

Investigating the Causal Effect of Socioeconomic Status on Quality of Care under a Universal Health Insurance System – A Marginal Structural Model Approach

Hoyune E. Cho

University of Michigan Medical School

Lu Wang

University of Michigan

Jung-Sheng Chen

Chang Gung Memorial Hospital

Mochuan Liu

University of North Carolina at Chapel Hill

Chang-Fu Kuo

Chang Gung Memorial Hospital

Kevin Chung (✉ kecchung@med.umich.edu)

University of Michigan Medical School <https://orcid.org/0000-0002-2822-3164>

Research article

Keywords: social determinants of health, preventable hospitalization, quality of care in the community, universal health insurance system, marginal structural model, causal effect relationship

Posted Date: November 13th, 2019

DOI: <https://doi.org/10.21203/rs.2.14965/v2>

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

[Read Full License](#)

Version of Record: A version of this preprint was published on December 23rd, 2019. See the published version at <https://doi.org/10.1186/s12913-019-4793-7>.

Abstract

Background: Social disparities in healthcare persist in the US despite the expansion of Medicaid under the Affordable Care Act. We investigated the causal impact of socioeconomic status on the quality of care in a setting with minimal confounding bias from race, insurance type, and access to care. **Methods:** We designed a retrospective population-based study with a random 25% sample of adult Taiwan population enrolled in Taiwan's National Health Insurance system from 2000 to 2016. Patient's income levels were categorized into low-income group (<25th percentile) and high-income group (\geq 25th percentile). We used marginal structural modeling analysis to calculate the odds of hospital admissions for 11 ambulatory care sensitive conditions identified by the Agency for Healthcare Research and Quality and the odds of having an Elixhauser comorbidity index greater than zero for low-income patients. **Results:** Among 2,844,334 patients, those in lower-income group had 1.28 greater odds (95% CI 1.24 – 1.33) of experiencing preventable hospitalizations, and 1.04 greater odds (95% CI 1.03 – 1.05) of having a comorbid condition in comparison to high-income group. **Conclusions:** Income was shown to be a causal factor in a patient's health and a determinant of the quality of care received even with equitable access to care under a universal health insurance system. Policies focusing on addressing income as an important upstream causal determinant of health to provide support to patients in lower socioeconomic status will be effective in improving health outcomes for this vulnerable social stratum.

Background

The two most important agendas to improve United States (US) healthcare are to enhance access to care and increase the quality of care delivered in an equitable manner.¹⁻⁴ Although the United States healthcare expenditures far exceeds other developed countries, the US ranks 30th in terms of morbidity and mortality.⁵ Despite efforts to expand delivery of health care through the Affordable Care Act,⁶ 12% of the population remained uninsured in 2016 with difficult access to primary care.⁷ Furthermore, middle and lower socioeconomic classes are less likely to have a regular source of care, less likely to receive preventive services, and more likely to experience delays in their care.⁸⁻¹¹ Studies have also found evidence for widening racial gap in health.¹² Social inequalities in access to healthcare persist in the US healthcare delivery system even with the expansion of Medicaid.^{8,13,14}

Medical care alone cannot adequately improve health or address health disparities by socioeconomic status (SES) in the US,¹⁵ and thus it is imperative to clearly delineate the causal relationship between a patient's SES and the quality of care received, in light of the recent government proposal for cutbacks on Medicaid.^{16,17} Many studies have investigated the influences of various social determinants of health, but few have recognized the longitudinal, compounding effect of SES on health.^{15,18,19} Furthermore, the confounding bias between variables of SES, race, and health outcomes were not investigated in the analysis.^{8,9,12,20-26} In fact, SES and race are deeply intertwined in their development and longitudinal continuum through multiple familial generations in the US.^{12,15} Even the direction of causality between SES, race, and health has been in debate.^{12,19} In addition, SES is correlated with the type of insurance

coverage, which further leads to disparities in access to care and health outcomes.¹⁰ Experts state that the confounding effect among race, SES, and health in the US cannot be sufficiently eliminated by statistical means.¹²

In contrast to our racially diverse nation with a multi-payer insurance system, Taiwan's population sample is homogeneous for race and insurance type, owing to the establishment of a universal, single-payer health insurance system in 1995.²⁷ Covering over 99% of the residents, the National Health Insurance (NHI) provides easily accessible, equitable care.^{27,28} Utilizing the NHI Database, we designed a population-based study with marginal structural modeling to more accurately investigate the longitudinal causal influence of SES on the quality of care in the community. Regarding income as a measure of SES, we hypothesized that without the confounding bias from race and insurance type, income is not a causal factor in determining the quality of care received. Results from this study will direct health policy towards improving quality of care by providing appropriate support for patients in lower SES.

Materials And Methods

Data Source

As one of the largest and most comprehensive national-level population databases in the world,²⁹ the NHI Database contains healthcare records of 30 million residents of Taiwan, including inpatient, outpatient, and pharmacy services.³⁰ The NHI Database is ideal for longitudinal epidemiological investigation because each beneficiary of the NHI has a unique identification number consistent across all datasets, and can be followed through multiple clinical encounters.³¹ To maximally retain the population heterogeneity to reflect the real-world impact of SES, we selected a random 25% sample from the eligible adult population in 2000 from the NHI Database by random sampling method based on Floyd's ordered hash table algorithm to ensure an equal probability of the eligible population being selected. We also excluded patients with missing data before 2006 for appropriate censoring and to ensure two-waves worth of longitudinal data at minimum (Figure 1). This study was approved by the Institutional Review Board at the Chang Gung Memorial Hospital.

Overview of Marginal Structural Modeling

Since patients' socioeconomic status is a time-varying variable, we adopted marginal structural model (MSM) analysis to investigate the causal relationship between patients' socioeconomic status as defined by income levels and the quality of care received measured by the rate of preventable hospitalization and the Elixhauser Comorbidity Index (ECI). MSM eliminates the confounding bias in estimating this causal relationship, where the other confounding effects vary over time from wave to wave, and these confounders may act as intermediate variables resulting from previous socioeconomic statuses at the same time.^{37,38} The observational data lack counterfactual outcomes that would have occurred under the

opposite exposure level or treatment decision.⁴⁰ Thus, cross-sectional observational studies use stratification or multivariable regression analyses to address potential confounding and make causal inferences. However, with longitudinal data for which the time-varying confounders (variables that are associated with both treatment and the subsequent outcome) and treatment by indication (variables affected by prior exposure and affect future exposure levels) co-exist, these traditional analyses cannot account for the dynamic effects of the time-varying covariates appropriately to make a causal inference.³⁹ MSM provides an approach to balance out the potential confounding effects in a longitudinal study by creating a pseudo-population to mimic the data from a sequentially-randomized trial, and use that to estimate the average effect of a treatment or an exposure on potential outcomes.⁴¹ We chose to use MSM to answer our research question because of these advantages that permit unbiased assessment of the causal effects of SES on the quality of care received in a longitudinal study. We conducted two separate models, each with a different outcome variable: one with the rate of preventable hospitalization and the other with the ECI.

Variables of Interest

We considered preventable hospitalization and comorbidity as representative outcome measures of the quality of care in the community.⁹ A patient was considered to have had a preventable hospitalization if the primary diagnosis code associated with the inpatient stay was included in the set of diagnosis and procedure codes defined by the Agency for Healthcare Research and Quality for 11 of ambulatory care sensitive conditions³² for which hospitalizations can be avoided with good outpatient care.^{20,21,33,34} We calculated each patient's Elixhauser comorbidity index (ECI)³⁵ by defining it as having ≥ 2 ambulatory visits or one hospital admission with a corresponding diagnosis code to ensure the validity of indices greater than zero. The first model with preventable hospitalization as outcome included ECI as a covariate, and was calculated as a categorical variable with 3 levels (0, 1-3, and ≥ 4). In the second model with ECI as outcome, ECI was defined as a continuous variable and calculated at years 2004, 2007, 2010, 2013, and 2016; thus ECI was excluded as a covariate from this model. Demographic and clinical covariates included in the analysis were patients' sex, age, income level, occupation type, urbanization of area of residence, number of outpatient visits, number of hospital admissions, and the physician density of the area of patients' residence. Occupation categories defined by the NHI program enrollment protocol³⁰ and income levels in Taiwan dollars were collected directly from the NHI Database and were converted to US dollars by the mean exchange rate during each corresponding calendar year.³⁶ We compared demographic characteristics by income groups for each time wave by analysis of variance test and chi-squared test. In our MSM analysis, values at year 2000 for each independent variable was defined as baseline values. Then, the baseline values were used to calculate the time-varying variables at each wave by taking the average value during each time interval for continuous variables and the mode for categorical variables at each wave and two years before each wave. For example, for the number of outpatient visits variable, the average value between 2001 and 2003 was considered as the number of

outpatient visits in 2003. To create income groups at each wave, the average income during the time-period up to each wave-year was used. For example, the average income from 2001 to 2008 was used to define the income level at 2009.

Inverse Probability of Treatment Weights

MSM can control the confounding effect of time-dependent confounders without over-adjusting by applying inverse probability of treatment weights (IPTW).⁴² At each time-point of follow-up, the probability of each patient receiving the treatment/exposure (or not receiving the treatment/exposure, whichever that actually took place) is estimated based on the baseline and time-varying covariates up to that time-point.⁴² Then, patients are weighted by the inverse of their predicted probabilities of receiving the observed treatment/exposure to create a pseudo-population without the covariate imbalances. Under-represented subjects, given their previous covariate values and treatment history, receive proportionally higher weights, and vice-versa for over-represented subjects. In this pseudo-population, the potential confounders are distributed evenly and thus we can estimate causal effects.³⁹ To reduce the variability and improve the precision of estimation, we applied the stabilized version of IPTW weights^{42,43} as follows: (see Formula 1 in the Supplementary Files)

$P\{A^*\}$ denotes the probability function, $A(k)$ represents the time-varying exposure at time k , $(k-1)$ represents the exposure history prior to time $(k-1)$, (k) are the time-dependent covariates through time k that are possible mediators as well as confounders, and $L(0)$ represents the vector of baseline covariates. Here, the numerator contains all covariates measured at baseline, and the denominator contains both baseline and time-varying confounders.⁴² In our model, weights larger than 50 were considered extreme, and the weights were truncated at 50 in our analysis. To ensure that the confounders were balanced after applying IPTW, we compared absolute standardized mean differences across different exposure groups calculated before and after the weight application.

Censoring/Attrition

To minimize selection bias from inconsistent study cohort at multiple time points, we used censoring weights to account for any loss to follow-up in the data by calculating for the probability of remaining uncensored up to each point of follow-up.^{44,45} We fit censoring models to predict the probability that a patient remained in the study for each time-interval that the patient actually remained in the study.³⁸ Each subject was weighted by the IPTW multiplied by the inverse probability of censoring weight.⁴²

MSM Analysis

We defined 5 time-waves with 3-year interval between 2000 (baseline) and 2016. The baseline values and time-varying covariates collected up to each wave-year were used to assess their relationships with the outcome variable at each wave-year (Figure 2).

Before applying MSM, we first explored the associations of income quartiles with the outcome and found that outcomes from the higher three income quartiles were very similar. Therefore, we decided to dichotomize the income level into low-income group (<25th percentile) and high-income group (\geq 25th percentile).

We used a two-stage approach to estimate treatment effects from MSM (Figure 2). During the first stage, we estimated each patient's probability of being in his or her income group at each time point, and used the inverse of this probability as weights to balance the potential confounding owing to the observed and non-randomized income levels. To acquire the treatment weights (here low-income exposure), we fitted logistic regression models with both baseline variables and time-dependent variables up to each wave (2003, 2006, 2009, 2012, and 2015). Censoring weights were also applied by estimating the survival probability, because the most common reason for discontinuing the NHI coverage was death. Censoring was present at time = t if the patient transferred out prior to or at the next time point t + 1. The final weight for each patient was calculated by multiplying both the treatment weights and censoring weight at each time point. The numerator for treatment weighting was derived by adjusting the baseline values and previous income binary groups in the model and the denominator was derived by adjusting the baseline values and time-dependent variables. The numerator and denominator for censoring weights were obtained from the cox proportional hazard models, with the response variable in cox model as the patients' binary censoring status.

Then, in the second stage, the causal parameter in the pseudo-population created with each individualized IPTW was recovered by fitting a weighted generalized linear model on the health outcomes (here preventable hospitalization and ECI). We applied IPTW to avoid biased estimation that happens when time-dependent confounders are inappropriately adjusted by stratification or traditional regression approaches. Furthermore, this methodology separates the time-dependent covariates confounder adjustment from the mediation adjustment in assessing the causal impact on the outcome.^{41,45,46}

Results

Sample Characteristics - Baseline

In the baseline year 2000, our study cohort included 2,844,334 patients with 32.3% in the low income group. By comparing demographic characteristics, our 25% random sample was not significantly different from the overall population (Supplement 1). Among patients in the low-income group, 59.1% lived in urban, and 34.5% in suburban areas. The distribution by place of residence for urban and suburban areas were similar between low-income and high-income groups, but the proportion of high-income group living in rural areas (10.2%) was greater than that of the low-income group (6.4%). In fact,

only 16.7% (44,262/220,148) of rural dwellers were categorized as low-income. Distribution by occupation type was similar between low and high-income groups for public employees (category 1, 21.4% vs. 19.0%) and private employees (category 2, 37.6% vs. 39.3%). However, for self-employed (category 3) and those related to the military (category 4), over 99% were included in the high-income group. Approximately 93.7% of veterans and those without permanent jobs (category 6) were categorized as low-income. The composition of patient mix by ECI was similar between the two income groups. The mean number of outpatient visits (calculated per 1000 patients) was 11.3 for low-income group and 11.5 for high-income group. The mean number of hospital admissions, also calculated for every 1000 patients, was 0.11 for low-income group and 0.10 for high-income group. In other words, patients with higher income utilized ambulatory care services more frequently whereas lower-income patients required inpatient services at a higher rate (Table 1).

Sample Characteristics – First wave to Fifth wave

The total number of patients in the study decreased from 2,844,334 in the first and second waves to 2,753,224 in the third, 2,644,668 in the fourth, and 2,538,246 in the fifth wave. Thus, the overall rate of attrition during the 16-year study period was 10.8%. The high-end limit in income ranges increased for both low and high-income groups from first wave to fourth wave, reflecting the growth of Taiwan economy over the years. The distribution of residence by urbanization stayed relatively constant throughout the study period.

The proportion of patients with ECI of 0 decreased over the years for both low and high-income groups, whereas those with index between 1 and 3 increased for both groups. There was an increasing trend in rate of health care utilization by the low-income group, as the number of outpatient visits increased from 2003 to 2015 whereas it stayed relatively constant for the high-income group. The number of inpatient admissions stayed the same for both income groups. The physician density in area of residence increased over the years for both income groups as well (Tables 1, 2, and 3).

Analysis of Causal Inference by MSM

The covariates between the low-income and high-income groups were not balanced across the waves (Tables 1, 2, and 3). We compared the absolute standardized mean differences before and after applying the IPTW and confirmed that the weighted data used in MSM analysis had the covariates balanced (Supplement 2). Through MSM analysis, we found that patients in low-income group had 1.28 times greater odds of incurring a preventable hospitalization in comparison to patients in the high-income group ($p < 0.001$). In other words, patients with lower income were 28% more likely to require inpatient-level treatment of their ambulatory care sensitive condition(s). We also found that patients in low-income group had 1.04 times greater odds of having a comorbidity ($p < 0.001$) (Table 4).

Discussion

In our study, patients in lower income group were 28% more likely to experience a preventable hospitalization, which indicates that they have a much higher risk of acquiring diseases severe enough to require inpatient-level treatment that would have been prevented with good ambulatory care. Our finding of lower income patients having 4% higher odds for comorbidity echoes this interpretation. Our analysis demonstrated that income is a causal factor in a patient's health status even under a universal health insurance system that grants equitable access to care. Furthermore, as the Taiwanese population is relatively homogenous in race, this finding is in a setting where racial disparities in healthcare that exist in the US is eliminated. In addition, our use of MSM shows that the relationship between a patient's SES and quality of care received is actually causal, rather than correlational.

Past studies that have investigated the influence of various demographic and health factors on preventable hospitalization are based on observational data^{8,10,11,20-22,25} and there is a concern with effect estimates that they may be biased from unobserved confounding among the variables used.⁴⁶ Moreover, these studies used cross-sectional designs that assume the magnitude of influence of each study variable is constant and takes immediate effect.⁴⁶ Rather, many of these patient and environment factors influence health at each life stage, with accumulating social advantages and disadvantages that translate to health advantages and disadvantages over time through complex causal pathways.^{15,21,46} In fact, a systematic review comparing estimates from conventional analysis and MSM found that there is a statistically significant difference in results between the two analytic designs.³⁹ Thus, our result by MSM analysis can be considered to be a more accurate estimate of the true effect of a patient's socioeconomic status on quality of care received.

Our study has established the causal impact of income on health, and that it is a social factor at the very root of health inequalities stemming from issues both inside and outside of the healthcare division.^{9,23,47} Upstream social determinants of health have been correlated to the onset and progression of various diseases as well as to overall mortality,⁴⁸ and experts stress that many of these factors influence each other as well.²¹ It is well accepted that strategies to solve health disparities should address the very generators of the inequalities,²³ yet in our current healthcare system, patients' health-related social needs are seldom evaluated or addressed.^{48,49} Information on these factors along with patients' clinical data would permit providers to tailor services and improve effectiveness of the care delivered.⁴⁸ Global health community considers universal health insurance coverage as a tool to improve access to care and further the population health, but our results suggest that providing universal health insurance is not enough to overcome the persistent inequalities in health by SES.⁵⁰ Policies that recognize income as an upstream causal determinant of health would be more effective in addressing the health disparities by SES and improve the quality of care for patients in lower socioeconomic group.^{51,52}

Based on the nature of claims data, our study is subject to error from the inherent uncertainty in coding practice. Also, the validity and consistency of IPTW is based on the assumption that there are no other

unmeasured confounders in MSM analysis,³⁷ which is not guaranteed in real-life situations. Furthermore, MSM holds under the condition that the measured covariates at each time point are sufficient to adjust for confounding, which cannot be tested with observed data.⁴⁴ However, our study cohort is comprised from a national database with demonstrated validity of recorded diagnoses⁵³⁻⁵⁵ that represents more than 99% of the population, and thus coding errors and the selection bias stemming from the study design are minimized. We excluded patients with missing data before 2006 for appropriate censoring and to ensure two-waves worth of data, and this may add to the selection bias. However, the excluded patients are small in proportion (10% of the registrants of the NHI in 2000, Figure 1) and thus should not lay a significant impact on our results. Lastly, our study timeframe of 16 years captures a small portion of a patient's lifetime.⁴⁶ Nevertheless, our results still successfully showcase the causal effect of income on a patient's health and quality of care received with methodology to address time-varying confounding bias within longitudinal population-based data.

Conclusion

Socioeconomic disparities in the quality of care delivered in the US persist despite the expansion of insurance coverage. Our study using data from a universal health insurance system found that income is a casual determinant of health even with equitable access to care. Patients in lower socioeconomic group would benefit greatly from policy interventions that recognize and address the upstream social determinants of health and the inequalities that emerge from them.

List Of Abbreviations

1. US = United States
2. SES = socioeconomic status
3. NHI = National Health Insurance
4. MSM = marginal structural model
5. ECI = Elixhauser comorbidity index
6. IPTW = inverse probability of treatment weights

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data and materials used for this study are available from one of the corresponding authors (CFK) upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

KCC: A Midcareer Investigator Award (2 K24-AR053120-06) from the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health

The funding body did not have any influence on the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

HEC: A Surgical Scientist Training Grant in Health Services and Translational Research (2 T32-GM008616-16A1) from the National Institutes of Health Ruth L. Kirschstein National Research Service Award

The funding body did not have any influence on the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

JSC and CFK: Maintenance Project for Center for Big Data Analytics and Statistics of Chang Gung Memorial Hospital (CLRPG3D0043, CORPG3G0111, CORPG3G0161)

The funding body did not have any influence on the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

HEC, LW, CFK, and KCC conceptualized the study. HEC, LW, JSC, ML, and CFK helped with the study's methodology. JSC and ML carried out the analysis. LW, ML, and CFK validated the results. JSC created tables and figures. HEC wrote the original draft of the manuscript. LW, JSC, and CFK reviewed and edited the manuscript. CFK and KCC supervised the entire study.

All authors have read and approved the manuscript.

Acknowledgements

Not applicable

References

1. Cutler D, Wikler E, Basch P. Reducing administrative costs and improving the health care system. *N Engl J Med.* 2012;367(20):1875-1878.

2. Weinick RM, Burns RM, Mehrotra A. Many emergency department visits could be managed at urgent care centers and retail clinics. *Health Aff (Millwood)*. 2010;29(9):1630-1636.
3. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med*. 2007;22 Suppl 3:391-395.
4. Ayanian JZ, Williams RA. Principles for Eliminating Racial and Ethnic Disparities in Healthcare. In: Williams RA, ed. *Eliminating Healthcare Disparities in America: Beyond the IOM Report*. Totowa, NJ: Humana Press; 2007:377-389.
5. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Aff (Millwood)*. 2008;27(3):759-769.
6. Koh HK, Sebelius KG. Promoting prevention through the Affordable Care Act. *N Engl J Med*. 2010;363(14):1296-1299.
7. OECD. *Universal Health Coverage and Health Outcomes, Final Report*. Paris, France 22 July 2016 2016.
8. Pappas G, Hadden WC, Kozak LJ, Fisher GF. Potentially avoidable hospitalizations: inequalities in rates between US socioeconomic groups. *Am J Public Health*. 1997;87(5):811-816.
9. Billings J, Zeitel L, Lukomnik J, Carey TS, Blank AE, Newman L. Impact of socioeconomic status on hospital use in New York City. *Health Aff (Millwood)*. 1993;12(1):162-173.
10. Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. *Jama*. 1992;268(17):2388-2394.
11. Chen PC, Tsai CY, Woung LC, Lee YC. Socioeconomic disparities in preventable hospitalization among adults with diabetes in Taiwan: a multilevel modelling approach. *Int J Equity Health*. 2015;14:31.
12. Williams DR, Collins C. US Socioeconomic and Racial Differences in Health: Patterns and Explanations. *Annual Review of Sociology*. 1995;21:349-386.
13. Billings J, Anderson GM, Newman LS. Recent findings on preventable hospitalizations. *Health Aff (Millwood)*. 1996;15(3):239-249.
14. Schoen C, Osborn R, Squires D, Doty MM. Access, affordability, and insurance complexity are often worse in the United States compared to ten other countries. *Health Aff (Millwood)*. 2013;32(12):2205-2215.
15. Braveman P, Egerter S, Williams DR. The social determinants of health: coming of age. *Annu Rev Public Health*. 2011;32:381-398.
16. Werner E. House GOP plan would cut Medicare, Medicaid to balance budget. *The Washington Post*. June 19, 2018, 2018.
17. Friedman MFF, Welch-Marahar M, Udow-Phillips M. Comparing Key Provisions: Affordable Care Act, American Health Care Act, and Better Care Reconciliation Act. Center for Healthcare Research & Transformation. <https://www.chrt.org/publication/comparing-key-provisions-affordable-care-act-american-health-care-act-better-care-reconciliation-act/>. Accessed July 11th, 2018.

18. Lynch J, Smith GD, Harper S, Hillemeier M. Is income inequality a determinant of population health? Part 2. U.S. National and regional trends in income inequality and age- and cause-specific mortality. *Milbank Q.* 2004;82(2):355-400.
19. Pickett KE, Wilkinson RG. Income inequality and health: a causal review. *Soc Sci Med.* 2015;128:316-326.
20. Falster MO, Jorm LR, Douglas KA, Blyth FM, Elliott RF, Leyland AH. Sociodemographic and health characteristics, rather than primary care supply, are major drivers of geographic variation in preventable hospitalizations in Australia. *Med Care.* 2015;53(5):436-445.
21. Blustein J, Hanson K, Shea S. Preventable hospitalizations and socioeconomic status. *Health Aff (Millwood).* 1998;17(2):177-189.
22. Culler SD, Parchman ML, Przybylski M. Factors related to potentially preventable hospitalizations among the elderly. *Med Care.* 1998;36(6):804-817.
23. Kavanagh A, Bentley RJ, Turrell G, Shaw J, Dunstan D, Subramanian SV. Socioeconomic position, gender, health behaviours and biomarkers of cardiovascular disease and diabetes. *Soc Sci Med.* 2010;71(6):1150-1160.
24. Smith BT, Lynch JW, Fox CS, et al. Life-course socioeconomic position and type 2 diabetes mellitus: The Framingham Offspring Study. *Am J Epidemiol.* 2011;173(4):438-447.
25. Will JC, Nwaise IA, Schieb L, Zhong Y. Geographic and racial patterns of preventable hospitalizations for hypertension: Medicare beneficiaries, 2004-2009. *Public Health Rep.* 2014;129(1):8-18.
26. Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. *Jama.* 1995;274(4):305-311.
27. Cheng TM. Reflections on the 20th anniversary of Taiwan's single-payer National Health Insurance System. *Health Aff (Millwood).* 2015;34(3):502-510.
28. Wu TY, Majeed A, Kuo KN. An overview of the healthcare system in Taiwan. *London J Prim Care (Abingdon).* 2010;3(2):115-119.
29. Hsing AW, Ioannidis JP. Nationwide Population Science: Lessons From the Taiwan National Health Insurance Research Database. *JAMA Intern Med.* 2015;175(9):1527-1529.
30. Tsai JC, Chen WY, Liang YW. Nonemergent emergency department visits under the National Health Insurance in Taiwan. *Health Policy.* 2011;100(2-3):189-195.
31. Lin YJ, Tian WH, Chen CC. Urbanization and the utilization of outpatient services under National Health Insurance in Taiwan. *Health Policy.* 2011;103(2-3):236-243.
32. AHRQ Quality Indicators - Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions. Rockville, MD: Agency for Healthcare Research and Quality, 2001. AHRQ Pub. No. 02-R0203.
33. Nyweide DJ, Anthony DL, Bynum JP, et al. Continuity of care and the risk of preventable hospitalization in older adults. *JAMA Intern Med.* 2013;173(20):1879-1885.

34. Moy E, Chang E, Barrett M. Potentially preventable hospitalizations - United States, 2001-2009. *MMWR Suppl.* 2013;62(3):139-143.
35. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* 1998;36(1):8-27.
36. OFX Historical Exchange Rates, Yearly Average Rates. . <https://www.ofx.com/en-ca/forex-news/historical-exchange-rates/yearly-average-rates/>. Accessed September 27th, 2018.
37. Lee KM. *Marginal Structural Modeling in Health Services Research*: Department of Health Policy & Management, Boston University School of Public Health; 2013.
38. Moodie EE, Stephens DA. Marginal Structural Models: unbiased estimation for longitudinal studies. *Int J Public Health.* 2011;56(1):117-119.
39. Suarez D, Borrás R, Basagana X. Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review. *Epidemiology.* 2011;22(4):586-588.
40. Mortimer KM, Neugebauer R, van der Laan M, Tager IB. An application of model-fitting procedures for marginal structural models. *Am J Epidemiol.* 2005;162(4):382-388.
41. Joffe MM HT, Feldman HI, Kimmel SE. Model Selection, Confounder Control, and Marginal Structural Models: Review and New Applications. *The American Statistician.* 2004;58(4):272-279.
42. Godin O, Elbejjani M, Kaufman JS. Body mass index, blood pressure, and risk of depression in the elderly: a marginal structural model. *Am J Epidemiol.* 2012;176(3):204-213.
43. Xu S, Ross C, Raebel MA, Shetterly S, Blanchette C, Smith D. Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value Health.* 2010;13(2):273-277.
44. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology.* 2000;11(5):561-570.
45. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550-560.
46. Do DP, Wang L, Elliott MR. Investigating the relationship between neighborhood poverty and mortality risk: a marginal structural modeling approach. *Soc Sci Med.* 2013;91:58-66.
47. Marmot M. Social determinants of health inequalities. *Lancet.* 2005;365(9464):1099-1104.
48. Adler NE, Stead WW. Patients in context—EHR capture of social and behavioral determinants of health. *N Engl J Med.* 2015;372(8):698-701.
49. Alley DE, Asomugha CN, Conway PH, Sanghavi DM. Accountable Health Communities—Addressing Social Needs through Medicare and Medicaid. *N Engl J Med.* 2016;374(1):8-11.
50. Sommers BD, Baicker K, Epstein AM. Mortality and Access to Care among Adults after State Medicaid Expansions. *New England Journal of Medicine.* 2012;367(11):1025-1034.
51. Lantz PM, Lichtenstein RL, Pollack HA. Health policy approaches to population health: the limits of medicalization. *Health Aff (Millwood).* 2007;26(5):1253-1257.

52. Marmot M. Universal health coverage and social determinants of health. *Lancet*. 2013;382(9900):1227-1228.
53. Cheng CL, Kao YH, Lin SJ, Lee CH, Lai ML. Validation of the National Health Insurance Research Database with ischemic stroke cases in Taiwan. *Pharmacoepidemiol Drug Saf*. 2011;20(3):236-242.
54. Cheng CL, Lee CH, Chen PS, Li YH, Lin SJ, Yang YH. Validation of acute myocardial infarction cases in the national health insurance research database in taiwan. *J Epidemiol*. 2014;24(6):500-507.
55. Kao WH, Hong JH, See LC, et al. Validity of cancer diagnosis in the National Health Insurance database compared with the linked National Cancer Registry in Taiwan. *Pharmacoepidemiol Drug Saf*. 2018;27(10):1060-1066.

Tables

Table 1. Sample Characteristics, Baseline and First Waves (No. (%), mean (95% CI))

	Baseline (2000)			First wave (2003)		
	Low Income	High Income	P value	Low Income	High Income	P value
	(29 - 586 USD)	(615 - 1,851 USD)		(26 - 543 USD)	(543 - 2,594 USD)	
N	694,162	2,150,172		711,684	2,132,650	
Place of residence						
Urban	410,251 (59.1%)	1,257,923 (58.5%)	<.001	429,044 (60.3%)	1,274,360 (59.8%)	<.001
Suburban	239,649 (34.5%)	672,101 (31.3%)		234,431 (32.9%)	632,389 (29.7%)	
Rural	44,262 (6.4%)	220,148 (10.2%)		48,209 (6.8%)	225,901 (10.6%)	
Occupation Category ^a						
1	148,891 (21.4%)	408,848 (19.0%)	<.001	150,682 (21.2%)	385,223 (18.1%)	<.001
2	260,675 (37.6%)	844,689 (39.3%)		212,327 (29.8%)	853,357 (40.0%)	
3	101 (0.0%)	456,637 (21.2%)		10,839 (1.5%)	446,917 (21.0%)	
4	0 (0.0%)	420,587 (19.6%)		3,201 (0.4%)	435,063 (20.4%)	
5	10,542 (1.5%)	884 (0.0%)		14,072 (2.0%)	643 (0.0%)	
6	273,953 (39.5%)	18,527 (0.9%)		320,563 (45.0%)	11,447 (0.5%)	
Elixhauser Comorbidity Index						
0	559,719 (80.6%)	1,755,767 (81.7%)	<.001	547,246 (76.9%)	1,675,097 (78.5%)	0.015
1-3	129,514 (18.7%)	383,916 (17.9%)		160,105 (22.5%)	450,250 (21.1%)	
≥4	4,929 (0.7%)	10,489 (0.5%)		4,333 (0.6%)	7,303 (0.3%)	
Outpatient Visits ^b	11.30 (11.27-11.33)	11.52 (11.50-11.53)	<.001	11.41 (11.38-11.43)	11.36 (11.35-11.38)	<.001
Inpatient Stays ^b	0.113 (0.112-0.114)	0.095 (0.094-0.095)	<.001	0.120 (0.119-0.121)	0.096 (0.096-0.097)	<.001
Physician Density of Residence ^b	1.573 (1.569-1.577)	1.712 (1.709-1.714)	<.001	1.655 (1.651-1.659)	1.809 (1.806-1.812)	<.001

^a Category 1 = civil servants, full-time or regularly paid personnel in governmental agencies and public schools, 2 = employees of privately owned enterprises or institutions, 3 = self-employed, other employees or paid personnel, and members of farmer and fishermen associations, 4 = military personnel, military school students, bereaved families of deceased military personnel, public service in lieu of military service, 5 = low-income

citizens, 6 = veterans and dependents, and citizens without a fixed profession from other areas.

^b Number per 1000 patients.

Table 2. Sample Characteristics, Second and Third Waves (No. (%), mean (95% CI))

	Second Wave (2006)			Third Wave (2009)		
	Low Income	High Income	P value	Low Income	High Income	P value
	(28 - 580 USD)	(580 - 4,077 USD)		(28 - 592 USD)	(592 - 4,087 USD)	
N	711,590	2,132,744		690,034	2,063,190	
Place of Residence						
Urban	435,359 (61.2%)	1,280,870 (60.1%)	<.001	429,769 (62.3%)	1,256,949 (60.9%)	<.001
Suburban	227,213 (31.9%)	630,682 (29.6%)		213,573 (31.0%)	598,487 (29.0%)	
Rural	49,018 (6.9%)	221,192 (10.4%)		46,692 (6.8%)	207,754 (10.1%)	
Occupation Category^a						
1	146,102 (20.5%)	357,037 (16.7%)	<.001	143,267 (20.8%)	351,295 (17.0%)	<.001
2	181,885 (25.6%)	887,469 (41.6%)		159,376 (23.1%)	846,449 (41.0%)	
3	19,462 (2.7%)	445,932 (20.9%)		20,309 (2.9%)	446,322 (21.6%)	
4	4,097 (0.6%)	433,736 (20.3%)		3,769 (0.5%)	412,798 (20.0%)	
5	17,015 (2.4%)	626 (0.0%)		18,369 (2.7%)	155 (0.0%)	
6	343,029 (48.2%)	7,944 (0.4%)		344,944 (50.0%)	6,171 (0.3%)	
Elixhauser Comorbidity Index						
0	501,825 (70.5%)	1,577,535 (74.0%)	<.001	484,908 (70.3%)	1,533,179 (74.3%)	<.001
1-3	204,453 (28.7%)	546,943 (25.6%)		201,450 (29.2%)	524,644 (25.4%)	
≥4	5,312 (0.7%)	8,266 (0.4%)		3,676 (0.5%)	5,367 (0.3%)	
Outpatient Visits ^b	12.24 (12.21-12.27)	11.63 (11.61-11.64)	<.001	12.06 (12.03-12.09)	11.53 (11.51-11.54)	<.001
Inpatient Stays ^b	0.137 (0.136-0.138)	0.104 (0.104-0.105)	<.001	0.130 (0.129-0.130)	0.101 (0.101-0.101)	<.001
Physician Density of Residence ^b	1.728 (1.725-1.732)	1.835 (1.832-1.837)	<.001	1.820 (1.817-1.824)	1.894 (1.891-1.896)	<.001

^a Category 1 = civil servants, full-time or regularly paid personnel in governmental agencies and public schools, 2 = employees of privately owned enterprises or institutions, 3 = self-employed, other employees or paid personnel, and members of farmer and fishermen associations, 4 = military personnel, military school students, bereaved families of deceased military personnel, public service in lieu of military service, 5 = low-income

citizens, 6 = veterans and dependents, and citizens without a fixed profession from other areas.

^b Number per 1000 patients.

Table 3. Sample Characteristics, Fourth and Fifth Waves (No. (%), mean (95% CI))

	Fourth Wave (2012)			Fifth Wave (2015)		
	Low Income	High Income	P value	Low Income	High Income	P value
	(32 - 650 USD)	(650 - 5,966 USD)		(37 - 694 USD)	(694 - 6,125 USD)	
N	661,473	1,983,195		633,181	1,905,065	
Place of Residence						
Urban	407,621 (61.6%)	1,192,881 (60.1%)	<.001	388,749 (61.4%)	1,141,552 (59.9%)	<.001
Suburban	209,417 (31.7%)	592,433 (29.9%)		201,148 (31.8%)	576,237 (30.2%)	
Rural	44,435 (6.7%)	197,881 (10.0%)		43,284 (6.8%)	187,276 (9.8%)	
Occupation Category ^a						
1	141,949 (21.5%)	344,832 (17.4%)	<.001	146,710 (23.2%)	394,423 (20.7%)	<.001
2	154,881 (23.4%)	803,378 (40.5%)		135,848 (21.5%)	763,839 (40.1%)	
3	15,196 (2.3%)	445,780 (22.5%)		10,742 (1.7%)	389,919 (20.5%)	
4	3,676 (0.6%)	384,836 (19.4%)		3,134 (0.5%)	354,372 (18.6%)	
5	23,255 (3.5%)	22 (0.0%)		25,266 (4.0%)	8 (0.0%)	
6	322,516 (48.8%)	4,347 (0.2%)		311,481 (49.2%)	2,504 (0.1%)	
Elixhauser Comorbidity Index						
0	436,629 (66.0%)	1,405,986 (70.9%)	<.001	406,967 (64.3%)	1,333,529 (70.0%)	<.001
1-3	221,470 (33.5%)	572,353 (28.9%)		223,665 (35.3%)	567,905 (29.8%)	
≥4	3,374 (0.5%)	4,856 (0.2%)		2,549 (0.4%)	3,631 (0.2%)	
Outpatient Visits ^b	12.30 (12.27-12.33)	11.59 (11.57-11.60)	<.001	12.43 (12.41-12.46)	11.61 (11.60-11.62)	<.001
Inpatient Stays ^b	0.126 (0.125-0.127)	0.099 (0.098-0.099)	<.001	0.121 (0.121-0.122)	0.097 (0.096-0.097)	<.001
Physician Density of Residence ^b	1.901 (1.897-1.904)	1.955 (1.952-1.957)	<.001	1.976 (1.972-1.980)	2.018 (2.015-2.020)	<.001

^a Category 1 = civil servants, full-time or regularly paid personnel in governmental agencies and public schools, 2 = employees of privately owned enterprises or institutions, 3 = self-employed, other employees or paid personnel, and members of farmer and fishermen associations, 4 = military personnel, military school students, bereaved families of deceased military personnel, public service in lieu of military service, 5 = low-income

citizens, 6 = veterans and dependents, and citizens without a fixed profession from other areas.

^b Number per 1000 patients.

Table 4. Analysis of Causal Impact of Low Income on Health with Marginal Structural Modeling

	Low Income (<25 th percentile) vs. High Income (≥25 th percentile)	
	Odds Ratio (95% CI)	P value
Preventable Hospitalization	1.28 (1.24 - 1.33)	<0.001
Elixhauser Comorbidity Index	1.04 (1.03 - 1.05)	<0.001

Figures

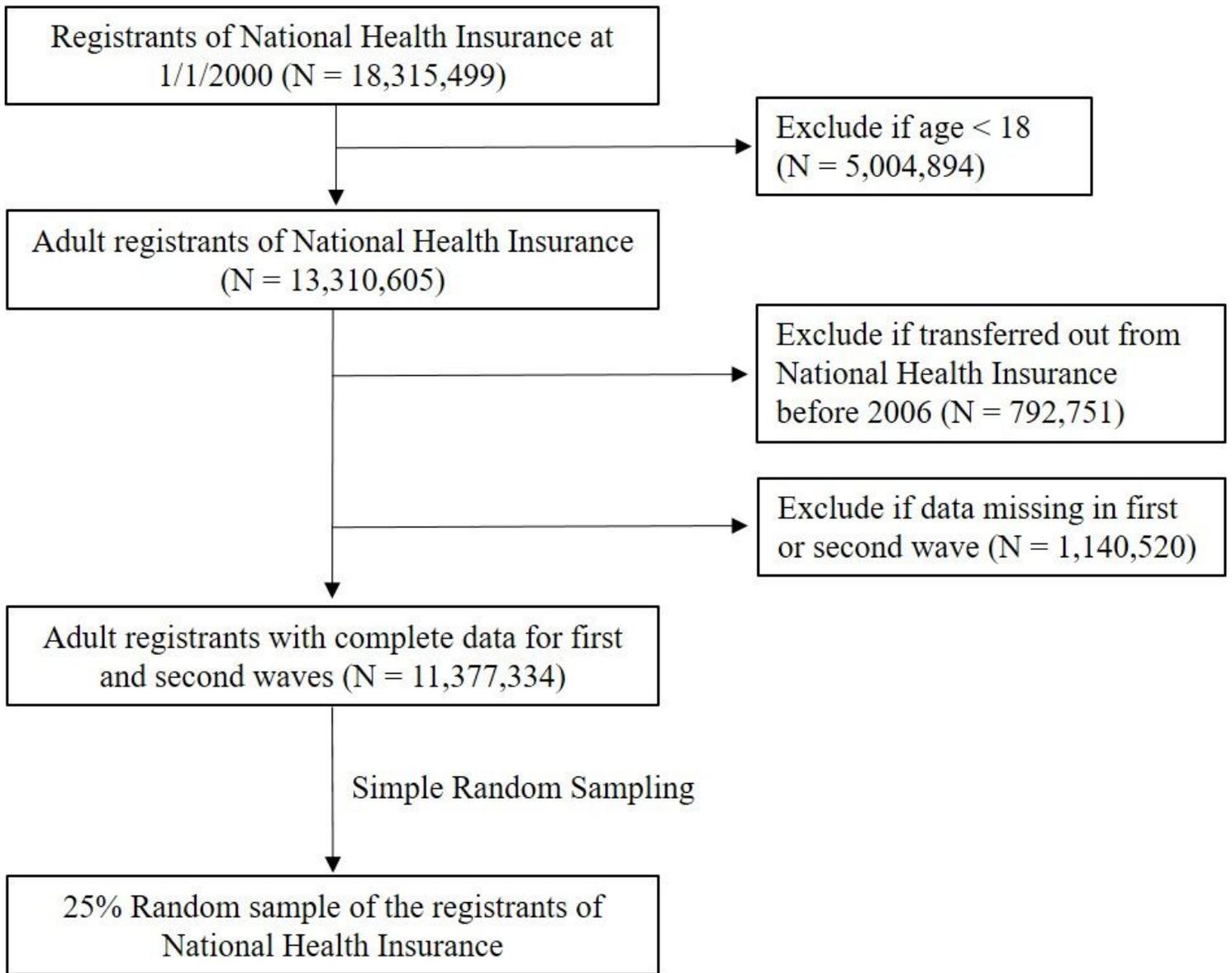


Figure 1

Selection flow for 25% random sample from Taiwan population

For patients enrolled before or in 2000 :

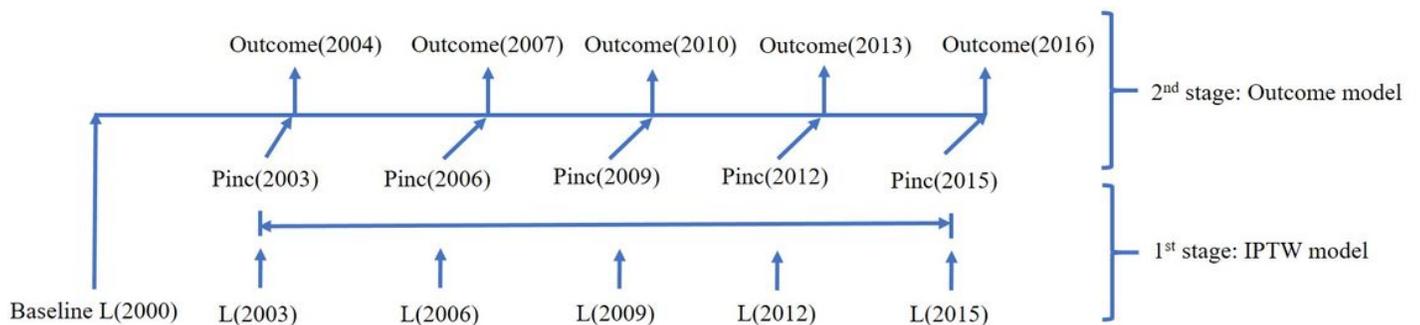


Figure 2

Marginal structural modeling conceptual framework *Baseline covariates ($L(0)$) were used to predict the low-income group exposure at each wave. The time-varying covariates ($L(t)$) was used to predict the low-income exposure at each wave, $Pinc(t)$, as the outcome variable in the first stage. Then, during the second stage, $Pinc(t)$ was used as the independent variable to examine the causal relationship between income and preventable hospitalization and comorbidity. All models in the first stage included baseline and prior low-income status to predict $Pinc(t)$.

Supplementary Files

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

- [Supplement209052019.xlsx](#)
- [Supplement108132019.docx](#)
- [Formula1.jpg](#)