

# Potential Errors in Health Disparities Research Resulting from Lack of Unique Patient Identifiers: Analysis of Diabetes-related Preventable Hospitalizations

Hyeong Jun Ahn PhD

## Abstract

All-payer, population-level hospital discharge data have been used to identify health disparities across racial/ethnic and other demographic groups. However, researchers are often unable to identify unique patients in the data sets if a unique patient identifier is not provided. The lack of the unique patient identifier can result in biased estimates of research outcomes using discharge data. This could then mislead the researchers, public, or policy-makers who utilize such biased results. This study examined estimation bias of health disparities due to rehospitalizations considering diabetes-related preventable hospitalizations using 6 years of state-level data from Hawai'i Health Information Corporation. Different analyses methods showed different probabilities of having multiple visits by age, race/ethnicity and payer subgroups. Charge analysis results also showed that ignoring the multiple visits could result in significance error. For a patient with multiple hospitalizations, rehospitalizations are often dependent upon the discharge status of previous visits, and the independence assumption of the multiple visits may not be appropriate. Ignoring the multiple visits in population-level analyses could result in severe health disparities significance errors. In this hospitalization charge analysis, the Chinese group was not significantly different than the White group (relative risk ratio - RR: [95% CI]: 0.93 [0.80, 1.08]), while the difference was significant (RR [95% CI]: 0.86 [0.77, 0.96]) when the multiple visits were ignored.

## Keywords

Multiple visits, significance error, unique patient identifier, GEE, hospitalization analysis

## Abbreviations

AHRQ = Agency for Healthcare Research and Quality  
AOR = adjusted odds ratio  
ARR = adjusted relative risk ratio  
DOD = Department of Defense  
DRPH = diabetes-related potentially preventable hospitalizations  
GEE = generalized estimating equation  
HHIC = Hawai'i Health Information Corporation  
ICD-9 = International Classification of Diseases – 9th revision  
– Clinical Modification  
OR = unadjusted odds ratio  
RR = unadjusted relative risk ratio

## Introduction

Diabetes is one of the most common chronic diseases in the United States (US). About 34.2 million people or 10.5% of the US population had diabetes in 2020.<sup>1</sup> Diabetes is present in almost 1 in 5 (19.4%) of hospitalizations in the general US population,<sup>2</sup> and more than 20% of patients with diabetes experience hospitalization each year. About 8.3% of adults with diabetes had multiple hospitalizations in 1988.<sup>3</sup>

Many diabetes-related hospitalizations are considered avoidable with good outpatient care.<sup>4</sup> Inpatient stays for uncontrolled diabetes, short-term and long-term diabetes complications, and lower-extremity amputations are specifically classified by the Agency for Healthcare Research and Quality (AHRQ) as diabetes-related potentially preventable hospitalizations (DRPH).<sup>5</sup> Decreasing such preventable hospitalizations is expected to result in improved quality of care and reduced health care costs.<sup>5-6</sup>

Important patterns exist among patients with diabetes. The percentage of people who are aged 65 years or older with diabetes was over 6 times higher than that among people aged 20 to 44 years (25.9% vs. 4.1%).<sup>7</sup> Racial and ethnic differences also exist. For instance, Native Hawaiians have higher rates of diabetes compared to other groups.<sup>8-13</sup>

All-payer, population-level hospital discharge data have been used to identify health disparities across racial/ethnic groups. One limitation in many all-payer, population-level discharge datasets is that researchers are unable to identify unique patients in the datasets. For instance, the widely-used Nationwide Inpatient Sample data from Healthcare Cost and Utilization Data<sup>14</sup> does not identify unique individuals due to privacy concerns, and thus cannot account for multiple hospitalizations by the same individuals in the estimation of rates and disparities.<sup>15</sup> Only 20 out of 47 participating states (43%) reported variables to track sequential visits, within or across facilities and hospitals for a patient within the state in their 2019 state-level inpatient databases. Even for states that have invested considerable time and money to create such unique identifiers, these data have not currently been widely used in health service research, especially across multiple years.

The lack of a unique patient identifier can result in biased estimates of research outcomes using discharge data. This could then mislead researchers, the public, or policymakers. For instance, the hospitalization patterns of patients with single visits might be quite different from that of patients with multiple visits. Significantly higher odds of having multiple stays were found for Hispanics and non-Hispanic Blacks relative to non-Hispanic Whites ( $P < .0001$ ), Medicare or Medicaid patients compared with privately insured ( $P < .0001$ ), and patients in low-income areas ( $P < .05$ ).<sup>16</sup> If similar issues are true across many commonly reported outcomes with discharge data, many existing analyses that do not account for these biases may misstate or misestimate health disparities.

Hawai'i state-level hospitalization data includes a unique patient identifier across multiple years. The goal of this study was to use 6 years of Hawai'i state-level population-level data to investigate the impact of not identifying unique patients related to the extent of racial/ethnic health disparities. The first goal of this study is to explore any errors in significance when multiple visits are not considered as well as patterns of multiple DRPHs. The second goal is to estimate the parameters using a generalized estimating equation (GEE). For more illustration, the charges of diabetes hospitalizations are explored as a practical example of significance error that could be seen due to the lack of unique patient identification in a health disparity investigation.

## Methods

Hawai'i Health Information Corporation (HHIC) inpatient data from 2007-2012 (n=640,824) was used, which includes detailed discharge information for all hospitalizations from all payers in Hawai'i. The HHIC data includes data on race/ethnicity of patients, insurer, age, gender, and International Classification of Diseases – 9th revision – Clinical Modification (ICD-9) codes. Long-term care and psychiatric hospital visits were excluded. The HHIC data has been used as the Hawai'i hospital data source for the major national inpatient database.<sup>4</sup>

DRPH were defined with AHRQ criteria using ICD-9 diagnosis and procedure codes including: (1) uncontrolled diabetes without mentioning of a short-term or long-term complication (ICD-9-CM principal diagnosis codes 250.02-250.03); (2) diabetes with short-term complications, eg, ketoacidosis, hyperosmolarity, coma (ICD-9-CM principal diagnosis codes 250.1-250.33); (3) diabetes with long-term complications, eg, renal, eye, neurological, circulatory, or complications not otherwise specified (ICD-9-CM principal diagnosis codes 250.4-250.93); and (4) lower-extremity diabetes-related amputations based on ICD-9 and procedure codes ICD-9-CM procedure codes for lower-extremity amputation in any field and diagnosis code for diabetes in any field.<sup>4</sup> If a trauma diagnosis code was in any field, the amputation was not considered a DRPH. As DRPH definitions generally exclude pregnancy, childbirth, and puerperium hospitalization, those by individuals under 18 years, and those transferring from another institution, these visits were excluded in the study as well. Additionally, 433 Tripler Army Medical Center (Tripler) visit records were excluded due to unknown race/ethnicity information. While most patients with Department of Defense (DOD) insurance are likely to visit Tripler, the DOD patients who were admitted to other hospitals might have Tripler hospitalizations that were excluded from this study. A total of 7 652 records from 4 964 patients were used for analysis focusing DRPH visits with the exclusion criteria.

The HHIC race/ethnicity variable was created from the race/ethnicity categories available consistently across all hospitals in Hawai'i. Only 1 primary race is reported across all hospitals, typically from patient self-report at intake. Mixed-race individuals are represented as only their primary race. Other variables, such as sex (male/female based on administrative data), age (grouped in categories 18-39; 40-64, 65+), payer (DOD, Medicare, Medicaid, Private, and Other), location of residence (living on O'ahu or another Hawaiian island) and substance use (Yes or No) were also included in our multivariable analyses. Hospital charges in dollars were included in the HHIC data.

Choosing 1 visit per person is essential for patient level analysis as 1 record per patient needs to be analyzed to avoid certain issues with multiple records. For example, for racial disparities, 1 race record for a patient needs to be selected. If a patient keeps changing their race over time, it will result in inconsistencies among the different race/ethnicity groups. Therefore, selecting 1 record per patient is important for patient level analysis. Three different options were examined (first, last, and random) to investigate any significant difference. To consider these options, visits by the same patient were sorted by discharge date with the first and last visits easily identified. A random number generator was used to select a single visit for those patients with multiple visits.

## Statistical Methods

Multivariable logistic regressions were used to compare the first, last and a random visit from individuals with multiple visits versus a single visit, adjusted for age, sex, race/ethnicity, payer type, location of residence and substance use. All hospitalizations ignoring repeated visits were also analyzed by the same multivariable logistic regression while GEE models were used to consider the unstructured correlation structure among repeated visits within each patient. Adjusted odds ratios (AOR) with 95% confidence intervals (CIs) were obtained from the logistic regression and GEE models.

Three different approaches were used for hospitalization charge analysis using charge information for: (1) all hospitalizations without considering repeated measures; (2) the first visit for a patient; (3) all hospitalizations for a patient considering repeated measures. Multivariable gamma regression models with log link were used to predict hospitalization charges by race/ethnicity adjusting for other factors such as age, sex, race/ethnicity, payer type and location of residence.<sup>17</sup> Adjusted relative risk ratios (ARRs) of charges with 95% CIs were estimated. Analyses were conducted using SAS version 9.3 (SAS Institute, Cary, North Carolina) and two-tailed *P*-value of less than .05 was regarded as statistically significant. Because HHIC data are de-identified, analysis does not involve human subjects so Institutional Review Board (IRB) review was not sought.

## Results

### Single Visit Analysis

Twenty-two percent of the DRPH patients (n=1084) had multiple visits while 3880 patients had single visit. In multivariable logistic analysis using the first visits, younger patients had more multiple hospitalizations than older age groups (65+ years) (AOR [95% CI]: 4.04 [3.07, 5.30] for 18-39 years; 2.18 [1.77, 2.67] for 40-64 years), and patients with Medicaid or Medicare insurance had more multiple hospitalizations than patients with private insurance (AOR [95% CI]: 1.47 [1.20, 1.80] for Medicaid; 1.59 [1.28, 1.98] for Medicare) (**Table 1**, second column). Native Hawaiians and other Pacific Islanders were not significantly different from Whites, while Asians were less likely to have multiple hospitalizations than Whites (AOR [95% CI]: 0.52 [0.34, 0.82] for Chinese; 0.68 [0.54, 0.87] for Japanese; 0.68 [0.52, 0.89] for Filipino; 0.59 [0.37, 0.93] for other Asian). Sex and substance use were not statistically significant in predicting DRPH multiple hospitalizations.

Multivariable adjusted last visit analysis results were similar to first visit analyses (**Table 1**, third column). Younger patients also had more multiple hospitalizations than older age groups (65+ years) (AOR [95% CI]: 4.28 [3.27, 5.59] for 18-39 years; 2.33 [1.91, 2.84] for 40-64 years), and patients with DOD insurance, Medicaid and Medicare had more multiple hospitalizations than private insurance (AOR [95% CI]: 3.53 [1.63, 7.67] for DOD; 3.40 [2.03-5.70] for Medicaid; 4.68 [2.76-7.93] for Medicare). Native Hawaiians were more likely to have multiple hospitalizations compared to Whites (AOR [95% CI]: 1.24 [1.02, 1.51]) while Chinese and Japanese were less likely to have multiple hospitalizations (AOR [95% CI]: 0.61 [0.39, 0.93] for Chinese; 0.72 [0.57-0.92] for Japanese). Filipino, other Pacific Islanders and other Asian were not significantly different from Whites. Sex was not significantly different for DRPH multiple hospitalizations, but substance use did differ significantly with those who had substance use more likely to have a multiple hospitalization (AOR [95% CI]: 1.42 [1.10, 1.85]).

In multivariable adjusted random visit analyses (**Table 1**, fourth column), younger patients had more multiple hospitalizations than older age groups (65+ years) (AOR [95% CI]: 4.13 [3.15, 5.41] for 18-39 years; 2.29 [1.87, 2.81] for 40-64 years), and patients with Medicaid and Medicare had more multiple hospitalizations than those with private insurance (AOR [95% CI]: 2.06 [1.36, 3.13] for Medicaid; 2.31 [1.49, 3.56] for Medicare). Native Hawaiians and other Pacific Islanders were not significantly different from Whites, while Asians had fewer multiple hospitalizations than Whites (AOR [95% CI]: 0.55 [0.36, 0.85] for Chinese; 0.68 [0.54, 0.86] for Japanese; 0.72 [0.55, 0.93] for Filipino; 0.60 [0.38, 0.95] for other Asian). Sex and substance use were not significantly different for DRPH multiple hospitalizations.

### All Hospitalizations Analysis

The analysis using all hospitalizations was conducted without consideration of repeated measures (**Table 1**, fifth column). Females were significantly more likely to have multiple hospitalizations than males (AOR [95% CI]: 1.15 [1.04, 1.27]). Other Pacific Islanders were significantly less likely to have multiple hospitalizations compared from Whites (AOR [95% CI]: 0.77 [0.63, 0.94]). Age and payer showed similar pattern with other analyses.

GEE model analysis results, which statistically incorporate multiple visits, provide the more accurate significance (**Table 1**, last column). The 2 notable differences with results from other analyses present were: (1) Chinese was not significantly different from Whites (AOR [95% CI]: 0.68 [0.33, 1.40]) and (2) substance use was significantly associated with multiple DRPH (AOR [95% CI]: 1.60 [1.25, 2.05]). Younger patients also had more multiple hospitalizations than older age groups (65+ years) (AOR [95% CI]: 6.95 [5.12, 9.44] for 18-39 years; 2.97 [2.38, 3.69] for 40-64 years), and patients with Medicaid and Medicare were more likely to have multiple hospitalizations compared to those with private insurance (AOR [95% CI]: 3.03 [2.08, 4.43] for Medicaid; 2.50 [1.67, 3.73] for Medicare). Native Hawaiians and other Pacific Islanders were not significantly different than Whites, while Japanese, Filipino and other Asians were less likely to have multiple hospitalizations than Whites (AOR [95% CI]: 0.64 [0.49, 0.84] for Japanese; 0.69 [0.51, 0.95] for Filipino; 0.64 [0.41, 0.99] for other Asian). Sex was not significantly different for DRPH multiple hospitalizations.

### Charge Analysis

Hospital charge analysis for preventable diabetes hospitalizations was used as a practical example to further illustrate potential errors in significance (**Figure 1**). Hospital charges for Chinese were significantly lower than those of Whites in hospitalization level analysis without knowing patient identification by assuming independence among multiple visits within a patient (ARR [95% CI]: 0.86 [0.77, 0.96] for Method 1). However, hospital charges for Chinese were not significantly different than those of Whites when the other 2 methods were used: Method 2: patient level analysis using first admissions (ARR [95% CI]: 0.92 [0.81, 1.05]) and Method 3: hospitalization level analysis with repeated measure consideration (ARR [95% CI] 0.93 [0.80, 1.08]).

Table 1. Adjusted Odds Ratios for Multiple Diabetes Potentially-Related Hospitalizations in Hawai'i, 2007-2012					
	Patient level			All Hospitalizations without repeat measures <sup>d</sup>	All Hospitalizations with repeat measures <sup>e</sup>
	First Hospitalization <sup>a</sup>	Last Hospitalization <sup>b</sup>	Random Hospitalization <sup>c</sup>		
	AOR [95% CI]	AOR [95% CI]	AOR [95% CI]	AOR [95% CI]	AOR [95% CI]
<b>Gender</b>					
Female	1.03 [0.90, 1.19]	1.02 [0.89, 1.18]	1.01 [0.88, 1.17]	1.15 [1.04, 1.27]*	1.15 [0.96, 1.37]
Male	Reference	Reference	Reference	Reference	Reference
<b>Age</b>					
18-39	4.04 [3.07, 5.30] *	4.28 [3.27, 5.59] *	4.13 [3.15, 5.41] *	6.95 [5.74, 8.42] *	6.95 [5.12, 9.44] *
40-64	2.18 [1.77, 2.67] *	2.33 [1.91, 2.84] *	2.29 [1.87, 2.81] *	2.97 [2.57, 3.43] *	2.97 [2.38, 3.69] *
65+	Reference	Reference	Reference	Reference	Reference
<b>Race/Ethnicity</b>					
Chinese	0.52 [0.34, 0.82] *	0.61 [0.39, 0.93] *	0.55 [0.36, 0.85] *	0.68 [0.52, 0.89] *	0.68 [0.33, 1.40]
Filipino	0.68 [0.52, 0.89] *	0.84 [0.64, 1.08]	0.72 [0.55, 0.93] *	0.69 [0.58, 0.83] *	0.69 [0.51, 0.95] *
Hawaiian	1.16 [0.95, 1.41]	1.24 [1.02, 1.51] *	1.15 [0.95, 1.40]	1.09 [0.94, 1.25]	1.09 [0.88, 1.35]
Japanese	0.68 [0.54, 0.87] *	0.72 [0.57, 0.92] *	0.68 [0.54, 0.86] *	0.64 [0.54, 0.76] *	0.64 [0.49, 0.84] *
Other Asian	0.59 [0.37, 0.93] *	0.64 [0.40, 1.02]	0.60 [0.38, 0.95] *	0.64 [0.47, 0.86] *	0.64 [0.41, 0.99] *
Other Pacific Islander	0.93 [0.71, 1.23]	1.28 [0.97, 1.67]	1.02 [0.78, 1.33]	0.77 [0.63, 0.93] *	0.77 [0.56, 1.04]
Other race	0.80 [0.61, 1.04]	0.76 [0.58, 1.01]	0.70 [0.53, 0.92] *	0.78 [0.64, 0.94] *	0.78 [0.58, 1.04]
White	Reference	Reference	Reference	Reference	Reference
<b>Payer</b>					
Department of Defense	1.09 [0.57, 2.08]	3.53 [1.63, 7.67] *	1.77 [0.86, 3.64]	1.62 [0.95, 2.78]	1.62 [0.75, 3.49]
Medicaid/Quest	1.47 [1.20, 1.80] *	3.40 [2.03, 5.70] *	2.06 [1.36, 3.13] *	3.03 [2.28, 4.04] *	3.03 [2.08, 4.43] *
Medicare	1.59 [1.28, 1.98] *	4.68 [2.76, 7.93] *	2.31 [1.49, 3.56] *	2.50 [1.86, 3.37] *	2.50 [1.67, 3.74] *
Other	0.97 [0.66, 1.42]	2.17 [1.29, 3.66] *	1.25 [0.82, 1.91]	1.51 [1.13, 2.01] *	1.51 [1.01, 2.24] *
Private	Reference	Reference	Reference	Reference	Reference
<b>Living O'ahu</b>					
Yes	0.99 [0.85, 1.16]	0.95 [0.81, 1.11]	0.99 [0.84, 1.15]	0.94 [0.84, 1.05]	0.94 [0.77, 1.14]
No	Reference	Reference	Reference	Reference	Reference
<b>Substance Use</b>					
Yes	1.12 [0.85, 1.47]	1.42 [1.10, 1.85] *	1.01 [0.77, 1.34]	1.60 [1.33, 1.92] *	1.60 [1.25, 2.05] *
No	Reference	Reference	Reference	Reference	Reference

\* = significant results with *P* value < .05

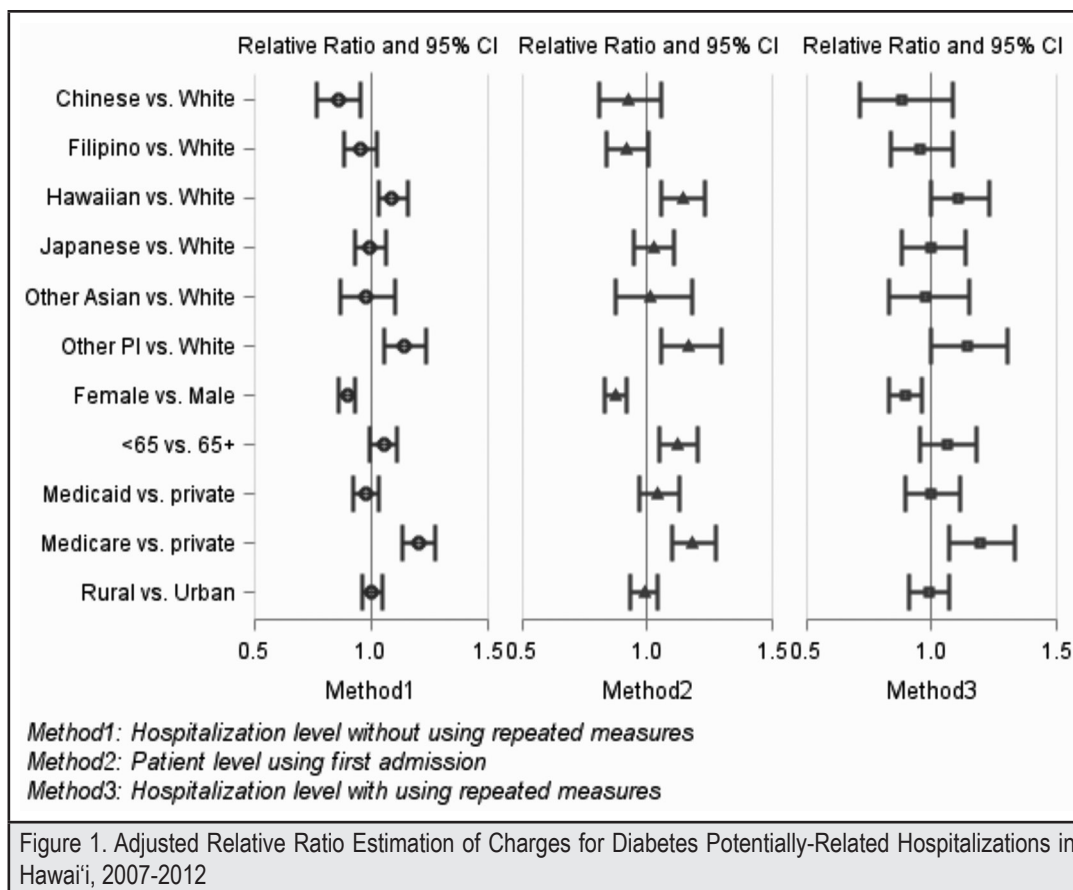
<sup>a</sup> First Hospitalization: the first visit out of multiple hospitalizations per patient.

<sup>b</sup> Last Hospitalization: the last visit out of multiple hospitalizations per patient.

<sup>c</sup> Random Hospitalizations: a randomly selected visit out the multiple hospitalizations per patient.

<sup>d</sup> All Hospitalizations without repeat measures: all visits were used without identifying patients.

<sup>e</sup> All Hospitalizations with repeat measures: all multiple visits were used with patient identification.



## Discussion

Different analytic methods showed different probabilities of having multiple visits by age, race/ethnicity, and payer subgroups. It was common for younger patients to have more multiple hospitalizations than older age groups (65+ years). There were more multiple hospitalizations for patients with Medicaid and Medicare than with private insurance. However, some results were dramatically different among the 5 analyses. For example, in contrast with results from other analyses, in the analysis using the last visit, DRPH multiple hospitalizations showed significant differences between Hawaiians and Whites. In the analysis using repeated measures which generates statistically more accurate significance levels, Chinese DRPH multiple hospitalizations were not significantly different from Whites, while the other 4 analyses all showed that the Chinese had significantly fewer multiple hospitalizations than Whites. The analysis results of substance use using repeated measures also showed different patterns than those with first and random visit analyses. The selection of certain visits to represent patient-level hospitalizations is beyond this study scope as the study goal was to illustrate that significant differences exist between single visits and multiple visits for DRPH stays irrelevant of the selection of visits or repeated measures.

This study highlights the importance of patient identification in using hospitalization data. When patient identification is unknown, researchers have to assume independence among visits of the same patient. That is, if a patient was readmitted, the new admission wouldn't be connected with previous visits and the significance corresponding to "all hospitalizations without repeat measures" and "all hospitalizations with repeated measures" should be same. This independence assumption may be reasonable for analysis of some types of disease that do not usually result in repeated hospital visits, (eg, skin and subcutaneous tissue infections-related hospitalizations). For these diseases that are of mild to modest severity and treated relatively easily, significant differences in patient characteristics (eg, race/ethnicity, payer type) between single and multiple hospitalizations might not be seen. But if a patient is admitted with a disease that has severe comorbidities, the patient is more likely to be readmitted with same medical issue. In such cases, the repeated visits will need to be considered in the analysis. When the interval between visits is short, 30-day readmission is a popular outcome measure for hospitals to track. For readmissions over longer periods, many factors that are outside of hospitals' control, (eg, other complicating illnesses, patients' own behavior, or care provided to patients after discharge), could play a role. The message from this DRPH analysis is

that rehospitalizations are often dependent on discharge status of previous visits for a patient, and independence assumption of the data may not be appropriate. If multiple visits were not appropriately accounted for, significance of health disparities would be severely affected, and the error could have impacts across racial/ethnic groups, comorbidity status, payer groups, and across the lifespan.

As an example of the potential error in significance, the hospital charge analysis showed that RR estimates could be dramatically different between the analyses with patient identification compared to those not taking patient identification into consideration. For Chinese, the RR of hospitalization charges is not significantly different from Whites when adjusted for multiple visits, but Chinese had significantly lower charges when multiple visits were not properly taken into consideration. Without the appropriate consideration of multiple visits, the results could mislead health policy makers, which could lead to misallocated effort in reducing health disparities that may not actually exist.

The current study has several limitations. Administrative data have known limitations and do not include many characteristics, including education, household income, English proficiency, and other factors that may help explain observed differences. The data studied is from a single state and is more than 10 years old. The limited years of data were accessed before the HHIC organization was dismantled and it was extremely difficult to update the data with recent years after reorganization. The results may not reflect national trends. While a strength of this study is the diversity of Asian American and Pacific Islander groups included, this may make the findings less directly applicable to national findings pertaining to other important racial/ethnic disparities, especially among African Americans and Latinos. While all individuals who live in Hawai'i during the 6 year period are included, it is not known if individuals with hospitalizations in the state just recently entered the state or have since left the state. Also, such migration may vary by demographic groups. But the overall impact should be relatively small for Hawai'i as the emigration and immigration of the state is not high. In the period 1995–2000, 125 160 people moved into the state and 201 293 moved out, for a net loss of 76 133.<sup>18</sup> The assignment of first and last visit by discharge date was also limited by the study period, and it may not capture the true first and last visit of a given patient. The potential error could be due to not only repeated measurements but also other reasons such as reported or unreported death and the fact that the GEE models also did not consider those events.

## Conclusions

Over 20% of the patients over the 6-year period had more than 1 DRPH captured by the HHIC hospitalization data and over 8% of these patients had at least 3 hospitalizations. For a patient with multiple hospitalizations, rehospitalizations are often dependent upon the discharge status of previous visits

and the independence assumption of the multiple visits may not be appropriate. Ignoring multiple visits in population-level analyses could result in severely distorting the significance of health disparities.

## Conflict of Interest

The author does not identify a conflict of interest.

## Funding

The author is partially supported by Ola Hawaii (2U54MD007601-36) and PIKO (5U54GM138062-02) from the National Institute of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

### Author's Affiliation:

-Department of Quantitative Health Sciences, John A. Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI

### Corresponding Author:

Hyeong Jun Ahn PhD; Email: hjahn@hawaii.edu

## References

1. Centers for Disease Control and Prevention. *National Diabetes Statistics Report*. Updated June 29, 2022. Accessed December 21, 2022. <https://www.cdc.gov/diabetes/data/statistics-report/index.html>
2. Frazee T, Jiang J, Burgess J. Hospital stays for patients with diabetes, 2008. HCUP Statistical Brief #93. August 2010. Agency for Healthcare Research and Quality, Rockville, MD. Accessed December 10, 2022. <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb93.pdf>
3. Aubert RE, Geiss LS, Ballard DJ, Cocanougher B, Herman WH. Diabetes-related hospitalization and hospital utilization. In *Diabetes in America*. 2nd ed. Bethesda, MD, National Institutes of Health, 1995, p. 553–570 (NIH publ. no. 95-1468).
4. Kim S. Burden of Hospitalizations Primarily Due to Uncontrolled Diabetes. *Diabetes Care*. 2007;30(5):1281-1282.
5. Prevention Quality Indicators overview. Rockville (MD): Agency for Healthcare Research and Quality; 2011. Accessed October 11, 2012. [http://www.qualityindicators.ahrq.gov/modules/pqi\\_overview.aspx](http://www.qualityindicators.ahrq.gov/modules/pqi_overview.aspx).
6. McCarthy D, How SKH, Schoen C, Cantor JC, Belloff D. Aiming higher: results from a state scorecard on health system performance, 2009. The Commonwealth Fund Commission on a High Performance Health System. Accessed March 31, 2011. <http://www.commonwealthfund.org/Content/Publications/Fund-Reports/2007/Jun/Aiming-Higher-Results-from-a-State-Scorecard-on-Health-System-Performance.aspx>.
7. Centers for Disease Control and Prevention. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States*, 2014. Atlanta, GA: U.S. Department of Health and Human Services. 2014. Accessed 31, 2011. [https://stacks.cdc.gov/view/cdc/23442/cdc\\_23442\\_DS1.pdf](https://stacks.cdc.gov/view/cdc/23442/cdc_23442_DS1.pdf)
8. Grandinetti A, Kaholokula JK, Theriault AG, Mor JM, Chang HK, Waslien C. Prevalence of glucose intolerance among native Hawaiians in two rural communities. *Diabetes Care*. 1998;21:549-554.
9. McNeely MJ, Boyko EJ. Type 2 Diabetes Prevalence in Asian Americans. *Diabetes Care*. 2004;27:66-69.
10. Grandinetti A, Keawe'aimoku J, Theriault A, Mor J, Chang HK. Prevalence of diabetes and glucose intolerance in an ethnically diverse rural community of Hawaii. *Ethn Dis*. 2007;17:250-255.
11. Maskarinec G, Grandinetti A, Matsuura G, et al. Diabetes prevalence and body mass index differ by ethnicity: the Multiethnic Cohort. *Ethn Dis*. 2009;19(1):49-55.
12. State of Hawaii, Behavioral Risk Factor Surveillance System, from 2005 to 2007; by demographic characteristic and ethnic group. Honolulu (HI): Hawaii State Department of Health. Accessed December 6, 2012. <http://hawaii.gov/health/statistics/brfss/ethnicity/0567/ethnicity567.html>
13. The Office of Minority Health. *Diabetes and Hispanic Americans*, Updated December 2011. Accessed December 7, 2012. <http://minorityhealth.hhs.gov/omh/browse.aspx?lvl=4&lvlid=63>
14. Healthcare Cost & Utilization Project. State Inpatient Databases (SID) Database. Accessed November 28, 2022. <http://www.hcup-us.ahrq.gov/sidoverview.jsp>
15. Healthcare Cost & Utilization Project. National (Nationwide) Inpatient Sample (NIS). Accessed November 28, 2022. <http://www.hcup-us.ahrq.gov/nisoverview.jsp#Data>
16. Jiang HJ, Stryer D, Friedman B, Andrews R. Multiple Hospitalizations for Patients With Diabetes. *Diabetes Care*. 2003 May 1;26(5):1421-1426. doi: 10.2337/diacare.26.5.1421.
17. Basu A, Manning WG, Mullahy J. Comparing alternative models: log versus Cox proportional hazard? *Health Econ* 2004;13:749–765.
18. City-Data.com. Migration. Accessed December 7, 2012. <http://www.city-data.com/states/Hawaii-Migration.html>