IBM MarketScan Research Databases at UCSF (formerly Truven Health Analytics Inc.)

Data Description

Data Source	Commercial healthcare claims from employers, health plans and hospitals; Medicare supplemental claims (i.e., payments reimbursed by commercial plans after Medicare payments; no CMS claims included)
Data Contents	• De-identified records of more than 250 million patients • Beneficiary enrollment and healthcare service line level claims • Sample data elements include: o demographic data (age, gender, employment status) o geographic location: MSA or state level o inpatient and outpatient medical information (admission date and type, principal and secondary diagnosis codes, discharge status, major diagnostic category, principal and secondary procedure codes, DRGs, length of stay, place of service, quantity of services) o financial information (total and net payments, payments to physician, hospital, total admission payments) o drug information (generic product name, average wholesale price, prescription drug payment, therapeutic class, days supplied, national drug code, refill number, therapeutic group)
Ages Included	0-64 (Commercial data); 65+ (Medicare Supplemental)
Date Range	2010 – 2018, patient ID's are fully longitudinal
Care Settings	Outpatient, inpatient, emergency, and outpatient pharmacy
Cost for Use	Free for UCSF with intramural/internal funding; $30k/study with extramural non-profit funding, $60k/study with extramural industry funding

Limitations	• Submitted charges not included in cost information; no benefit design information available aside from high level insurance product type (PPO/HMO/etc.) • Demographics do not include race, income, language, or other social determinants; non-representative sample • Difficult to differentiate individual health care providers and attribute patients to unique clinicians • Population over age 65 represents only individuals who receive commercial insurance in addition to Medicare.
Operational Concerns	• Requires experienced programmer for data management and analysis. Population Health Data Initiative (PHDI) provides limited programmer support to perform initial extracts and brief consultations only. However, we are happy to work with researchers to help identify potential staff collaborators or craft future grant proposals that include funding for programming assistance.

How to Access

Please submit a CTSI Consultation Request Form.

Examples of Studies Using this Data

Franc, B. L., Copeland, T. P., Thombley, R., Park, M., Marafino, B., Dean, M. L., Boscardin, W., Rugo, H. S., Seidenwurm, D., Sharma, B., Johnston, S. R., & Dudley, R. (2018). Geographic Variation in Postoperative Imaging for Low-Risk Breast Cancer, Journal of the National Comprehensive Cancer Network J Natl Compr Canc Netw, 16(7), 829-837. Retrieved Sep 20, 2019, from https://jnccn.org/view/journals/jnccn/16/7/article-p829.xml

Gibson TB, Herring SA, Kutcher JS, Broglio SP. Analyzing the Effect of State Legislation on Health Care Utilization for Children With Concussion. JAMA Pediatrics. 2015;169(2):163-168. doi:https://doi.org/10.1001/jamapediatrics.2014.2320

Song Z, Ji Y, Safran DG, Chernew ME. Health Care Spending, Utilization, and Quality 8 Years into Global Payment. N Engl J Med. 2019;381(3):252-263. doi:https://doi.org/10.1056/NEJMsa1813621

Mokhlesi B, Ham SA, Gozal D. The effect of sex and age on the comorbidity burden of OSA: an observational analysis from a large nationwide US health claims database. European Respiratory Journal. 2016;47(4):1162-1169. doi:https://doi.org/10.1183/13993003.01618-2015