Comparative Health System Performance Initiative: Compendium of U.S. Health Systems, 2022, Hospital Linkage File, Technical Documentation

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract Number: 47QRAA18D00BQ/75Q80123F80002

Prepared by:

Mathematica, Washington, DC
David Jones
Eric Dehus
Katie Hancock
Laura Kimmey
Eugene Rich

AHRQ Publication No. 24-0011-2-EF
December 2023 (updated February 2024)

This page has been left blank for double-sided copying.

Contents

Acknowledgments....v
I. Introduction....1
II. Data Sources....3
III. Methodology....5
A. Hospital-to-Hospital Matching....5
B. Linking Hospitals to Health Systems....6
IV. Hospital Linkage File Contents....9
V. Caveats and Limitations....11
A. Definition of a Health System....11
B. Differences Across Data Sources....11
C. Multiple Hospitals Sharing the Same CCN....12
D. Mergers and Acquisitions....12
E. HCRIS Data Availability....12
Appendix A. Non-Federal General Acute Care Flag....13
Appendix B. Summary of Hospital Variables Created Using HCRIS Data....15
Appendix C. Data Dictionary – Hospital Linkage File....17

Tables

II.1 Data sources used to identify health systems....3
III.1 Match scores and descriptions....5
IV.1 Hospitals in the Compendium hospital linkage file....9
A.1 Definitions of non-Federal general acute care hospitals....13

This page has been left blank for double-sided copying.

Acknowledgments

We would like to acknowledge a number of organizations and individuals who provided support and guidance during the development of the Compendium of U.S. Health Systems, 2022, Hospital Linkage File:

- Agency for Healthcare Research and Quality (AHRQ): Mike Furukawa, Dan Miller, Leeann Comfort, Jay Crosson, Lingrui Liu, Bruce Seeman, and Doreen Bonnett
- Data vendors: IQVIA™ and the American Hospital Association (AHA)
- Mathematica staff: Miaomiao Shen, Andrew McGuirk, Michael Barna, Linda Molinari, Genna Cohen, Charles Bush, Danielle Whicher, Emily Hague, Caroline Mack, Rachel Machta, and Ken Peckham

This page has been left blank for double-sided copying.

I. Introduction

In 2015, the Agency for Healthcare Research and Quality (AHRQ) created the Comparative Health System Performance (CHSP) Initiative to study how health care systems promote evidence-based practices in delivering care.[i] AHRQ's goal is to understand the factors that affect health systems' use of patient-centered outcomes research (PCOR) and identify best practices in disseminating and using PCOR.

AHRQ supports the ongoing work of the CHSP Initiative by providing this Compendium of U.S. Health Systems, a list of health systems in the United States with details about their structure, staffing, and program participation. The Compendium is available on the AHRQ website: https://www.ahrq.gov/chsp/index.html.

As part of the CHSP Initiative, AHRQ and Mathematica developed a number of publicly available data resources for researchers, policymakers, and other stakeholders who want to understand how health systems can improve the value of healthcare. These data resources include the 2016, 2018, 2020, and 2021 Compendium of U.S. Health Systems and hospital linkage files; 2016 and 2018 group practice linkage files; and now the 2022 Compendium of U.S. Health Systems and hospital linkage file.

AHRQ developed the systems list using information from several data sources that identify systems and their members. In addition to the names and locations, the systems list includes characteristics such as the number of physicians and hospitals in the system and the number of hospital discharges from system hospitals.

AHRQ also created files linking health systems with their member hospitals (referred to as the "hospital linkage file" in this document). The hospital linkage files include information on hospitals, such as hospital name, street address, city, State, and ZIP Code. In addition, the files include hospital identifiers such as the Centers for Medicare & Medicaid Services (CMS) certification number (CCN), health system name, and Compendium health system ID. Starting with the 2020 hospital linkage file, AHRQ added variables describing various hospital characteristics such as the number of discharges, ownership type, and teaching status.

The hospital linkage files are publicly available so users can identify hospitals within health systems and then, using additional data sources, examine aspects of systems and their members, such as cost and quality of care.

This document summarizes the approach taken to create the 2022 hospital linkage file. In section II, we summarize the data sources used to create the file. Section III describes the methodology used to create and refine the file. In section IV, we describe the variables contained in the linkage file. The document concludes with a brief list of caveats that should be considered when using the hospital linkage file.

[i] Additional information on the CHSP Initiative can be found at https://www.ahrq.gov/chsp/index.html.

This page has been left blank for double-sided copying.

II. Data Sources

Although many data sources are useful for studying health system components and their attributes, only a handful explicitly identify systems and indicate relationships among organizations that can be used to identify which collections of organizations constitute a system. These data sources include the following:

- IQVIA OneKey[ii]
- AHA Annual Survey Database

We use these two data sources because they identify hospitals, link hospitals to systems, and are nationally representative. Table II.1 briefly describes each data source, including the name of the data holder, intended purpose of the data, and health system and hospital identifiers.

Table II.1. Data sources used to identify health systems

Source	Data Holder	Intended Purpose	System Identifiers	Hospital Identifiers and Hospital Count
OneKey, December 2022	IQVIA	Reference database for sales/marketing purposes	Integrated delivery network	CCN n=7,636
America Hospital Association (AHA) Annual Survey, 2021	AHA	Membership database for policy research and industry monitoring	Hospital system	AHA ID and CCN n=6,201

The OneKey data include frequent updates of information on health systems, physicians, and hospitals nationwide. The data contain system- and facility-level information on staffing, beds, and facility type. The data also describe relationships that providers have with hospitals and group practices via ownership, management, leasing, purchasing, and contracting mechanisms. Data are collected through a combination of telephone surveys and administrative sources.

AHA data are based on an annual census of hospitals. We used 2021 AHA data because the 2022 data were not available at the time we created the 2022 systems list and hospital linkage file. An overview of the data collection methods for the OneKey and AHA data can be found in the Compendium of U.S. Health Systems, 2022, Technical Documentation.[iii]

[ii] IQVIA maintains two integrated databases relevant to the study of health system performance under the umbrella of Healthcare Relational Services: OneKey Organizations, formerly known as HCOS, and OneKey Professionals, formerly known as HCPS. Throughout the document, we refer to these databases jointly as the OneKey data.

[iii] Available on the Compendium web page at https://www.ahrq.gov/chsp/data-resources/compendium.html.

In addition to the OneKey and AHA data used to identify hospitals and link them to systems, we used CMS hospital cost report data available through the Healthcare Cost Report Information System (HCRIS) to create variables describing various hospital characteristics (e.g., number of discharges, ownership type, and teaching status). We merged the hospitals in the hospital linkage file to the HCRIS data using their CCNs. We describe these variables added to the hospital linkage file in section IV.

III. Methodology

The AHA and OneKey data sources contained 6,201 and 7,636 hospitals, respectively. In this section, we first describe the hospital-to-hospital matching approach to link hospitals in the AHA and OneKey data to create a single set of unique hospitals. We then describe the approach to link hospitals to systems.

A. Hospital-to-Hospital Matching

We matched AHA to OneKey data to identify a set of unique hospitals across the two data sources. First, we matched hospitals using the CCN. Next, using data for hospitals in AHA and OneKey that did not match on CCN, we used the SAS COMPGED function to determine the extent to which hospital names matched. COMPGED assigns a score to each potential match, which reflects the degree to which hospital names match alphanumerically. A COMPGED score of 0 reflects an exact name match; a higher score indicates a lower quality match. Finally, we examined the geographic proximity of hospitals between the paired data sources based on geocoded street address, city, State, and ZIP Code and exact text matches on city, State, and ZIP Code.

The name and geographic proximity matching often resulted in a one-to-many match for each pairwise comparison; for example, many AHA hospitals had multiple potential OneKey hospital matches based on name or geography. We assigned a match score to each potential match in order of the quality of the match (lower scores indicate higher quality matches), as shown in Table III.1.

All hospital pairs with a match score of 1 (CCN matches) were assessed as "matched" and set aside as matches. For the rest of the matches, we identified a best possible match using the lowest match score and closest geographic proximity. Next, for each best possible match, we visually inspected the hospitals' names and addresses, and when necessary, conducted a web search to confirm whether the match was correct. For example, if a particular hospital in IQVIA had two matches with a match score of 3 and two matches with a match score of 4, we visually inspected the match that had a match score of 3 and the closest geographic proximity of the two hospitals. When necessary, we conducted a web search to confirm whether the two hospitals were in fact the same hospital.

Table III.1. Match scores and descriptions

Match Score	Description
1	Exact match on CCN
2	COMPGED≤150 and (ZIP Codes match or street names match and hospitals are within 1 mile of each other)
3	Hospitals are within one-half mile of each other
4	COMPGED≤150 on full hospital name (e.g., Jones Medical Center)
5	COMPGED≤150 on truncated hospital name (e.g., Jones Medical)
6	Hospitals are within 10 miles of each other
7	Exact numeric match on ZIP Code
8	Exact text match on city and State

We identified 7,946 unique hospitals between the two data sources. We excluded 1,107 hospitals from the hospital linkage file that did not report a CCN or an AHA ID. We excluded them because, although they did not match any other hospitals based on name or address, without a CCN or AHA ID, we cannot be confident that they are unique hospitals. In addition, without a CCN or AHA ID, it will be difficult for users of the linkage file to confidently match these hospitals to other data sources for analysis. We then excluded 75 hospitals that were not located in one of the 50 states or the District of Columbia. These exclusions resulted in a hospital linkage file with 6,764 unique hospitals. Among those hospitals, 5,570 matched across the two data sources on the CCN. An additional 254 hospital matches were the result of matching on name and geographic proximity. Finally, 940 hospitals in the hospital linkage file did not match with a hospital from another source, based on CCN, name, or location.[iv]

B. Linking Hospitals to Health Systems

As described in the Compendium of U.S. Health Systems, 2022, Technical Documentation, hospitals need to be assigned to health systems for two reasons.[v] First, health systems contained in the Compendium are required to have at least one non-Federal general acute care hospital. Second, the list of health systems includes system attributes that are hospital-level variables aggregated to the system level, such as:

- Number of hospitals per system,
- Number of general acute care hospitals per system,
- Extent to which system hospitals are located in multiple States, and
- Number of hospital discharges per system.

Both data sources identify hospitals included in systems. Before linking hospitals to health systems, we excluded those that did not report a CCN or an AHA ID or were not located in one of the 50 states or the District of Columbia. To assign the remaining 6,764 hospitals to health systems, we applied the following rules:

1. We linked hospitals reported as being assigned to only one health system to that system.[vi,vii] Of the 6,764 unique hospitals with a CCN or AHA ID reported across the two data sources, several hundred were assigned to two or more systems.

[iv] If a hospital had the same address and name in IQVIA and AHA data but different CCNs, we reported the IQVIA and AHA information in separate observations.

[v] Available at https://www.ahrq.gov/chsp/data-resources/compendium.html.

[vi] Hospitals linked to one health system include those in which both data sources agree on the system assignment, as well as cases in which one of the data sources does not list the hospital in a system. For example, if one of the data sources assigns a hospital to a system, and the other data source does not assign a system for the hospital, it is listed as being in the system.

[vii] We identified one hospital that was incorrectly assigned to a system in the underlying data. Caldwell Memorial Hospital in Louisiana was incorrectly assigned the University of North Carolina Health Care System. We removed the linkage between this hospital and system.

2. We manually reviewed the hospitals identified as being assigned to two or more health systems due to discrepancies between the data sources. In most cases, the multiple systems were in fact the same system with a different name or systems nested within each other (that is, subsystems and parent systems). In the former cases, we updated the list of systems to indicate that these systems were a match and the hospital belonged to this system; in the latter case, we assigned the hospital to the parent system. We also identified cases in which a change in ownership occurred, which was reflected in one data source but not another. We updated these linkages to reflect the change.
3. The remaining cases were hospital-level joint ventures, in which multiple systems have a formal relationship with a hospital. In these cases, we linked the hospitals to a system using the following decision rules, in order:
    - When it was clear, based on a manual review of systems' and hospitals' websites, that one system was the majority owner or taking responsibility for running the day-to-day operations of the hospital, we linked the hospital to that system.
    - In the absence of other information, we linked the hospital to the system whose headquarters location was closest to the location of the hospital; that is, we aimed to link it to a local system over regional or national systems.

After applying the exclusion criteria to remove systems that do not meet the Compendium definition of a health system and linking hospitals in subsystems to their parent systems, we had 4,173 hospitals linked to the 640 systems in the Compendium of U.S. Health Systems, 2022. (For a detailed description of the exclusion and linking process, see the Compendium of U.S. Health Systems, 2022, Technical Documentation.)

The result of linking hospitals to systems is a hospital linkage file that contains one row for each hospital with a CCN or AHA ID, along with its identifying information (CCN, name, and address). In addition, if the hospital is in a Compendium system, the file includes the identifying information of its system. No hospital is linked to more than one health system.

This page has been left blank for double-sided copying.

IV. Hospital Linkage File Contents

The hospital linkage file contains 6,764 hospitals (Table IV.1). A total of 4,173 hospitals are linked to Compendium health systems, of which 3,606 are non-Federal general acute care hospitals.[viii] The hospital counts within the Compendium list of health systems sum to match the counts linked to systems in the hospital linkage file. An additional 2,591 hospitals contained within the hospital linkage file are not linked to a Compendium health system, and 1,127 of these are non-Federal general acute care hospitals. The remaining 2,031 hospitals are specialty hospitals, such as surgical, long-term acute care, and rehabilitation hospitals.

Table IV.1. Hospitals in the Compendium hospital linkage file

Hospital Group	N
Total hospitals in Compendium linkage file	**6,764**
Linked to Compendium health systems	**4,173**
Non-Federal general acute care hospitals	3,606
Not linked to Compendium health systems	**2,591**
Non-Federal general acute care hospitals	1,127

The hospital linkage file contains 25 variables, including each hospital's name; location (address, city, State, and ZIP Code); a flag for non-Federal general acute care hospitals; a unique hospital ID assigned as part of the development of the Compendium; and, if applicable, their linked Compendium health system ID, name, and location.[ix]

The hospital linkage file contains two linking variables. Medicare CCN (ccn) can be used to link hospitals to external data sources. The unique health system identifier (health_sys_id) can be used to link the hospital to the systems in the Compendium of U.S. Health Systems. The linkage file does not include the AHA ID, as release of this variable is not permitted under our AHA data use agreement.

In addition, starting in the 2020 hospital linkage file, AHRQ added variables created using HCRIS data that describe various hospital characteristics, including:

- Number of hospital beds,
- Number of hospital discharges,
- Number of full-time-equivalent interns and residents,
- Children's hospital flag,
- Major teaching hospital flag,
- Very major teaching hospital flag,
- Teaching intensity (resident-to-bed ratio),

[viii] See Appendix A for a description of how we identified non-Federal general acute care hospitals.

[ix] When hospital names or locations differed across data sources, we reported information in the hospital linkage file from the AHA and OneKey data, in priority order.

- High Disproportionate Share Hospital (DSH) patient percentage (top quintile among hospitals),
- Uncompensated care burden,
- High uncompensated care burden (top quintile among hospitals),
- Ownership type,
- Total patient revenue, and
- Net patient revenue.

The values for these variables are missing when the hospital does not appear in the HCRIS data or the specific values are missing in the data. These variables are all hospital-level versions of the system-level variables included in the Compendium of U.S. Health Systems.[x] Appendix B contains descriptions of each of the hospital-level variables created using the HCRIS data. Appendix C contains a data dictionary for all variables included in the hospital linkage file.

[x] See the Compendium of U.S. Health Systems, 2022, Technical Documentation for descriptions of the system-level variables: https://www.ahrq.gov/chsp/data-resources/compendium.html.

V. Caveats and Limitations

This release of the hospital linkage file enables users to link Compendium health systems with their member hospitals. When using the linkage file, users should bear in mind a few caveats and limitations to the current methods for assigning hospitals to health systems.

A. Definition of a Health System

The Compendium definition of a health system follows:

> *A health system includes at least one hospital and at least one group of physicians that provides comprehensive care (including primary and specialty care) who are connected with each other and with the hospital through common ownership or joint management.*[xi]

This definition is further operationalized as requiring that a health system:

1. Include at least one non-Federal acute care hospital,
2. Include, in total, at least 50 physicians, and
3. Include at least 10 primary care physicians.

Hospitals within organizations that do not meet this definition are not identified in the hospital linkage file as being part of a system. Also, hospitals in the hospital linkage file are associated with parent systems; relationships between hospitals and subsystems are not identified.

B. Differences Across Data Sources

We attempted to maximize the coverage of hospitals in the United States by using multiple data sources. However, the data sources used in the development of the hospital linkage file vary in ways that affect their characterization of hospitals and their linkages to systems (for example, they use different data collection methods). Also, AHA data were from 2021, whereas the OneKey data were from 2022. In addition, some hospitals may be missing from the two data sources, and thus would not be reflected in the hospital linkage file.

Similarly, we used the AHA and OneKey data to determine whether hospitals were designated as a non-Federal general acute care hospital. However, differences in this designation across data sources may result in some general acute care hospitals not being flagged as such or hospitals being flagged that are not general acute care hospitals (see Appendix A).

Finally, the mechanism and criteria for assigning hospitals to health systems differ by data source, which leads to differences in linkages. We encountered differences in health system

[xi] Foundation models of health system organization are considered a form of joint management. Joint participation in an accountable care organization is not by itself indicative of joint management. In addition, "group" is not synonymous with a separately organized medical group. A hospital that employs community-based physicians who provide comprehensive care (but are not organized as a medical group) would be considered a health system.

assignments across data sources that we had to adjudicate. Also, we take a fairly inclusive approach, only requiring that one data source indicate a hospital is assigned to a system.

C. Multiple Hospitals Sharing the Same CCN

The Compendium hospital linkage file does not include hospitals reporting the same CCN as separate entries in the hospital linkage file, for example, separate hospital facilities or campuses with different sites of care under the same parent hospital. Rather, the file includes a single entry for the CCN in these cases. Reporting by CCN, and thus reporting some hospitals together in a single entry in the linkage file, is a limitation for those aiming to identify all individual hospital facilities and campuses and their linkages to health systems. However, there are data sources available (for example, the AHA units file) that split some CCNs into multiple hospitals.

Those with access to data that disaggregates hospitals within CCNs can use the linkage file to assign the multiple hospitals under a CCN to the correct system; for example, if CCN 111 includes hospital 222 and 333, and the CCN is in a system, hospitals 222 and 333 can be assigned to the system. This approach assumes that all hospitals under a CCN are linked to the same system.[xii]

D. Mergers and Acquisitions

The linkage file reflects hospitals in the United States at the end of 2022. However, because OneKey and AHA data vary in the periodicity of their updates, lags may occur in updating changes to systems and their hospitals, such as mergers, acquisitions, and name changes that occurred before the end of 2022. Thus, the period represented by the hospital linkage file aligns with the periods covered by the data sources (calendar year 2021 for AHA and the end of 2022 for OneKey), with some level of updating accomplished through manual review of hospital-system linkages that varied by data source.

E. HCRIS Data Availability

As described in Section IV, the hospital linkage file contains information from HCRIS reports for each hospital. However, because the HCRIS data only include information extracted from reports submitted, validated, and verified as of the time of download, some hospitals are not present in the data. Overall, 979 (14 percent) hospitals included in the linkage file did not have a report in HCRIS.

Among all hospitals in Compendium health systems, 429 hospitals (10 percent) were not in the HCRIS data; among the non-Federal general acute care hospitals in systems, 293 hospitals (8 percent) were not in the data. The majority of systems (450 out of 640) have HCRIS data for all of their component hospitals.

[xii] We include hospitals with AHA IDs but missing CCNs. The hospital linkage file has 132 hospitals without CCNs but with AHA IDs. Many of these (although not all) are State hospitals (such as correctional facility hospitals), Federal hospitals, or children's specialty hospitals. It is possible these facilities do not have a CCN because they do not bill Medicare.

Appendix A. Non-Federal General Acute Care Flag

In the Compendium of U.S. Health Systems, 2022, all health systems were required to have at least one non-Federal general acute care hospital. The hospital linkage file includes a flag that identifies non-Federal general acute care hospitals—those in Compendium systems and not in systems. In this appendix, we provide additional details about the construction of the non-Federal general acute care hospital flag.

The non-Federal general acute care hospital flag in the hospital linkage file denotes hospitals that were identified as non-Federal general acute care hospitals in at least one data source (AHA or IQVIA) used to develop the Compendium of U.S. Health Systems, 2022. Table A.1 describes how non-Federal general acute care hospitals were identified in each data source.

Table A.1. Definitions of non-Federal general acute care hospitals

Data Source	Description of Non-Federal General Acute Care Hospitals
AHA	(1) Not Department of Defense, Public Health Service, Veterans Affairs, Federal other, Indian Health Service, or Department of Justice facilities and (2) the hospital provides one or more of the following services: general medical and surgical or children's general medical and surgical care or is a critical access hospital or a major or minor teaching hospital based on resident-to-bed ratio.
IQVIA	(1) Not a government or Veterans Affairs-owned/run facility and (2) flagged as an acute care general hospital, critical access hospital, or children's hospital.

Note: Beginning in 2020, the AHA Annual Survey Database combined Air Force, Army, and Navy under the Department of Defense. In previous years, they were reported separately.

If a hospital was considered a general acute care hospital in at least one of these data sources, we flagged it as such in the hospital linkage file.

This page has been left blank for double-sided copying.

Appendix B. Summary of Hospital Variables Created Using HCRIS Data

Hospital Variable	Definition
Number of hospital beds	Count of total hospital beds reported in the HCRIS data.
Number of hospital discharges	Count of total hospital discharges reported in the HCRIS data.
Number of full-time-equivalent interns and residents	Count of total full-time-equivalent interns and residents reported in the HCRIS data.
Children's hospital flag	A flag indicating that the hospital is identified as a children's hospital in the HCRIS data.
Major teaching hospital flag	A flag indicating that the hospital is a major teaching hospital, defined as a resident-to-bed ratio greater than or equal to 0.25.[xiii]
Very major teaching hospital flag	A flag indicating that the hospital is a very major teaching hospital, defined as a resident-to-bed ratio greater than or equal to 0.60.
Teaching intensity (resident-to-bed ratio)	A ratio of the number of residents to the number of beds at a hospital. Note, a minor teaching system is typically defined as a resident-to-bed ratio greater than zero but less than 0.25.
High Disproportionate Share Hospital (DSH) patient percentage (top quintile among hospitals)	The DSH patient percentage is defined as: DSH Patient Percentage = (Medicare SSI days/total Medicare days) + (Medicaid, non-Medicare days/total days). Hospitals paid under the Inpatient Prospective Payment System (IPPS) report the DSH patient percentage to CMS, which uses it to calculate Medicare DSH payments. It is a measure of the hospital's overall caseload of low-income insured patients. A hospital is categorized as having a high DSH patient percentage if their value is in the top quintile[xiv] among IPPS hospitals.
Uncompensated care burden	Uncompensated care is defined as the sum of the hospital's charity care and bad-debt expense adjusted by the hospital-specific ratio of cost to charges. Uncompensated care burden is calculated as the ratio of uncompensated care to total operating expense.
High uncompensated care burden (top quintile among hospitals)	A hospital is categorized as having a high uncompensated care burden if they are in the top quintile of uncompensated care burden among all hospitals in the data.[xv]

[xiii] Thresholds used to identify hospitals with high teaching intensity were selected to align with previous literature, including Patel MS, Volpp KG, Small DS, et al. Association of the 2011 ACGME resident duty hour reforms with mortality and readmissions among hospitalized Medicare patients. JAMA 2014;312(22):2364-73. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5546100/. Accessed January 22, 2024.

[xiv] Quintiles are used to align with prior health services research on defining the safety net, including: Office of the Assistant Secretary for Planning and Evaluation. Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs: a Report Required by the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014. Washington, DC: U.S. Department of Health and Human Services; December 2016. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//171041/ASPESESRTCfull.pdf. Accessed January 22, 2024.

[xv] This definition extends work from previous literature on defining safety net hospitals, including Zuckerman S, Bazzoli G, Davidoff A, et al. How did safety-net hospitals cope in the 1990s? Health Aff 2001;20(4):159-68; and Bazzoli GJ, Kang R, Hasnain-Wynia R, et al. An update on safety-net hospitals: coping with the late 1990s and early

Hospital Variable	Definition
Ownership type	Type of ownership reported in the HCRIS data. This is a categorical variable taking the following values: nonprofit=1, public/government=2, church-operated=3, and for-profit/investor owned=5.
Total patient revenue	Hospital's total patient revenue.
Net patient revenue	Hospital's total patient revenue less contractual allowances and discounts on patients' accounts.

2000s. Health Aff 2005;24(4):1047-56. https://www.healthaffairs.org/doi/10.1377/hlthaff.24.4.1047. Accessed January 22, 2024. In these studies, high uncompensated care burden for hospitals was defined as the top decile of uncompensated care burden.

Appendix C. Data Dictionary – Hospital Linkage File

Variable Name	Variable Type	Description
compendium_hospital_id	Character	Unique hospital ID created by the CHSP Initiative and used across all CHSP analyses[a]
ccn	Character	CMS certification number (hospital ID)
hospital_name	Character	Hospital name
hospital_street	Character	Hospital street address
hospital_city	Character	Hospital city
hospital_state	Character	Hospital State
hospital_zip	Character	Hospital ZIP Code
acutehosp_flag	Numeric	Flag for non-Federal general acute care hospitals
health_sys_id	Character	Unique Compendium health system ID (assigned by the CHSP Initiative) of the system linked to the hospital (same variable as in the Compendium of U.S. Health Systems)
health_sys_name	Character	Health system name (same variable as in the Compendium of U.S. Health Systems)
health_sys_city	Character	Health system city (same variable as in the Compendium of U.S. Health Systems)
health_sys_state	Character	Health system State (same variable as in the Compendium of U.S. Health Systems)
hos_beds	Numeric	Number of hospital beds
hos_dsch	Numeric	Number of hospital discharges
hos_res	Numeric	Number of full-time-equivalent interns and residents
hos_children	Numeric	Children's hospital flag
hos_majteach	Numeric	Major teaching hospital flag
hos_vmajteach	Numeric	Very major teaching hospital flag
hos_teachint	Numeric	Teaching intensity (resident-to-bed ratio)
hos_highdpp	Numeric	High Disproportionate Share Hospital (DSH) patient percentage (top quintile among hospitals)
hos_ucburden	Numeric	Uncompensated care burden
hos_highuc	Numeric	High uncompensated care burden (top quintile among hospitals)
hos_ownership	Numeric	Ownership type
hos_net_revenue	Numeric	Total patient revenue
hos_total_revenue	Numeric	Net patient revenue

[a] If a 2022 hospital had the same CCN and/or AHA identification number as a 2021 hospital, then we assigned the 2022 hospital the same unique hospital identification number as 2021.