (3.0, 3.0)	3.3 (3.3, 3.3)
Māori				
None	2.0 (2.0, 2.0)	2.0 (2.0, 2.0)	2.3 (2.3, 2.3)	2.2 (2.2, 2.3)
Admission source flag	2.4 (2.4, 2.4)	2.2 (2.2, 2.2)	2.8 (2.8, 2.8)	3.1 (3.1, 3.1)
Discharge flag	2.7 (2.7, 2.7)	2.7 (2.7, 2.7)	3.0 (3.0, 3.0)	3.4 (3.4, 3.4)
Dates	2.8 (2.8, 2.8)	2.8 (2.8, 2.8)	3.1 (3.1, 3.1)	3.3 (3.3, 3.3)
Non-Māori				
None	2.0 (2.0, 2.0)	1.9 (1.9, 1.9)	2.0 (2.0, 2.0)	2.0 (2.0, 2.1)
Admission source flag	2.4 (2.4, 2.4)	2.2 (2.2, 2.2)	2.5 (2.5, 2.5)	2.7 (2.7, 2.7)
Discharge flag	2.8 (2.8, 2.8)	2.6 (2.6, 2.6)	2.9 (2.9, 2.9)	3.0 (3.0, 3.0)
Dates	2.8 (2.8, 2.8)	2.7 (2.7, 2.7)	2.9 (2.9, 2.9)	3.2 (3.2, 3.2)

Table 4. The number and percentage (with 95% confidence intervals) of encounters that included an interhospital transfers when each of the bundling methods are used to group Ambulatory Sensitive Hospitalizations for the four Geographic Classification of Health (GCH) categories

	Geographic classification for health							
	U1		U2		R1		R2-3	
	N	Percent (95% confidence interval)	*N*	Percent (95% confidence interval)	*N*	Percent (95% confidence interval)	*N*	Percent (95% confidence interval)
Total								
Admission source flag	9514	2.2% (2.1%, 2.2%)	5917	3.8% (3.7%, 3.9%)	3446	4.1% (4.0%, 4.3%)	2655	5.7% (5.7%, 5.9%)
Discharge flag	13,559	3.2% (3.0%, 3.2%)	6929	4.5% (4.4%, 4.6%)	4856	5.9% (5.8%, 6.1%)	4137	9.0% (8.8%, 9.3%)
Dates	16,326	3.8% (3.7%, 3.8%)	8000	5.3% (5.2%, 5.4%)	6298	7.8% (7.6%, 8.0%)	5482	12.3% (12.0%, 12.6%)
Māori								
Admission source flag	1492	1.8% (1.7%, 1.9%)	1325	2.8% (2.6%, 2.9%)	725	3.6% (3.4%, 3.9%)	979	5.5% (5.2%, 5.9%)
Discharge flag	2233	2.7% (2.6%, 2.9%)	1636	3.4% (3.3%, 3.6%)	1127	5.7% (5.4%, 6.1%)	1577	9.2% (8.7%, 9.6%)
Dates	2774	3.5% (3.3%, 3.6%)	1951	4.2% (4.0%, 4.3%)	1526	7.9% (7.5%, 8.3%)	2042	12.1% (11.6%, 12.6%)
Non-Māori								
Admission source flag	8022	2.2% (2.2%, 2.3%)	4592	4.3% (4.2%, 4.4%)	2721	4.3% (4.1%, 4.5%)	1676	5.7% (5.5%, 6.0%)
Discharge flag	11,326	3.2% (3.1%, 3.2%)	5293	5.0% (4.9%, 5.2%)	3729	6.0% (5.8%, 6.2%)	2560	8.9% (8.6%, 9.3%)
Dates	13,552	3.9% (3.8%, 3.9%)	6049	5.8% (5.7%, 6.0%)	4772	7.8% (7.6%, 8.0%)	3440	12.4% (12.0%, 12.7%)

Using any bundling method increased the length of stay compared with not bundling events. The dates method resulted in the largest increase in mean total length of stay (between 0.8 and 1.2 days longer). For both Māori and non-Māori, the length of stay for people who lived in rural areas was impacted most by bundling.

3.2.1. Hospital transfers

Table 4 presents the percentage (and 95% CIs) of encounters that contained an interhospital transfer for each of the three bundling methods. In all GCH categories, the dates based bundling method identified the highest percentage of admissions that contained an interhospital transfer. The biggest increases were seen in the R1 and R2-3 GCH categories with an approximately 47% increase in the number of admissions that contained interhospital transfers compared with using the admission source flag. This pattern was similar for Māori and non-Māori, although compared with non-Māori, there were fewer interhospital transfers for Māori that lived in urban areas, but not rural areas. There was a lower percentage of encounters (0.4% to 3.0%) that contained at least one intrahospital transfer compared to encounters that contained at least one interhospital transfer. The dates-based bundling method identified triple the number of encounters with an intrahospital transfer compared to the other two bundling methodologies (supplementary table 4). The results were similar across the four GCH categories.

3.3. Excluding interhospital transfers from ASH encounters

Excluding ASH encounters that included an interhospital transfer reduced the age-sex standardized incidence rate for all GCH categories when any of the three bundling methods were used, with R1 and R2-3 having a disproportionate reduction in incidence rate (Table 5). Using the dates method resulted in a 3% reduction for those living in U1 areas and 10% reduction for those living in the most rural R2-3 category compared with no bundling method. This has the effect of increasing the relative difference between rural and urban areas with slightly lower IRRs for R1 (0.90 vs 0.93) and R2-3 (0.45 vs 0.48). Standardized incidence rates and IRRs are presented in Table 5.

4. Discussion

This NZ-based study using the NMDS demonstrates that there was a considerable reduction (up to 8% for those in R2-3) in the incidence rates of ASH when consecutive

Table 5. The age-sex standardized incidence rate (SIR) and incidence rate ratios (SIRR) with 95% confidence intervals for Ambulatory Sensitive Hospitalizations (ASH) after admissions with more than one facility (interhospital transfers) were excluded

Bundling method	U1	U2	R1	R2-3
Total				
None				
SIR[a]	3062 (3052, 3072)	3683 (3663, 3703)	2891 (2870, 2913)	1530 (1515, 1545)
SIRR[b]	1.00 (1.00, 1.00)	1.20 (1.20, 1.21)	0.94 (0.94, 0.95)	0.50 (0.50, 0.50)
Admission source flag				
SIR	2978 (2968, 2988)	3519 (3500, 3539)	2755 (2733, 2776)	1420 (1406, 1435)
SIRR	1.00 (1.00, 1.00)	1.18 (1.18, 1.19)	0.93 (0.92, 0.93)	0.48 (0.48, 0.48)
Discharge flag				
SIR	2935 (2925, 2944)	3483 (3464, 3503)	2685 (2664, 2706)	1353 (1339, 1367)
SIRR	1.00 (1.00, 1.00)	1.19 (1.18, 1.19)	0.91 (0.91, 0.92)	0.46 (0.46, 0.46)
Dates				
SIR	2868 (2859, 2878)	3407 (3388, 3427)	2593 (2572, 2614)	1280 (1266, 1294)
SIRR	1.00 (1.00, 1.00)	1.19 (1.18, 1.19)	0.90 (0.90, 0.91)	0.45 (0.44, 0.45)
Māori				
None				
SIR	4638 (4606, 4670)	5318 (5271, 5366)	4060 (4004, 4116)	2307 (2273, 2340)
SIRR	1.00 (0.99, 1.01)	1.15 (1.14, 1.15)	0.88 (0.87, 0.88)	0.50 (0.49, 0.50)
Admission source flag				
SIR	4498 (4467, 4530)	5073 (5028, 5120)	3819 (3765, 3873)	2095 (2064, 2127)
SIRR	1.00 (0.99, 1.01)	1.13 (1.12, 1.14)	0.85 (0.84, 0.85)	0.47 (0.46, 0.47)
Discharge flag				
SIR	4423 (4393, 4454)	5001 (4956, 5047)	3677 (3624, 3730)	1967 (1936, 1998)
SIRR	1.00 (0.99, 1.01)	1.13 (1.12, 1.14)	0.83 (0.83, 0.84)	0.44 (0.44, 0.45)
Dates				
SIR	4319 (4289, 4349)	4878 (4833, 4924)	3521 (3469, 3573)	1857 (1827, 1887)
SIRR	1.00 (0.99, 1.01)	1.13 (1.12, 1.14)	0.82 (0.81, 0.82)	0.43 (0.43, 0.43)
Non-Māori				
None				
SIR	2861 (2851, 2871)	3242 (3220, 3265)	2679 (2655, 2704)	1278 (1260, 1295)
SIRR	1.00 (1.00, 1.00)	1.13 (1.13, 1.14)	0.94 (0.93, 0.94)	0.45 (0.44, 0.45)
Admission source flag				
SIR	2783 (2773, 2794)	3095 (3072, 3118)	2558 (2534, 2582)	1195 (1178, 1212)
SIRR	1.00 (1.00, 1.00)	1.11 (1.11, 1.12)	0.92 (0.92, 0.92)	0.43 (0.43, 0.43)
Discharge flag				
SIR	2743 (2733, 2753)	3066 (3044, 3089)	2501 (2477, 2525)	1144 (1128, 1161)
SIRR	1.00 (1.00, 1.00)	1.12 (1.11, 1.12)	0.91 (0.91, 0.92)	0.42 (0.42, 0.42)
Dates				
SIR	2681 (2671, 2691)	3001 (2978, 3023)	2419 (2396, 2443)	1083 (1067, 1100)
SIRR	1.00 (1.00, 1.00)	1.12 (1.12, 1.12)	0.90 (0.90, 0.91)	0.40 (0.40, 0.41)

Rates and IRRs are shown using the four bundling methods and by the four Geographic Classification for Health categories. U1 was used as the reference category.

[a] SIRR: Age-sex standardized incidence rates per 100,000 person years. Standardized to New Zealand Māori 2018 Census Estimated Resident Population.

[b] SIR: Standardized incidence rate ratio.

hospital events are bundled into encounters. The dates method, where events that occur 1 day or less apart are grouped together to create encounters, resulted in the biggest reduction in the incidence rates for ASH compared to using the admission or discharge flags. However, perhaps surprisingly, the interpretation of the relative rural-urban IRRs did not differ based on the bundling method used.

Bundling episodes using the dates method into encounters also lengthened the mean total length of stay compared to other methods, especially for patients who reside in the rural areas of NZ. The percentage of hospital encounters that contained an interhospital transfer was also increased. Excluding these interhospital transfers reduced the incidence rates, especially for rural areas, which resulted in slightly reduced incidence rate ratios for the R1 and R2-3 compared with U1 and U2 areas. The overall trends that were observed in this study were similar for both Māori and non-Māori.

4.1. Strengths and weaknesses

A strength of this study is being able to use the comprehensive NMDS as a consistent national data source. Researchers in other health jurisdictions, where individual patients do not have unique identifiers, have had to develop complex methods of grouping events into encounters, especially when there have been interhospital transfers [4,6,7]. While the use of the NMDS is an advantage, the results of this study may have limited generalizability in other countries, where a similar dataset is not available. However, the principles of grouping related and consecutive events into encounters remain important and relevant.

While ASH is commonly used as an important health systems indicator [31,32,37], many of the diagnoses in the ASH definition may result in brief admissions that would not involve an interhospital transfer or movement within the same facility (‘‘statistical discharge’’). Other diagnoses may result in more transfers, which could result in larger rural-urban differences between the bundling techniques.

Only primary diagnosis codes were considered to assign an ASH to an encounter. It has been demonstrated that all primary and secondary diagnoses codes need to be considered for analyses that include acute myocardial infarction (a condition i.e. part of the ASH definition) [38], although it has not been similarly established that this is the case for other diagnoses. It is also unclear which event diagnosis (eg, the first, last or any) or diagnoses within the bundled encounter to use. Additionally, there may be differences in coding practices between urban and rural hospitals, with rural hospitals anecdotally less likely to have access to trained coding professionals [14]. The category of hospital of admission (rural vs. urban) was not considered in this study. While it is unlikely that ASH conditions result in lengthy admissions, the choice of bundling method might influence the mean length-of-stay. Where the condition of interest may result in very long admissions, these potential outliers should be explicitly accounted for.

Ethnicity data is poorly recorded in the NMDS, with Māori undercounted by 16% [39]. This has implication for interpreting outcomes for Māori, especially for more rural communities where the percentage of Māori are higher.

4.2. Policy implications

It is clearly important that health researchers define or explain how multiple events within a single encounter are dealt with in their analyses. If multiple events are not accounted for an overcount of the incidence of some conditions (ASH in this example) may occur. Additionally, bundling events into encounters is required to accurately identify the date or facility of admission and failure to do so may influence the results of studies where this is of importance, such as identifying interhospital transfers [2], readmission rates or time to event studies [12,18]. When these transfers are excluded from analyses [17], our study showed that patients that live in more rural areas will be disproportionally affected, which may miss or underestimate important rural-urban health disparities. Some interhospital transfers or statistical discharges might also be wrongly identified as unintended readmissions if they are not properly accounted for. Depending on the research question, bundling of events into encounters should occur prior to applying sampling techniques to large datasets.

Using the date method to group consecutive events into encounters is consistent with the international best-practice and has been shown in this study to result in the lowest, and likely most realistic incidence rates compared to other methods [3,6,7,10,40]. Compared to using other methods, bundling by dates has been demonstrated to reduce the overcounting of incidence rates by up to 30% depending on the diagnosis of interest [3,6,7,40]. The date method might make identifying unintended early readmissions more difficult but relying on admission or discharge flags is also likely to be unreliable [18].

Custodians of administrative health data that contains hospital discharge records could consider including a date-based algorithm or improve the coding of admission or discharge flags (which occurs manually, and therefore prone to error or misinterpretation) to enable consecutive events to be bundled into encounters. It should be easy to differentiate physical from administrative transfers. For injury related admissions, a date of injury is already recorded in the NMDS. This makes it possible to identify related admissions that can occur several months after the initial injury [1,4].

Future research could focus on the effect of different bundling techniques on the rural-urban incidence rates of other health conditions, the impact of bundling events into encounters in other health datasets (eg, national nonadmitted patient collection) and examine the impact of considering both primary and secondary diagnosis codes when identifying conditions of interest.

5. Conclusion

This study identifies variation in the GCH stratified incidence rates of ASH in NZ of up to 8.3% depending on the bundling method. However, this did not result in a change in the interpretation of rural-urban differences for ASH. The date-based method of bundling resulted in an increase in the mean length of stay and percentage of ASH encounters that required interhospital transfers. Patients who live in the most rural communities were affected most. Researchers should carefully consider how to deal with events and encounters when using hospital discharge data.

Ethics

Ethics approval was obtained from the University of Otago Human Research Ethics Committee (HD19/069).

CRediT authorship contribution statement

Rory Miller: Writing − review & editing, Writing − original draft, Visualization, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gabrielle Davie:** Writing − review & editing, Validation, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Sue Crengle:** Writing − review & editing, Validation, Methodology, Funding acquisition, Conceptualization. **Jesse Whitehead:** Writing − review & editing, Funding acquisition, Conceptualization. **Brandon De Graaf:** Writing − review & editing, Visualization, Funding acquisition, Conceptualization. **Garry Nixon:** Writing − review & editing, Validation, Supervision, Funding acquisition, Conceptualization.

Data availability

The data used in this study is not available. Requests for the dataset can be made to Te Whatu Ora (https://www.tewhatuora.govt.nz/our-health-system/data-and-statistics/how-to-access-data/).

Declaration of competing interest

There are no competing interests for any author.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2024.111400.

References

[1] New Zealand Ministry of Health. Ministry of Health NZ. National Minimum Dataset (hospital events); 2021. Available at: https://www.health.govt.nz/nz-health-statistics/national-collections-and-surveys/collections/national-minimum-dataset-hospital-events. Accessed October 1, 2023.

[2] Cormack D, Masters-Awatere B, Lee A, Rata A, Boulton A. Understanding the context of hospital transfers and away-from-home hospitalisations for Māori. N Z Med J 2022;135(1565):10.

[3] Davie G, Samaranayaka A, Langley JD, Barson D. Estimating person-based injury incidence: accuracy of an algorithm to identify readmissions from hospital discharge data. Inj Prev 2011;17(5):338−42.

[4] Vallmuur K, McCreanor V, Cameron C, Watson A, Shibl R, Banu S, et al. Three Es of linked injury data: episodes, encounters and events. Inj Prev 2021;27(5):479.

[5] Lopez D, Katzenellenbogen JM, Sanfilippo FM, Woods JA, Hobbs MS, Knuiman MW, et al. Transfers to metropolitan hospitals and coronary angiography for rural Aboriginal and non-Aboriginal patients with acute ischaemic heart disease in Western Australia. BMC Cardiovasc Disord 2014;14(1):58.

[6] Fransoo R, Yogendran M, Olafson K, Ramsey C, McGowan KL, Garland A. Constructing episodes of inpatient care: data infrastructure for population-based research. BMC Med Res Methodol 2012 Dec;12:133.

[7] Lopez D, Nedkoff L, Knuiman M, Hobbs MST, Briffa TG, Preen DB, et al. Exploring the effects of transfers and readmissions on trends in population counts of hospital admissions for coronary heart disease: a Western Australian data linkage study. BMJ Open 2017;7(11):e019226.

[8] Westfall JM, McGloin J. Impact of double counting and transfer bias on estimated rates and outcomes of acute myocardial infarction. Medical Care 2001;39:459−68.

[9] Westfall JM. Methods for handling inter-hospital transfer in acute myocardial infarction research. Clin Med Cardiol 2008;2:CMC.S331.

[10] Kerr AJ, Lee M, Jiang Y, Grey C, Wells S, Williams M, et al. High level of capture of coronary intervention and associated acute coronary syndromes in the all New Zealand acute coronary syndrome quality improvement cardiac registry and excellent agreement with national administrative datasets (ANZACS-QI 25). N Z Med J 2019;132(1492):19−29.

[11] Chan WC, Wright C, Tobias M, Mann S, Jackson R. Explaining trends in coronary heart disease hospitalisations in New Zealand: trend for admissions and incidence can be in opposite directions. Heart 2008;94:1589−93.

[12] Rumball-Smith J, Sarfati D, Hider P, Blakely T. Ethnic disparities in the quality of hospital care in New Zealand, as measured by 30-day rate of unplanned readmission/death. Int J Qual Health Care 2013;25:248−54.

[13] Lee S, Miller R, Lee M, White H, Kerr AJ. Outcomes after ST-elevation myocardial infarction presentation to hospitals with or without a routine primary percutaneous coronary intervention service (ANZACS-QI 46). N Z Med J 2020;133(1524):18.

[14] Miller R, Nixon G, Turner RM, Stokes T, Keenan R, Grey C, et al. Outcomes and access to angiography following non-ST segment elevation acute coronary syndromes in patients who present to rural or urban hospitals: ANZACS-QI 72. N Z Med J 2023;136(1573).

[15] Seibt S, Gilchrist CA, Reed PW, Best EJ, Harnden A, Camargo CA, et al. Hospital readmissions with acute infectious diseases in New Zealand children $<$ 2 years of age. BMC Pediatr 2018;18(1):98.

[16] Jeffs E, Williman J, Brunton C, Gullam J, Walls T. The epidemiology of listeriosis in pregnant women and children in New Zealand from 1997 to 2016: an observational study. BMC Public Health 2020;20:116.

[17] Baker MG, Barnard LT, Kvalsvig A, Verrall A, Zhang J, Keall M, et al. Increasing incidence of serious infectious diseases and inequalities in New Zealand: a national epidemiological study. Lancet 2012;379(9821):1112−9.

[18] Woods TJ, Ngo L, Speck P, Kaambwa B, Ranasinghe I. Thirty-day unplanned readmissions following hospitalisation for atrial fibrillation in Australia and New Zealand. Heart Lung Circ 2022;31(7):944−53.

[19] Woelk V, Speck P, Kaambwa B, Fitridge RA, Ranasinghe I. Incidence and causes of early unplanned readmission after hospitalisation with peripheral arterial disease in Australia and New Zealand. Med J Aust 2022;216:80−6.

[20] Hariharaputhiran S, Peng Y, Ngo L, Ali A, Hossain S, Visvanathan R, et al. Long-term survival and life expectancy following an acute heart failure hospitalization in Australia and New Zealand. European J of Heart Fail 2022;24(9):1519−28.

[21] Whitehead J, Davie G, de Graaf B, Crengle S, Fearnley D, Smith M, et al. Defining rural in Aotearoa New Zealand: a novel geographic classification for health purposes. N Z Med J 2022;135(1559):24−40.

[22] Whitehead J, Blattner K, Miller R, Crengle S, Ram S, Walker X, et al. Defining catchment boundaries and their populations for Aotearoa New Zealand's rural hospitals. J Prim Health Care 2023;15(1):14−23.

[23] Miller R, Stokes T, Nixon G. Point-of-care troponin use in New Zealand rural hospitals: a national survey. N Z Med J 2019;132(1493):13.

[24] Blattner K, Stokes T, Nixon G. A scope of practice that works "out here": exploring the effects of a changing medical regulatory environment on a rural New Zealand health service. Rural Rem Health 2019;19(5442).

[25] Ranasinghe I, Barzi F, Brieger D, Gallagher M. Inter-hospital transfer (IHT) for acute myocardial infarction (AMI): a geospatial analysis of transfer patterns. Heart Lung Circ 2012;21:S253−4.

[26] Ministry of health: manatū Hauora. Ministry of health NZ. 2021. National health index questions and answers. Available at: https://www.health.govt.nz/our-work/health-identity/national-health-index/nhi-information-health-consumers/national-health-index-questions-and-answers. Accessed October 1, 2023.

[27] World Health Orgranization. International statistical classification of diseases and related health problems. 10th revision, Fifth edition. Geneva, Switzerland: World Health Organization; 2016. Available at: https://icd.who.int/browse10/Content/statichtml/ICD10Volume2_en_2019.pdf. Accessed October 1, 2023.

[28] Health Information Standards Organisation. Ethnicity Data Protocols. NZ Ministry of Health; 2017. Available at: https://www-health-govt-nz.ezproxy.otago.ac.nz/system/files/documents/publications/hiso-10001-2017-ethnicity-data-protocols-v2.pdf. Accessed October 1, 2023.

[29] StatsNZ :tatauranga Aotearoa. 2018 census. 2020. Available at: https://www.stats.govt.nz/2018-census/. Accessed October 1, 2023.

[30] Ministry of Health. Saving Lives: Amenable mortality in New Zealand, 1996 - 2006. Wellington: MInistry of Health; 2010. Available at: https://www.health.govt.nz/publication/saving-lives-amenable-mortality-new-zealand-1996-2006. Accessed October 1, 2023.

[31] METEOR: Metadata Online Registry. National Healthcare Agreement: PI 18−Selected potentially preventable hospitalisations, 2019. Australian Institute of Health and Welfare; 2019. Available at: https://meteor.aihw.gov.au/content/698904. Accessed October 1, 2023.

[32] Ministry of Health: Manatū Hauora. Ambulatory sensitive (avoidable) hospital admissions. MInistry of Health. Wellington; 2023. Available at: https://nsfl.health.govt.nz/accountability/performance-and-monitoring/data-quarterly-reports-and-reporting/ambulatory-sensitive. Accessed October 1, 2023.

[33] Nixon G, Davie G, Whitehead J, De Graaf B, Crengle S, Lawrenson R, et al. GCH:Domicile Concordance File (MoH-HDOM-t-GCH2018 2024. Available at: https://asset.turboweb.co.nz/219/file/ap9ernxp1ffwye07lhkm/MoH_HDOM_to_GCH2018.xlsx. Accessed October 5, 2024.

[34] StatsNZ: Tatauranga Aotearoa. Statistical standard for geographic areas 2023. Available at: 2022. https://www.stats.govt.nz/methods/statistical-standard-for-geographic-areas-2023/. Accessed October 1, 2023.

[35] Newman S. Standardized rates and age−period−cohort analysis. In: Biostatistical Methods in Epidemiology. John Wiley & Sons, Ltd; 2002:249−61. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/0471272612.ch12. Accessed October 1, 2023.

[36] Crengle S, Davie G, Whitehead J, de Graaf B, Lawrenson R, Nixon G. Mortality outcomes and inequities experienced by rural Māori in Aotearoa New Zealand. Lancet Reg Health West Pac 2022;28:100570.

[37] Becker NV, Karmakar M, Tipirneni R, Ayanian JZ. Trends in hospitalizations for ambulatory care−sensitive conditions during the COVID-19 pandemic. JAMA Netw Open 2022;5(3):e222933.

[38] Kerr AJ, Wang TKM, Jiang Y, Grey C, Wells S, Poppe KK. The importance of considering both primary and secondary diagnostic codes when using administrative health data to study acute coronary syndrome epidemiology (ANZACS-QI 47). Eur Heart J Qual Care Clin Outcomes 2020;7(qcaa056):548−55.

[39] Harris RB, Paine SJ, Atkinson J, Robson B, King PT, Randle J, et al. We still don't count: the under-counting and under-representation of Māori in health and disability sector data. N Z Med J 2022;135(1567):54−78.

[40] Du W, Hayen A, Finch C, Hatfield J. Comparison of methods to correct the miscounting of multiple episodes of care when estimating the incidence of hospitalised injury in child motor vehicle passengers. Accid Anal Prev 2008;40(4):1563−8.