HCCI’s 2.0 Commercial Claims Research Dataset
What is HCCI?

HCCI’s mission is to *get to the heart* of the key issues impacting the U.S. health care system — by using the best data to get the best answers.

Our values are simple:

- Health care spending information should be transparent and easy to understand.
- Robust analytics can drive improvements in quality and value.
- Health care claims data should be accessible to all who have important questions.
170 million people in the US get health insurance through an employer – half the US population

Largest population, by far, of people with health insurance

- Historically, there has been more data about populations with other kinds of insurance (e.g. Medicare)
- Led to a lack of data and insights about costs and spending in the employer-sponsored insurance market

There are very few large, geographically diverse, multi-payer commercial claims datasets with real cost information in the US

HCCI fills this gap by licensing access to a large dataset of commercial claims to

- Academic researchers at any college or university
- Government agency researchers
- Some research-focused non-profits
Enabling World Class Research

HCCI’s current research partners include:

- Dartmouth
- MD Anderson Cancer Center
- NBER
- MEDPAC
- Weill Cornell Medicine
- Northwestern
- University of Minnesota
- Penn
- Society of Actuaries
- The Commonwealth Fund
- Robert Wood Johnson Foundation
- Laura and John Arnold Foundation
- The Pew Charitable Trusts
Public Reporting: Benchmarks and Resources

Annual Health Care Cost and Utilization Report

- Tracks metrics year-over-year and over 5 years
- Weighted to reflect all people with employer-based insurance
- Downloadable and state-level data tables and data viz tools available

Healthy Marketplace Index
supported by Robert Wood Johnson Foundation

- Tracks metrics year-over-year and various trend timeframes
- Reports for 112 metropolitan areas in the U.S.
- Downloadable data tables and interactive articles available
What Data Does HCCI Provide?
HCCI’s “2.0” Dataset

- Over 1 billion commercial claims per year
- Over 55 million covered lives per year
- Updated annually
- All year of data made available to all researchers
- Years available in 2020: 2012-2018
- All 50 states and D.C.
- De-identified, HIPAA and anti-trust compliant
HCCI provides:

- Streamlined claims data to the fields used for research studies
- Descriptive statistics for each field
- Explanation of fields beyond simple definitions
- Common code repository
  - Data cleaning
  - Grouping claims into utilization measures
  - Identifying patient cohorts
- Crosswalks available alongside data on the server
  - Geographic crosswalks based on 5-digit ZIP code
  - Diagnosis crosswalks for ICD-9 and ICD-10
  - Service category crosswalks for DRG, CPT, NDC
What Makes HCCI Data Different?

- Sourced directly from payers
  - Adjudicated claims with 6+ months of runout
- Covers all ESI lives the payer manages, either as the direct payer or ASO
  - Not sampled from certain employers, plan types, etc.
- Contains real payment information
  - Allowed amounts
  - Out-of-pocket payments
- Contains encrypted patient IDs to enable longitudinal analyses
- Provider IDs – encrypted NPIs
- Five-digit ZIP code of provider and patient
- Data views are statistically deidentified and certified by experts
Many HCCI publications include state-level interactive tools.
HCCI Data Enables Spending and Price Analyses

Figure 2: Factors Contributing to Growth in Spending per Person
(2018 dollars)

Figure 22: Average Price of Inpatient Admissions

Figure 23: Average Out-of-Pocket Price of Inpatient Admissions
Accessing HCCI 2.0 Data
The dataset is being built as we speak

We will open the application process in December 2020

Research teams will gain access to the data starting in January 2020

Pricing details:

<table>
<thead>
<tr>
<th>PER PROJECT ANNUAL ACCESS FEE</th>
<th>PER PROJECT STUDENT ANNUAL ACCESS FEE</th>
<th>INSTITUTIONAL MEMBERSHIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$45,000 (2020)</td>
<td>$15,000 (2020)</td>
<td>Customized for each partner organization</td>
</tr>
</tbody>
</table>

HCCI partners with foundations and organizations to subsidize or sponsor research projects e.g. Health Data for Action
HCCI Data Access: Who and How

WHO?
Researchers affiliated with:
- Academic institutions
- Government agencies
- Some non-profit orgs

HOW?
HCCI reviews proposals for:
- Feasibility and suitability
- Conflicts of interest
- Approach and methodological rigor

HOW?
HCCI reviews exports and work products to ensure:
- Suppression rules are followed
- No material misrepresentation of the data

WHAT ELSE?
HCCI can help research teams disseminate and promote their work

Application
Scientific Review
Data Integrity Review
Publish!

Access Granted
Available Resources and Updates

- Data dictionary and application are now live on HCCI’s website!
- Sign up for HCCI’s newsletter to receive periodic updates
- While you’re on our website…
  - Check out HCCI’s public use files and research studies for examples of what the data can do

HCCI Website:
healthcostinstitute.org

Email Us Anytime:
data@healthcostinstitute.org
Thank you!

Questions?