LOWN INSTITUTE
HOSPITALS INDEX

White Paper Series

Lown Institute Hospitals Index for Social Responsibility

2024 Methodology

About this White Paper

This white paper is part of a series analyzing specific metrics in the Lown Institute Hospitals Index for Social Responsibility. This paper was written by Vikas Saini, Shannon Brownlee, Valérie Gopinath, Paula Smith, Kelsey Chalmers, and Judith Garber.

Suggested Citation: Saini, V., Brownlee, S., Gopinath, V., Smith, P., Chalmers, K., Garber, J. 2024 Methodology of the Lown Institute Hospitals Index for Social Responsibility. Needham, MA: The Lown Institute. 2024.

Contents

Overview

About the Lown Institute Hospitals Index

The Lown Institute Hospitals Index (LIHI) is the first ranking of hospital social responsibility, evaluating nearly 4,000 hospitals on their performance across health outcomes, value, and equity.

The Lown Institute Hospitals Index, published annually since 2020, has up to 54 metrics distributed across four tiers (see Figures 1 & 2 below).

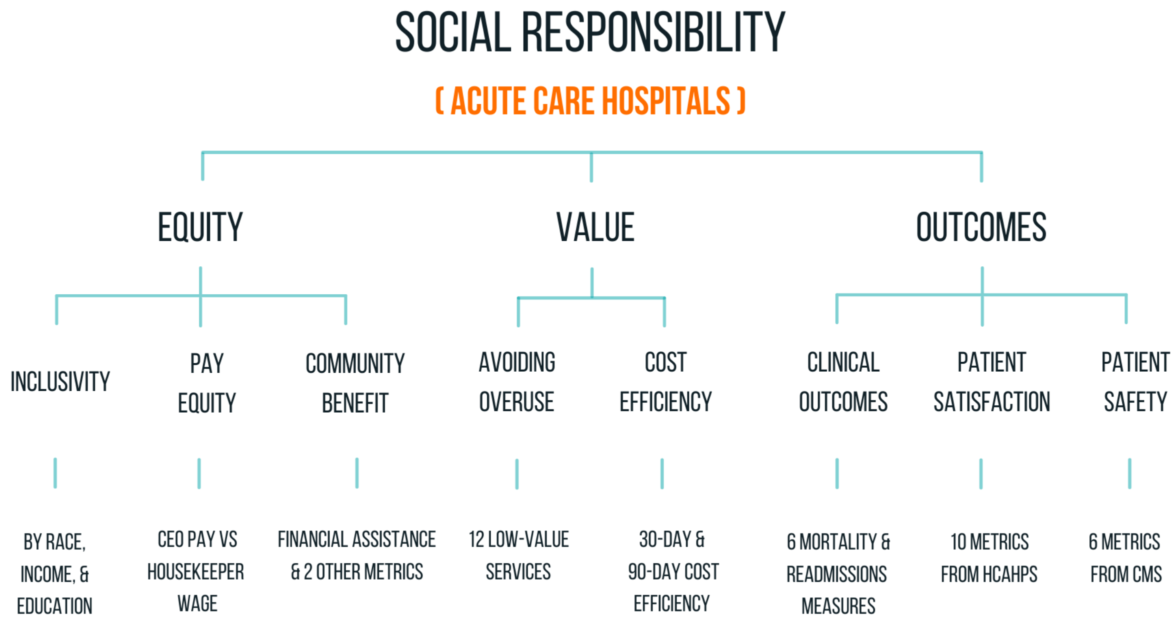

Figure 1: The Lown Institute Hospital Index for Social Responsibility metrics diagram for Acute Care Hospitals.

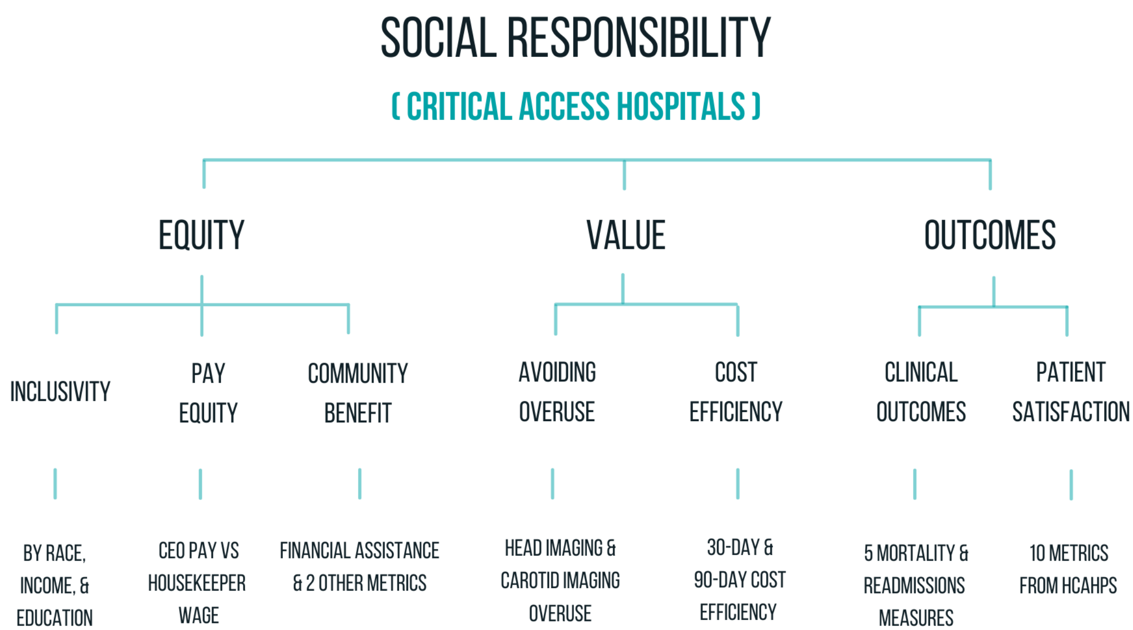

Figure 2: The Lown Institute Hospital Index for Social Responsibility metrics diagram for Critical Access Hospitals.

New on the Index in 2024

- This year we created separate rankings for short-term Acute Care and Critical Access Hospitals. Therefore ranks and grades are now relative to their peer group, which we felt was a more fair assessment for these distinct hospital types. The rankings consist of 2,784 Acute Care Hospitals and 896 Critical Access Hospitals. We identified the hospital type designation based on CMS Care Compare as of January 2024. In the system rankings, we have one set of results since systems can include both Acute Care and Critical Access Hospitals.
- To calculate our composite social responsibility grade, we increased the weight of equity (from 30% to 40%) and decreased the weight of outcomes (from 40% to 30%). Previously, outcomes and value had a disproportionate effect on the final grade due to their internal correlation with each other. Now hospitals' final social responsibility result is more correlated with their equity performance as we had originally intended.
- We added a new threshold for hospitals' social responsibility results. In order to receive an A on social responsibility, hospitals cannot have a C or lower grade on either equity, value or outcomes.
- We added a new Tier 4 measure: unplanned hospital visits post outpatient surgery within clinical outcomes.
- 246 fewer hospitals are included in the LIHI rankings compared to last year; a total of 3,680 hospitals in the full dataset. One of the main reasons fewer hospitals are included this year is the requirement that hospitals have sufficient Medicare claims to be eligible for a clinical outcomes score.
- We adjusted the grade cut-offs for our inclusivity measure. Previously, we awarded As to those hospitals which had much higher proportions of patients from minoritized communities or low income or educated areas, relative to their surrounding community demographics. Now, hospitals will also receive an A if their patient demographics are near equal to their community demographics.
- We adjusted the methodology for our pay equity measure. Previously, we calculated a ratio of CEO pay to average worker pay where the worker pay was based on nonmedical HCRIS wage categories. In this year's LIHI, we calculated a ratio of CEO hourly pay to the housekeeping average hourly wage to better reflect the lowest paid workers.

- In the community benefit detailed measure for financial assistance, we made use of prior year's HCRIS data when the most recent year was not available. In the detailed measure for community investment we used 2022 IRS data when 2021 data was not available. For financial assistance, we adjusted the score cutoffs to better distinguish between 1-star and 2-star hospitals.
- We included Critical Access Hospitals in this year's overuse rankings (see *Avoiding overuse* for more).

Criteria for hospitals to be included in the Index

The Lown Hospitals Index for Social Responsibility includes 3,680 hospitals (2,784 Acute Care Hospitals and 896 Critical Access Hospitals) in the US.

Non-Acute Care and non-Critical Access hospitals, federal hospitals (e.g. Veterans Health Administration) and those outside of the 50 states and Washington, D.C. were excluded. We restricted hospitals to those in the January 2024 update of Care Compare, a website run by the Centers for Medicare and Medicaid Services (CMS). We selected hospitals based on hospitalizations of beneficiaries age 65 and over during 2019 through 2022: we used Medicare inpatient fee-for-service (FFS) claims from 2020-2022 as well as Medicare Advantage inpatient encounter data from 2019-2021. We excluded specialty hospitals with more than 45% admissions for orthopedic, more than 45% for cardiac, more than 80% surgical procedures, more than 80% elective surgeries (among hospitals with > 45% surgical procedures). We further eliminated hospitals that were closed as of February 2024 by checking against press releases/news reports on closures. Hospitals with patient volume below 50 annual patient stays were also eliminated as well as hospitals that did not perform any surgery. This left a list of 3,680 hospitals: 523 for-profits and 3,157 nonprofits.

Hospital characteristics

We defined Safety Net hospitals as the top 20% of hospitals in the Index based on their dual-eligibility ratio, the proportion of patient stay days eligible for both Medicare and Medicaid.

Information on hospital characteristics was taken from the CMS Care Compare 2024 database and internally verified against commercial datasets.

Data sources for the Index

Table 1: Source(s) used for each component metric in the Lown Index.

Metric	Year(s)	Source(s)
Pay Equity	2021 (Fiscal Year Ending)	IRS Form 990, SEC Filings, Healthcare Cost Report Information System (HCRIS) hospital data, state databases with public employee salaries
Community Benefit	2021 (Fiscal Year Ending)	IRS Form 990, HCRIS hospital data
Inclusivity	2021-2022	Census Bureau Data (American Community Survey), CMS Medicare claims, CDC Facility ID data set
Avoiding Overuse	2019-2022	Medicare FFS claims 2020-2022 and Medicare Advantage encounter files 2019-2021 inpatient, outpatient, and carrier
Cost Efficiency	2020-2022	Medicare FFS inpatient, outpatient, and carrier claims data
Clinical Outcomes	2019-2022	Medicare FFS claims 2020-2022 and Medicare Advantage encounter files 2019-2021 inpatient, outpatient, and carrier with 1-year lookback
Patient Safety	July 2020 to March 2023	CMS Patient Safety Indicators and Health care Associated Infections from Care Compare
Patient Satisfaction	April 2022 to March 2023	CMS Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey

Equity

Our equity measure includes inclusivity, community benefits, and pay equity (weighted 40, 40 and 20% respectively).

Pay equity

For pay equity, we obtained data for Chief Executive Officer (CEO) compensation from three different sources corresponding to the tax status of the hospital. Compensation data on for-profit, publicly-traded hospital systems was obtained from Securities and Exchange Commission's (SEC) Edgar database. Public, non-federal hospital CEO salaries were gleaned from available payroll data and other public records. For nonprofit hospitals required to file with the IRS, we accessed the IRS 990 filings from the IRS website.

Last year, we generated our own comprehensive dataset that linked CMS hospital data to IRS tax filings. We leveraged this previous year's mapped data to generate our hospital tax dataset. This left 122 hospitals that were either new to our LIHI hospitals list or filed under a different tax identification number for fiscal year ending 2021. These were subsequently mapped to IRS filings using our address algorithm. Using the IRS dataset and checking against commercial datasets, we used text matching algorithms to identify CEO names and then manually verified the result. We were able to find 2,207 hospital CEO salaries of the original 2,476 (89.1%) hospitals using this strategy.

In cases in which CEO pay was unavailable, values were imputed using predictive mean matching (PMM) based on CEO pay for nonprofit hospitals, combined with other variables such as bed size and hospital revenue. Pay for 2,480 nonprofit hospitals was used to impute values for three populations with unavailable pay: 523 for-profit hospitals, 406 public hospitals, and 270 nonprofit hospitals whose 990 forms did not contain the full executive compensation information. Since the imputation dataset was composed of only nonprofit hospitals, we applied additional compensation ratios to PMM results for both for-profit and public hospitals. The for-profit compensation ratio was calculated using the average for-profit system salary divided by the average nonprofit system salary. For public hospitals, the compensation ratio was based on the average public hospital salary divided by the average nonprofit hospital salary. For hospitals within systems (two or more hospitals), we distributed the system CEO's salary among the constituent hospitals using the percentage of total revenue each hospital generated.

We obtained average housekeeping wages from the CMS Healthcare Cost Report Information System (HCRIS). We included both housekeeping and housekeeping under contract wage categories in our review, deferring to the category which had the higher total hours logged. To avoid sparsity issues, we reviewed 2020-2022 HCRIS wage index records for housekeeping and used 2021 wage data when available. For hospitals that had no housekeeping wage information in HCRIS for 2020-2022, we substituted an average of housekeeping HCRIS data from hospitals in the same state and with the same urban or rural status. We then estimated hourly wages for CEOs based on the work hours listed in their IRS forms, defaulting to an average of 46.69 when the hours were not listed, and calculated a ratio of CEO pay to average worker pay.

Of the full list of 3,680 hospitals, 4 hospitals were not ranked due to missing HCRIS data.

Star rating and grade

In order to calculate stars for pay equity, the pay ratios were first winsorized and then z-normed. The range for those z-scores were split into five regions, where hospitals in the lowest regions were assigned five stars and hospitals in the highest region were assigned 1 star. This approach applied to both Critical Access and Acute Care hospitals.

The winsorized Z-scores were assigned letter grades based on where those scores fell on a normal distribution curve using 15%, 35%, 75% as the cutoffs. As before, this approach applied to both Critical Access and Acute Care Hospitals.

Pay equity limitations

Data anomalies may exist if different hospitals reported the same person with a name other than their legal name, or added middle initials to name, as well as any number of text-based inconsistencies within tax records or public documentation. We have done our best to minimize these issues using algorithms and manual review.

In regards to salary imputation, the linear regression extrapolations were dependent on the nonprofit hospital salary population for modeling, since nonprofits were the most represented hospital type. With respect to for-profit hospitals, only system-level information was available through SEC filings. The for-profit compensation ratio assumes that the gap between the salaries of for-profit hospital system CEO salaries versus nonprofit hospital system CEO salaries disseminates down to the individual hospital level. This assumption is dependent on the nature of the data available. There are a number of factors that go into the determination of a hospital CEO's salary and we mainly used revenue as our basis for estimation.

The average HCRIS wage substitution may not reflect the specific market characteristics of a hospital without wage index data since the substitution only relies on an average for states and urban or rural status.

Community benefit

The community benefit metric measures hospital spending on financial assistance and community health initiatives, as well as their service of Medicaid patients. Community benefit is a composite of three details: financial assistance, Medicaid revenue, and community investment, which included several categories of community benefit spending that we deemed to be meaningful. We did not use several other categories of community benefit reported on 990 forms, including: shortfall from Medicaid and other government means-tested insurance programs (shortfall is the difference between the amount Medicaid or other programs pay and the costs hospitals claim for caring for such patients); health professionals training (which is already largely subsidized by the federal government); and research. For this metric, our goal was to focus on spending that directly benefits community health and the upstream factors that affect it.

Out of the 3,680 hospitals in the full LIHI data set, we ranked 3,629 hospitals on community benefit (2,752 Acute Care Hospitals and 877 Critical Access Hospitals). We measured and ranked the Acute Care Hospitals and Critical Access Hospitals separately. For 51 hospitals with missing data for two or three details, they were not ranked on community benefit. For 2,430 hospitals with data available for all three metrics, each metric was z-normed and weighed equally in the composite at one-third of the total community benefit score. For 1,199 hospitals with data for two of the metrics available, each metric was z-normed and weighed equally in the composite as half of the total score.

Financial assistance

Financial assistance (also known as *charity care*) is free or discounted care provided to patients eligible for assistance based on their income. We measured spending on financial assistance as a share of total hospital expenses as reported in the Centers for Medicare and Medicaid's Hospital Cost Reports (HCRIS) for Fiscal Year Ending 2021.

We adjusted hospital financial assistance spending based on hospitals' three-year average patient services margin using HCRIS data from 2019-2021. We expect that hospitals with more of a financial cushion are able to give more assistance. The financial assistance scores were weighted based on the Z-scored financial assistance spending (80%) of the total score and Z-scored average patient services margin (20%). We ranked 3,627 hospitals on financial assistance; data for 53 hospitals were unavailable. For hospitals without financial assistance or expenses data for 2021, we used HCRIS data from 2020.

Community investment

We measured hospital spending on community investment, as a share of total functional expenses, using Fiscal Year Ending 2021 or the most recent year available of Internal Revenue Service (IRS) tax filings. Private nonprofit hospitals are required to report community benefit spending to the IRS to maintain nonprofit status. IRS data on these community benefits were available for 2,466 hospitals. For 22 hospitals, 2022 IRS was used, as 2021 was not available.

Data for for-profit and government hospitals were not available, as these types of hospitals are not required to file Form 990. For 33 hospitals that filed with a Schedule E (indicating university), their total operating expenses were taken from financial audits when available or from HCRIS rather than from the IRS form. Hospitals that reported spending over 100% of their budget, and hospitals that file with foundations were excluded.

Community investment comprises a subset of hospital community benefit spending including the following categories on IRS Form 990 Schedule H:

- Community Health Improvement Activities: Programs to improve community health, such as health education programs, immunizations, blood pressure screening, etc.
- Subsidized healthcare services: Health services provided to the community at a loss to the hospital, that serve a specific community need and would not exist without the hospital providing it. These often include mental health and substance use programs, labor & delivery, burn units, and others.
- Contributions to community groups: Grants and in-kind contributions to community groups
- Community building activities: Programs that address the "upstream" factors that impact health, such as housing, environment, and food security.

Fair Share Spending

The Lown Institute publishes a report annually on hospital fair share spending, which is not included within the Social Responsibility composite score. Fair share spending compares spending on financial assistance and community investment with the estimated value of hospital tax breaks. This year, this metric was reported only for private nonprofit hospitals with IRS data for fiscal year ending 2021 (2,425 hospitals included).

Hospitals that dedicated at least 5.9 percent of overall functional expenses to financial assistance and meaningful community investment were considered to have spent their fair share. The 5.9 percent threshold is based on established research into the valuation of the nonprofit tax exemption. Hospital expenses were retrieved from hospitals' IRS Form 990. For hospitals filing with universities where Schedule E was submitted, financial audit or CMS cost report information was used to calculate expenses and net income.

Spending on financial assistance and community investment was retrieved from hospitals' IRS Schedule H Form 990 (NOTE: This is different from the LIHI financial assistance metric which is sourced from CMS cost reports). For group filings, hospital expenses and community investment spending were prorated across hospitals based on their share of system revenue.

For spending on community investment, we include a subset of community benefit categories reported to the IRS that provide a direct benefit to community health (see section on *community investment* above).

Medicaid revenue

In general, hospitals in states that expanded Medicaid spend less on financial assistance because fewer patients need financial assistance. To account for hospitals' service of Medicaid patients and differences in state policy, we included a metric to estimate the proportion of the hospital's patients that are covered by Medicaid. We measured net Medicaid revenue as a proportion of net patient revenue using HCRIS data for fiscal year ending 2021. Hospitals with a greater share of patient revenue from Medicaid were ranked better on this metric. We ranked 3,643 hospitals on Medicaid patient revenue; 37 hospitals did not have data available. For hospitals without 2021 Medicaid data available, Medicaid revenue that was negative or above 100% of patient revenue, or negative net patient revenue, we instead used 2020 HCRIS data where available.

Star rating

To assign stars to Acute Care Hospitals for each detail (financial assistance, Medicaid revenue, and community investment), we calculated the ranges of the Z-scores and divided this into five regions. Hospitals in the highest region were assigned 5 stars (the best performers), while hospitals in the lowest region were assigned 1 star. With this method, many hospitals may receive the same number of stars if their raw scores are similar.

This year for financial assistance we Winsorized the Z-scores at the highest range, replacing values more extreme than the 95th percentile, to make the star ratings more meaningful, as most of these raw values are on the low end.

We used the same method to assign stars to Critical Access Hospitals.

Community benefit grade

For Acute Care Hospitals, the composite community benefit score was the averaged result of a hospital's detailed scores (equally weighted). Hospitals needed to have at least two of the detail scores to receive a composite score. To determine the grade cut-offs, we selected values using a standard normal distribution for the quantiles 15%, 35%, 75%. Hospitals with a Z-score in this top quantile received an A, those in the next quantile received a B, and so on. For Critical Access Hospitals, we calculated the composite composite community benefit score and assigned grades using the same approach.

Community benefit limitations

The measurement of community benefits spending by hospitals has improved since the 2010 Affordable Care Act (ACA) clarified reporting requirements for IRS Form 990, which all hospitals must use for reporting their spending. However, there are still several limitations to the data that are available. For hospitals that did not file a Form 990, the score was based on financial assistance as a share of total expenses and share of Medicaid revenue, but we could not take into account other types of community benefits. Therefore, community benefit spending by public and for-profit hospitals may be undercounted.

We used CMS's HCRIS data set to be able to compare financial assistance spending and Medicaid revenue across hospital types; however, this data set also has potential limitations. Hospitals are not required to have the cost reports audited by independent accounting firms and only some reports are audited by the federal government. Therefore, the financial assistance amount on these forms can be subject to inaccuracies or misrepresentations.[1] Financial assistance offered by hospital physicians is not always captured in HCRIS, which may underestimate financial assistance spending by hospitals with a salaried-physician model. Other research has found that some hospitals may include spending on "uninsured discounts" as financial assistance in HCRIS (discounts the hospital gives to patients who aren't eligible for Medicaid or financial assistance).[2] Some hospitals may not report their revenue from Medicaid Managed Care programs to CMS; for these hospitals, their share of Medicaid revenue will be underestimated.

The national data we have only allows us to measure the amount of spending on community benefits, not the impact that spending had on community health. While we have focused on a few categories of community benefits we have deemed most meaningful to the health of communities, we lack data on whether the spending by top-ranking hospitals is directed towards community health priorities identified in the Community Health Needs Assessment, which every nonprofit hospital is required by the ACA to conduct. We hope our research will facilitate efforts to increase transparency around hospital community benefit spending and permit local citizens, officials, and organizations to hold their hospitals accountable to their social mission to improve community health.

Inclusivity

Inclusivity is a novel metric we developed and first published in 2020 to measure the degree to which a hospital's patient population reflects the demographics of its community area.

We used inpatient admissions recorded in the Medicare Provider Analysis and Review (MEDPAR) in 2022 which allows the use of both Fee-For-Service and Medicare Advantage claims. We included distinct counts of beneficiaries at a hospital by the patient's mailing zip code. We map each zip code to a Zip Code Tabulation Area (ZCTA) using a 2020 crosswalk. We exclude patients with a zip code that was a "post office or large volume customer" to remove people using PO boxes.

---

[1] Bai G, Zare H, Eisenberg MD, Polsky D, Anderson GF. Analysis Suggests Government And Nonprofit Hospitals' Charity Care Is Not Aligned With Their Favorable Tax Treatment: Study examines government and nonprofit hospital charity care expenses compared to charity care obligations arising from the organizations' favorable tax treatment. Health Affairs. 2021 Apr 1;40(4):629-36.

[2] Bannow, T. HCA reports almost $1 billion more in charity care to Medicare than to its shareholders, drawing more taxpayer money. StatNews. 2024 Apr

Hospitals are excluded from all inclusivity results if they have 5% or more of their patients in zip codes which did not map to a ZCTA, or if their remaining patients in ZCTAs were fewer than 50.

We defined the community area by using the ZCTA of the hospital's Medicare patient population. We sorted ZCTAs by the number of patients in each (the horizontal axis in Figure 2), and then selected the ZCTA at the turning point of the cumulative patient count (the vertical axis in Figure 2). We then select the radius of the community area as the maximum distance across all ZCTAs before this cut off, after excluding any outlier distances across these ZCTAs defined by the median distance plus three times the median absolute deviation.

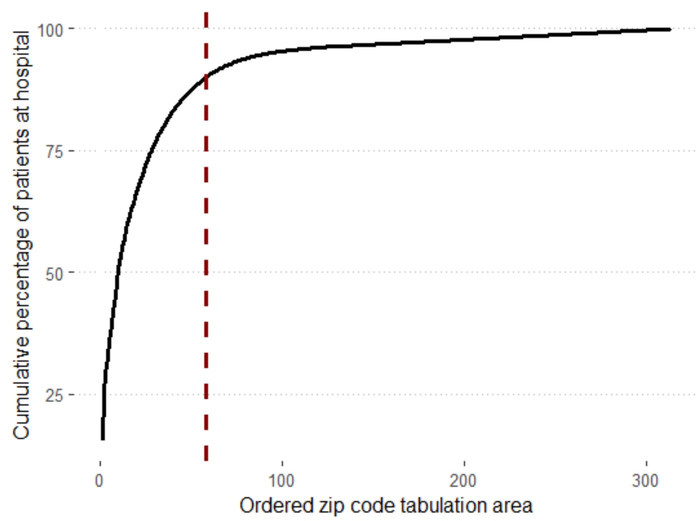

Figure 3. *First step in defining the catchment area. ZCTAs are ordered by the number of patients on the horizontal axis. The red line is the zip code 'turning point', shown here for illustration purposes only.*

We used this ZCTA with the greatest distance from the hospital as the basis for finding the maximum driving travel time from the hospital. We used OpenStreetMap (© OpenStreetMap contributors, data available under the Open Database License[3]) and the Open Source Routing Machine[4] to find this travel time as well as the area around the hospital with a boundary of this travel time. The community area is the intersection between the area defined by the geodesic distance radius and this travel time area. This excludes any areas that may be within the radius distance, but take longer to travel to than other ZCTAs from the hospital.

[3] https://www.openstreetmap.org/copyright

[4] http://project-osrm.org/

To calculate the demographic estimates for the community area population we used the 2022 five-year estimates from the US Census Bureau's American Community Survey data for people over the age of 65 on race, income, and education levels within the community area ZCTAs. We calculated each score using the total population counts and the levels of income and education and proportions of race for each ZCTA. We weighted the demographic contributions of ZCTAs to the community area by distance to the hospital, and the ZCTA area within the community area (for ZCTAs that overlapped with the travel time boundary).

We created the hospital score by using the ZCTA demographic data of the patients' ZCTAs. We then compared the community area score to the hospital score: a ratio for income and education levels, and a score summarizing the differences between the racial group populations.

We combined these three values for the overall Inclusivity score (equally weighted). There are some hospitals which have a community area with low racial and ethnic diversity: we do not include a racial inclusivity score for hospitals with a community area where the probability that two randomly selected persons being the same race is greater than 95%. These hospitals only receive an income and education inclusivity score.

Several adjustments were made for hospitals with a single Medicare provider ID with multiple campuses. These were identified using the CDC's facility identifier data set, released in 2021 (and updated continuously since). We essentially use the same steps as we do for a single campus to find the set of ZCTAs that are within a certain threshold of patient counts. Instead of the distance to the single facility address, however, we use the minimum distance of each ZCTA to any campus location. From the subset of patient ZCTAs matched to each facility, we select the maximum distance for each facility within the defined cut-off (similar to Figure 2 approach). Using the distance calculated above, for each facility we find the travel time to this point. We then draw an area around each facility based on this travel time. The final community area for the hospital is the union of all of its facility areas.

Star rating

We assign stars for the income, education and racial inclusivity measures based on the Z- score and using cut-offs based on a normal distribution. Hospitals with a Z- score greater than 1.5 receive 5 stars, while hospitals with a score less than -1.5 receive 1 star. Since the income, education and racial inclusivity scores are not normally distributed, this results in only a small number of hospitals receiving 1 and 5 stars (the more extreme results only).

Inclusivity grades

We calculate the inclusivity grade based on the sum of the adjusted scores for income, education and racial inclusivity. We transform these scores using min-max transformation (so they are on similar scales), with an adjustment such that a 'zero' score is fixed. To receive an A on inclusivity, hospitals must have a summed score greater than zero and have a minimum of two-stars on the income, education and racial inclusivity scores. We designate hospital grades B to D by splitting the negative range of scores into three equal bins.

Inclusivity limitations

Our method is based on zip codes, and assumes that people within a zip code are equally likely to visit one hospital within a community area. For example, if a zip code had an 80% population of low-income earners and 20% high-income earners, we assume that patients going to the hospital from this zip code should match this ratio. We would not be able to observe if all of the patients going to the hospital from this zip code were actually high-income earners (that is, the 80% population of low-income earners was completely excluded), and we would give the hospital a better income score than if we had actual income data for hospital patients. Conversely, we would not be able to know if the 20% of high-income earners was completely absent from the hospital's patient population. Such data is not available.

Our community area weighting for ZCTA demographics is based on an assumed relationship between the distance to the hospital and the likelihood of a person going to this hospital. This relationship may not be consistent across all hospitals and areas.

Our inclusivity score rewards hospitals that effectively "over-serve" communities with lower average income and education attainment and higher minority populations. That's by design. Hospitals whose community area demographics and patient demographics are reasonably aligned receive a mid-range score in the percentile ranking of the inclusivity scores. For example, a hospital may be in a community area that is quite wealthy and their entire patient demographics might reflect this demographic fact. In such a situation it will be difficult to improve upon a middling score. We did not include a racial inclusivity score for hospitals where most people in the community area were the same race (defined as the probability of selecting two persons that are the same race being greater than 95%).

For full details, please see our paper on inclusivity at MedRxiv.[5]

Value

Value includes cost efficiency and avoiding overuse measures (weighted 60 and 40% respectively).

Avoiding overuse

This component includes rates of overuse of 12 low-value medical services (see Table 2 below). For Acute Care Hospitals, we rated hospitals on all 12 services. This year, we created a separate ranking for Critical Access Hospitals and rated them on 2 of the services, carotid artery imaging for fainting and head imaging for fainting. We selected these since they have a high volume of low-value (numerator) instances.

Table 2. Overuse definitions for 12 low-value services

Name of low-value service	Description of service	When is it overuse?	How is overuse measured?
Arthroscopic knee surgery	Surgery to remove damaged cartilage or bone in the knee using an arthroscope (tiny camera)	Overuse when it's for patients with osteoarthritis or for "runner's knee" (damaged cartilage). Excluding patients with meniscal tears.	Overuse measured as the proportion of arthroscopic knee surgery that met our criteria for overuse.
Carotid artery imaging for fainting	A test to screen for carotid (neck) artery disease. Includes CT, Magnetic resonance angiography, and duplex ultrasound	Overuse for patients where syncope (fainting) is the primary diagnosis on the claim and no history of syncope in the past two years. Exclusions for stroke/TIA, retinal vascular occlusion/ischemia, nervous and musculoskeletal symptoms.	Measured as the proportion of patients who came to the hospital with fainting but no other symptoms of serious disease and received carotid artery imaging.

[5] Saini V, Chalmers K. Segregated Patterns of Racial and Socioeconomic Inclusivity of Access to Hospital Care Among the Medicare Population. medRxiv. 2021 Jan 1;2021.05.24.21257551. https://www.medrxiv.org/content/10.1101/2021.05.24.21257551v1

Carotid endarterectomy	Procedure to remove plaque buildup from a carotid (neck) artery in a patient to prevent stroke	Overuse when performed on female patients without stroke symptoms or history of stroke.	Measured as the proportion of carotid endarterectomies that met our criteria for overuse, out of all the CEAs performed.
Colonoscopy screening	Examination of the entire colon from the rectum to the cecum	Overuse for patients 86 years or older or if screening colonoscopy is repeated within 9 years	Measured as the proportion of colonoscopies that met our criteria for overuse, out of all colonoscopies for patients 75+ without gastrointestinal symptoms, colorectal neoplasia within 12 months, and no colectomy, colorectal cancer, colon polyps, inflammatory bowel disease, family history of colorectal cancer recorded within 10 years.
Coronary artery stenting	Procedure to place a stent or balloon in a coronary artery	Overuse when performed on patients with stable heart disease (not having a heart attack or unstable angina). Excluding patients with current and past diagnosis of unstable angina as well as patients having a heart attack.	Measured as the proportion of coronary stents that met criteria for overuse, out of all the stents placed.
EEG for fainting	A test of the electrical activity of the brain	Overuse for patients where syncope (fainting) is the primary diagnosis on the claim and no history of syncope in the past two years.	Measured as the proportion of patients who fainted but no other symptoms of serious disease who received an EEG.
EEG for headache	A test of the electrical activity of the brain	Overuse for patients with headache as the primary diagnosis on the claim and no history of headache in the past two years. Also exclusions for epilepsy and recurrent seizures, convulsions, and abnormal involuntary movements.	Measured as the proportion of patients who came to the hospital with headache but no other symptoms of serious disease who received an EEG.

Head imaging for fainting	A CT scan or MRI of the head	Overuse for patients where syncope (fainting) is the primary diagnosis on the claim and no history of syncope in the past two years. Exclusions for epilepsy or convulsions, cerebrovascular diseases including stroke/TIA and subarachnoid hemorrhage, head or face trauma, altered mental status, nervous and musculoskeletal system symptoms, including gait abnormality, meningismus, disturbed skin sensation, speech deficits, personal history of stroke/TIA.	Measured as the proportion of patients who came to the hospital with fainting but no other symptoms of serious disease and received an MRI or CT scan.
Inferior vena cava filter (IVC)	Procedure to place a filter (a medical device) in the large vein in the abdomen to prevent blood clots from moving to the lungs	Overuse for all patients	Measured as the number of times an IVC filter was overused, as proportion of total hospital volume.
Renal artery stenting	Procedure to place a stent or balloon in the renal (kidney) artery in a patient with high blood pressure or cholesterol (plaque) buildup in the artery	Overuse when done for hypertension or plaque buildup. Excluding patients that had diagnosis of fibromuscular dysplasia of renal artery (abnormal twisting of the blood vessels)	Measured as the number of times a renal artery stent or balloon was overused, as a proportion of total hospital volume.
Spinal fusion/ laminectomy	Procedure to remove part of a spinal vertebra or fuse vertebrae together	Overuse for patients with low-back pain, excluding patients with radicular symptoms, herniated disc, radicular pain, scoliosis, radiculopathy, sciatica, trauma, discitis, spondylosis, or myelopathy. Additionally, exclude cases with both stenosis with neural	Measured as the proportion of spinal fusion or laminectomy procedures that met our criteria for overuse, out of all the spinal fusion / laminectomies performed.

		claudication and spondylolisthesis when spinal fusion is performed; exclude cases with stenosis with neural claudication when laminectomy alone is performed (without spinal fusion).	
Vertebroplasty	Procedure to inject cement into the vertebrae to relieve pain from spinal fractures	Overuse for patients with spinal fractures caused by osteoporosis. Excluding claims with bone cancers, myeloma, or hemangioma.	Measured as the proportion of patients that came in with spinal fractures caused by osteoporosis who received vertebroplasty.

We chose these services from the overuse literature. Renal stenting and inferior vena cava filters have been shown in high-quality clinical trials to be ineffective and are nearly always considered overuse. The remaining interventions are considered overuse when prescribed to patients with certain diagnoses or conditions. For example, a patient with stable angina is considered an inappropriate candidate for a cardiac stent and use of a stent in this case is considered low value or overuse. Similarly, a patient with syncope does not require an EEG.

We used 100% Medicare claims FFS inpatient and outpatient claims from January 2020 to December 2022 as well as Medicare Advantage inpatient and outpatient encounter data from January 2019 to December 2021 to count instances when these services were used. In addition, we used carrier claims from both FFS and Medicare Advantage for assessing prior history of diagnoses. Hospitals without capacity to perform a service, as reflected in their claim history, were excluded from the rating for that particular service.

For the 12 services, we calculated the overuse rate as the instance of low value services (numerator) among a defined cohort (denominator). For those services which are low-value in most cases (renal stenting and inferior vena cava filter), we used the total patient volume as the denominator. For the remaining services where there was some benefit in certain circumstances, we used a service-specific (for the procedures) or diagnosis-specific denominator (for tests, imaging, and vertebroplasty).

We used an empirical Bayes reliability adjustment on the rates so hospitals with smaller denominator volumes had their rates shifted towards the overall mean. We rank hospitals on each individual low value service based on the reliability adjusted scores.

Star rating

Avoiding overuse stars for each low-value service are based on the reliability adjusted overuse rates. We standardized the reliability adjusted overuse rates using a minimum-maximum transformation (so they were between zero and one). For carotid endarterectomy, colonoscopy screening, coronary artery stenting, and head imaging for fainting, we then Z-transformed and converted the sign so hospitals with the highest Z-score have the lowest overuse. We used cutoffs of 1.5, 0.5, -0.5, and -1.5. Hospitals with Z-score 1.5 or higher had the best performance (lowest overuse) and received 5-stars. For arthroscopic knee surgery, carotid imaging for fainting, EEG for fainting, EEG for headache, IVC, renal artery stenting, spinal fusion/ laminectomy, and vertebroplasty, the data were not normally distributed. For these services, hospitals within the lowest fifth of the reliability adjusted overuse rate ranges received 5-stars. Hospitals that had zero numerator counts, regardless of their reliability adjusted rate also received 5-stars. For low-value services measured using a service-specific denominator (carotid endarterectomy, coronary artery stenting, spinal fusion/laminectomy, and knee arthroscopy), we calculated the numerator decrease needed to improve the star rating. If this amount was small based on the hospital's volume of similar services, their star rating was upgraded to account for chance and specificity of the overuse rating.

Avoiding overuse grade

We combined the standardized overuse rates for each low-value service into a single overuse composite score. We used capacity criteria to determine whether hospitals were eligible to receive the composite score. Acute Care Hospitals with capacity to perform at least 6 services received a composite score. Critical Access Hospitals with capacity to perform both carotid artery imaging and head imaging received a composite score.

For Acute Care Hospitals, we calculated the overuse score as the weighted sum of the 12 standardized values for each low-value service. The weights were determined by the count of total low-value services among Acute Care Hospitals nationally in our data set. If a hospital had no capacity for a service, we redistributed this weight to their other service results. For Critical Access Hospitals, we calculated the overuse score as the weighted sum of the standardized values for carotid imaging for fainting and head imaging for fainting. The weights were determined by the count of low-value services among Critical Access Hospitals nationally in our data set.

To assign grades for Acute Care Hospitals, we Winsorized the composite overuse score (and replaced values more extreme than the 5% or 95% quantile) and then converted these to Z-scores with high Z-score representing good performance (low overuse composite score). We selected grade cutoff values using a standard normal distribution for the quantiles 15%, 35%, 75%. Hospitals with a Z- score in this top quantile received an A, those in the next quantile received a B, and so on. We used the same method to assign grades to Critical Access Hospitals.

Avoiding overuse limitations

We used low-value services well-established in the literature, but the true definition of overuse almost always depends on the clinical circumstances, which are not necessarily captured in claims data. Furthermore, errors of coding and reporting by providers could have resulted in errors in our estimates. Particularly for low-volume hospitals, these estimates may be subject to sampling error resulting in changing rates from year to year. Our goal was to estimate rates at the level of the hospital, not of an individual practitioner.

We tried to avoid rewarding hospitals for avoiding overuse when they do not in fact have the capacity to perform such a service. The capacity assessment we developed as an indicator is defined using lists of procedure codes that are much broader than the inappropriate ones. However, as with the measurement of overuse itself, our capacity assessment is claims-based and subject to errors at very low volumes. It is possible that some hospitals have been included and rewarded when they do not, in fact, have true capacity to perform the service.

Cost efficiency

The cost efficiency component measures the clinical outcomes hospitals achieve over the cost of care on the Medicare FFS population. This metric encompasses two details: 30-day mortality and cost, and 90-day mortality and cost.

Calculating 30-day and 90-day episode costs

We measured 30- and 90-day total, standardized Medicare FFS payments for patients hospitalized in 2020 to 2022. We excluded any hospitalizations that were transfers from another hospitalization, had denied Medicare payments, if patients left against medical advice, patients were admitted with or had COVID-19, or where the primary payer was not Medicare.

For each hospitalization, we found the claim payment amount in all claims within 30 or 90 days from the admission date. These claims included: inpatient, outpatient, carrier, skilled nursing facility, home health agencies, durable medical equipment and hospice claims. We excluded any claims where Medicare denied the payment.

We prorated any claims that started but did not finish in the 30- or 90-day period after the index hospitalization. For example, if a patient had another inpatient visit starting on day 29 after their first hospitalization, and finishing on day 31, then only two-thirds of this inpatient claim payment would be included in the patient's total 30-day payment.

Medicare adjusts their payment amounts to hospitals and other providers based on various geographic factors. To account for this, we calculated standardized payments using the Virtual Research Data Center's public use files of 2019 to 2021 Hospital Referral Regions (HRR) standardized ratio tables for patients over 65. As the 2022 data was not available, we re-used 2021 HRR estimates for 2022 claims. These tables have separate values for each claim type (inpatient, outpatient, etc.). Our standardized payment amount was the hospital's HRR standardized payment value for the claim year divided by the HRR actual payment value, multiplied by the claim payment amount.

A hospitalization's 30-day and 90-day standardized payments were the total sum of the standardized payments across each claim type.

Risk-standardized payments

We risk adjusted the 30-day and 90-day standardized payments for each hospitalization using hierarchical logistic regression models. We estimated separate models for Acute Care and Critical Access Hospitals. All measurement and results were carried out separately for Acute Care and Critical Access Hospitals. The response variable in the model was the episode standardized payment per survival day where survival day was the number of days the patient survived in the 30-day or 90-day episode. Model predictions provided the risk-standardized payment per survival day with hospital effects (predicted) and without hospital effects (expected). For Acute Care Hospitals, the hospital effects included hospital volume and hospital proportion of admissions in each of 15 service divisions as described in *Clinical Outcomes*. For Critical Access Hospitals, the hospital effects did not include volume and included hospital proportion of admissions in each of the 15 service divisions.

The risk-standardized payment per survival day for each hospitalization was multiplied by the number of survival days to get the predicted and expected episode cost for each hospitalization.

We then calculated the mean risk-standardized predicted cost (P) and expected cost (E) for each hospital. A hospital's risk-standardized payment (RSP) is the hospital's P/E ratio multiplied by the national average episode cost. We calculated 30-day and 90-day RSP for each hospital using this method.

Cost efficiency metric

Our goal for the cost efficiency metric was to reward hospitals with low mortality rates and low costs, and give the lowest scores to hospitals with high mortality rates and high costs. We also decided to bias our scores to give hospitals with high costs and low mortality a higher score than hospitals with low costs and high mortality. This is because we believe that if there is a trade-off between costs and mortality, we should favor better mortality rates compared to lower costs.

In order to operationalize this metric, we mapped the respective 30-day and 90-day risk standardized mortality rates developed on the FFS population and risk-standardized payments on a cartesian plane. We transformed the mortality rates and payments using a min-max transformation, so the range of values of the two variables were equal.

We then created a point on this plane that represented the ideal (most cost efficient) hospital, with the lowest mortality rate and payment value. We then used vectors to calculate the distance and angles between every single hospital in the data set and this ideal hospital using polar coordinates.

We then multiplied these two values, the distance and the angle, between a hospital's results and the best, theoretical hospital to generate our cost efficiency metric. We included the angle in the metric to ensure that if there were two hospitals with an equal distance from the ideal hospital on the payment-mortality plane, hospitals with lower mortality would receive a better score than hospitals with higher mortality. Larger angles reflected higher mortality rates, while smaller angles higher payments.

Star rating

To assign stars for Acute Care Hospitals, we Winsorized the risk-standardized payments (and replaced values more extreme than the 5% or 95% quantile) and then converted these to Z-scores with high scores representing good performance (small distance and angle from best theoretical hospital). We calculated the range of these values and divided this into five regions. Hospitals in the highest region were assigned 5 stars (the best performers), while hospitals in the lowest region were assigned 1 star. We used the same method to assign stars to Critical Access Hospitals.

Cost efficiency grade

For Acute Care Hospitals, the composite cost efficiency score was calculated as the average of the 30 and 90-day cost efficiency scores. We Z-transformed the composite score. Hospitals that did better than the mode result received an A. We split the range below the mode into 3 regions to assign grades B, C, and D. We used the same method to assign grades to Critical Access Hospitals.

Cost efficiency limitations

We included Medicare payments and not payments from other payers, such as patient contributions or other insurers. This means we might be underestimating the true costs of some patient episodes.

Since our data includes Medicare beneficiaries and standardized costs, we cannot examine price variation as part of our metric. A hospital might be highly cost efficient because they have low readmissions and avoid unnecessary care, but they may charge high prices to non-Medicare patients.

Our cost standardization method is specific to the HRR. Some HRRs are quite large, and there might be more specific adjustments made within these regions not accounted for in our standardization approach.

While mortality and cost are adjusted for underlying patient risk, it is likely that some environmental and social factors that impact patient outcomes may not be accounted for in our risk adjustment. That means hospitals caring for the poorest and sickest patients may appear to do worse on mortality and cost (see *clinical outcomes limitations* for more).

Outcomes

Our outcomes composite is created from three measures (clinical outcomes, patient safety, and patient satisfaction). Acute Care Hospitals had scores weighted 62.5%, 25% and 12.5%, respectively. The majority of Critical Access Hospitals do not have enough data for a patient safety result, therefore we do not report patient safety for these hospitals. Their outcomes result is calculated based on clinical outcomes (weighted 83.3%) and patient satisfaction (weighted 16.7%).

Clinical outcomes

Clinical outcomes comprises six sub-measures: in-hospital, 30-day and 90-day mortality; 7-day and 30-day readmission; and unplanned hospital visits post outpatient surgery within 7 days. All of the measurement and results for Acute Care and Critical Access Hospitals was performed separately on the split populations. We do not report unplanned hospital visits post outpatient surgery for Critical Access Hospitals, as the majority of these hospitals do not have high outpatient volumes.

We chose the mortality endpoints to cover measurements in CMS' inpatient quality reporting programs as well as a more extended period when mortality is a function of both hospital and community. Similarly, for readmission, we wanted both a shorter interval that would better reflect inpatient care, and longer follow-up that would reflect post-hospital community support. Readmission was calculated with greater weight on risk-standardized rate of 7- relative to 30-day readmission.

Hospitalizations and readmissions were identified from the 100 percent Medicare inpatient fee-for-service claims files for years 2020 through 2022 as well as from Medicare Advantage inpatient encounter files for 2019 and 2021. We applied additional restrictions to our mortality cohort. We excluded hospitalizations that were transfers from another hospitalization, had denied Medicare payments (applied only to FFS claims), if patients left against medical advice, or where the primary payer was not Medicare (applied only to FFS claims). We excluded Covid-19 confirmed and probable admissions. We excluded psychiatric and rehab admissions as well as admissions with a diagnosis of metastatic cancer or where survival cannot be influenced. We used Medicare FFS outpatient and carrier claims as well as Medicare Advantage outpatient and carrier encounter files to include a 1-year lookback so that risk adjustment was based on patient conditions from both the hospitalization and the prior year. Beneficiary characteristics and death date were obtained from the Medicare Beneficiary Summary file.

For unplanned hospital visits following outpatient surgery, we identified our denominator cohort as low- and moderate-risk surgeries performed at hospital outpatient departments (HOPDs) from the 100 percent Medicare inpatient fee-for-service claims files for years 2020 through 2022 as well as from Medicare Advantage inpatient encounter files for 2019 and 2021. We used CMS's list of covered ambulatory surgery center procedures updated annually[6]. From this cohort, we defined our outcome as an unplanned hospital visit within 7-days. This includes inpatient admissions directly following surgery as well as any ED visit, observation stay, or unplanned inpatient admission following discharge from the HOPD (defined using CMS's unplanned admission algorithm).

Mortality, readmission, and unplanned hospital visit rates were risk adjusted using the Risk Stratification Index (RSI), an algorithm in the public domain that the Lown

[6] ASC Payment Rates - Addenda | CMS

Institute trained using more than 18 million patient stays from inpatient FFS. RSI has been tested on several different national and hospital-based datasets and has been shown to predict outcomes with greater discrimination compared with other publicly available risk adjustment tools.[7,8,9]

Risk-standardization

We risk adjusted the mortality, readmission, and unplanned hospital visit rates for each hospitalization using hierarchical logistic regression models. We estimated separate models for Acute Care and Critical Access Hospitals. All measurement and results including scores, rankings, stars, and grades were carried out separately for Acute Care and Critical Access Hospitals. Model predictions provided the risk-standardized mortality, readmission, or unplanned hospital visit with hospital effects (predicted) and without hospital effects (expected) for each hospitalization.

For each hospital, we then calculated the predicted (P) and expected mortality (E) based on all of its hospitalizations. A hospital's risk-standardized mortality (RSMR) is the hospital's P/E ratio multiplied by the national observed mortality rate. We calculated in-hospital, 30-day, and 90-day risk-standardized mortality using this method. Similarly, we used the P/E ratio and national observed event rate to calculate risk-standardized 7 and 30-day readmission (RSRR) as well as risk-standardized unplanned hospital visits post outpatient surgery (RSUVR).

In addition to the patient conditions in RSI, we included model effects to account for differences in hospital volume (Acute Care Hospitals only) and hospital proportion of admissions in each of 15 service divisions. We used the 15 service divisions defined in CMS Hybrid Hospital Wide Mortality Methodology along with AHRQ CCS procedure and diagnosis categories to classify each inpatient admission into one of 9 non-surgical or 6 surgical divisions.[10] At the patient level, we included model effects for dual eligibility and end stage renal disease.

---

[7] Chamoun GF, Li L, Chamoun NG, Saini V, Sessler DI. Validation and calibration of the risk stratification index. Anesthesiology. 2017 Apr;126(4):623-30.

[8] Sessler DI, Sigl JC, Manberg PJ, Kelley SD, Schubert A, Chamoun NG. Broadly applicable risk stratification system for predicting duration of hospitalization and mortality. The Journal of the American Society of Anesthesiologists. 2010 Nov 1;113(5):1026-37

[9] Chamoun GF, Li L, Chamoun NG, Saini V, Sessler DI. Comparison of an updated risk stratification index to hierarchical condition categories. Anesthesiology. 2018 Jan;128(1):109-16.

[10] https://qualitynet.cms.gov/inpatient/measures/hybrid/methodology

Star rating

To assign stars for Acute Care Hospitals, we Winsorized the risk-standardized endpoint rates (and replaced values more extreme than the 1% or 99% quantile) and then converted these to Z-scores with high scores representing good performance (low risk-standardized rate). Hospitals that did better than the mode result received 5 stars. We split the range below the mode into 4 regions to assign star levels 1 through 4. We used the same approach to assign stars to Critical Access Hospitals.

Clinical outcomes grade

For Acute Care Hospitals, the composite clinical outcomes score was calculated as the weighted sum of the 6 endpoint risk-standardized rates after Winsorization and Z-transformation with high Z-score representing good performance (low risk-standardized rate). We used weights from the table below. For Critical Access Hospitals, the composite clinical outcomes score was calculated as the weighted sum of the 5 endpoint risk-standardized rates after Winsorization and Z-transformation with high Z-score representing good performance (low risk-standardized rate). We Z-transformed the composite clinical outcomes score for Acute Care Hospitals. Hospitals that did better than the mode result received an A. We split the range below the mode into 3 regions to assign grades B, C, and D. Similarly, for Critical Access Hospitals, we Z-transformed their composite clinical outcomes score and assigned grades using the mode.

	Acute Care Hospital weighting (%)	Critical Access Hospital weighting (%)
In-hospital mortality	30	30
30-day mortality	25	25
90-day mortality	15	20
7-day readmissions	10	15
30-day readmissions	5	10
Unplanned hospital visits post outpatient surgery	15	-

Hybrid Hospital-Wide (All-Condition, All-Procedure) Risk-Standardized Mortality Measure with Electronic Health Record Extracted Risk Factors Methodology Report Version 2.0.

Clinical outcomes limitations

While our clinical outcomes metrics adjust for underlying patient risk, it is likely that some environmental and social factors that impact patient outcomes, such as the availability of healthy food, access to preventive care, pollution, and others, may not be accounted for in our risk adjustment. Patients living in neighborhoods with poor environmental and social conditions often come to the hospital with more advanced cases of a given disease, and these patients are often discharged from the hospital into situations where they are less able to get the continuing care they need. For example, a patient who leaves the hospital for an apartment on the fifth floor of a walk up with no grocery store nearby might not do as well as a patient who can hire an aide to help them recover at home. That means hospitals caring for the poorest and sickest patients may appear to do worse on patient outcomes unrelated to the quality of their care.

For more details, please see our paper in the journal *Medical Care*.[11]

Patient safety

For patient safety we used well-established indicators provided by CMS on its Care Compare website for hospitalizations, such as rates of pressure ulcers, accidental punctures, and central intravenous line infections. We included the CMS composite measure (PSI-90), which comprises 10 separate indicators of patient safety, as well as 5 hospital acquired infection (HAI) measures. We included a reliability adjustment for the HAI measures using the reported numerator and denominator counts from Care Compare. For more detail and a listing of the measures used, please see the CMS webpage on hospital acquired conditions.[12]

Star rating

To assign stars, we Winsorized the adjusted CMS scores (and replaced values more extreme than the 5% or 95% quantile) and then converted these to Z-scores. We calculated the range of these values and divided this into five regions. Hospitals in the

---

[11] Saini V, Gopinath V. [Application of the Risk Stratification Index to Multilevel Models of All-condition 30-Day Mortality in Hospitalized Populations Over the Age of 65.] Medical Care. 2021 Sep;59(9):836.

[12] https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Downloads/HAC-Reduction-Program-Fact-Sheet.pdf

lowest region were assigned 5 stars (the best performers), while hospitals in the highest region were assigned 1 star.

Patient safety grade

Very few Critical Access Hospitals had these measures reported on CMS Care Compare, and therefore these hospitals do not have a patient safety grade. For a patient safety overall score, hospitals had to have had at least three of the PSI-90 or HAI results. The overall patient safety score was the averaged result of a hospital's patient safety scores (equally weighted), after Winsorization and Z-transformation. To determine the grade cut-offs, we selected values using a standard normal distribution for the quantiles 15%, 35%, 75%. Hospitals with a Z-transformed score in this top quantile received an A, those in the next quantile received a B, and so on.

Patient satisfaction

CMS Care Compare was also the source for our patient satisfaction ranking. CMS uses the annual Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey to give a rating of patient experience across 10 factors. We took the average of the 10 linear mean scores of these factors published on the 2024 Care Compare site (covering submitted data from 2022 to 2023), which also reports a percentage of patients with each summary response. The linear mean scores for each component are adjusted for patient-mix and survey-mode by CMS.

We chose to include hospitals with between 50 and 100 responses after data analysis indicated that imputation of these scores would be reasonable to account for CMS's mean calculations and adjustment. We calculated scores for these hospitals by extrapolating to the nearest median score of hospitals with similar survey responses. For more detail and a listing of the 10 measures used, please see the CMS webpage on patient experience.[13]

Star rating

To assign stars, we Winsorized the adjusted CMS scores (and replaced values more extreme than the 5% or 95% quantile) and then converted these to Z-scores. We calculated the range of these values and divided this into five regions. Hospitals in the highest region were assigned 5 stars (the best performers), while hospitals in the lowest region were assigned 1 star.

---

[13] https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalHCAHPS

Patient satisfaction grade

Like all hospital measures reported in 2024 and onwards, we compared Critical Access Hospitals and Acute Care Hospitals against their peers. The overall patient satisfaction score was the averaged score of the hospital's Linear Mean Scores, after Winsorization and Z-transformation. To determine the grade cut-offs, we selected values using a standard normal distribution for the quantiles 15%, 35%, 75%. Hospitals with a Z-transformed score in this top quantile received an A, those in the next quantile received a B, and so on.

Hospital grades and ranks for equity, value and outcomes

To roll up these components for the Tier 2 ranking, we first take the *grade point average* (GPA) of the component grades within each category. Similar to a college student's GPA value, we assign a 4 to an A, 3 to a B, 2 to a C, and 1 to a D. The GPA of the category is the weighted average of these values, with weights specific to each category and provided in the previous sections.

Hospitals with a weighted GPA of 3.3 or higher get an A for Tier 2; a GPA of 2.7 or higher receives a B; a GPA of 1.8 receives a C; while anything less than this receives a D overall.

To determine the hospital ranks, we first sort by each overall grade and then their grade within each Tier 3 metric (ordered by the size of the metric's weighting in the composite score). Final ranks within these grades are based on the composite score. This process ensures that hospitals with top Tier 2 results all receive an *A* grade on the Tier 3 measures, and that a poor performance in one Tier 3 measure is not compensated with an exceptional score in another measure.

Putting it together: Grade and rank for social responsibility

The Lown social responsibility GPA is the weighted sum of a hospital's grades in the three categories: Equity, value and outcomes. This year, we changed the weights of equity (40%; previously 30%) and outcomes (30%; previously 40%) in order to have hospitals' equity scores more predictive of their final social responsibility result, which was our original intention when we developed the ranking.

Value and outcomes are two highly correlated measures, due to the use of mortality rates in both clinical outcomes and cost efficiency, meaning hospitals which perform well on outcomes tend to perform well on value. This meant hospitals performing well on outcomes, but not equity, still performed well on social responsibility because of the likely extra boost in their scores from their value result. Increasing the equity weight to 40% has attenuated this effect.

In addition to the above change, we have also implemented a rule such that a hospital cannot receive an A grade on social responsibility if they have a C or lower on either equity, value or outcomes.

The social responsibility rankings are determined by first sorting the hospitals by their Tier 2 grades, then their GPA, and then the weighted sum of their Tier 2 scores.

We dropped hospitals from the social responsibility ranking if they did not have an equity, value and outcomes result (116 hospitals). The results for these hospitals on other metrics are still visible.

Hospital Systems

We looked at hospital systems as a secondary unit of analysis. We used our amended version of Agency for Healthcare Research and Quality's (AHRQ) 2022 compendium of US healthcare systems as a baseline for our hospital system definition. AHRQ defines a health system as "at least one hospital and at least one group of physicians providing comprehensive care, and who are connected with each other and with the hospital through common ownership or joint management." We again made necessary changes to the AHRQ 2022 definitions to update them to 2024 status, such as accounting for system mergers and closures. Our goal was to see how these systems compared against each other within the various tiers of the Lown Index. System scores were calculated using only those hospitals in the system that were included in this year's Lown Index hospitals list. A system may have additional hospitals that were not included in the ranking.

For all metrics except patient satisfaction and cost efficiency, we consolidated hospital component scores to the system level by calculating an average of each hospital metric across the system of hospitals weighted by patient volume from 2019 to 2022. To calculate patient satisfaction scores at the system level, we computed a hospital average weighted by the number of completed surveys recorded within the 2024 Care Compare dataset. We calculated the system scores for cost efficiency by computing a hospital average weighted by 2019 to 2022 patient volume from Fee for Service (FFS) claims.

In order to integrate both Critical Access and Acute Care hospitals into one system score, we deferred to using the Acute Care hospital approach to calculating scores if the system contained at least one Acute Care general hospital. This is due to the fact that in the majority of cases, Acute Care general hospitals see a majority of annual patient volume for systems compared to critical access hospitals. There was one system that only contained Critical Access Hospitals and thus, its scores were calculated using the Critical Access Hospital approaches outlined in the previous sections.

Hospital systems limitations

In almost all cases, we used the 2022 AHRQ definition of health system.

We used a weighted average across all hospitals within a system to calculate systems scores. Hospitals with higher patient volume are weighted higher within our systems rollup. We could have, alternatively, summed the numerator and denominators for all metrics within each system and calculated a system score that arguably could have reflected the culture of a system. However, we found that this approach meant that the system scores were most dependent on the hospitals with the largest patient volumes, and results from smaller volume hospitals made little impact on the system results. Our weighted average approach combines the results of individual hospitals, and therefore is likely a closer reflection of the combination of individual hospital's cultures opposed to the system culture as a whole. Finally, when ranking systems by state, the system will appear in that state's rankings if a system includes at least one hospital in that state.

COVID-19 Burden

A note on not updating the COVID-19 burden data: We have not updated the COVID-19 burden data published on the Lown Institute Hospitals Index page from its original 2022 version. Our clinical outcome measures in 2022 covered 2018 to 2020, and this year our clinical outcome measures cover 2020 to 2022. The pandemic period of 2020 is still included in this time period. As COVID-19 treatments and mortality rates changed through 2021, the COVID-19 burden for a hospital in 2021 may be different than in 2020. We therefore decided to leave our original COVID-19 burden data the same as our 2022 version. The methodology outlined hereafter regarding Covid-19 burden is the same as last year's methodology.

We did not incorporate COVID-19 specific case counts or hospitalizations into the rankings. Hospital staff across the country faced an unprecedented challenge in caring for patients these past few years through the COVID-19 pandemic. We made the decision to not compare or adjust these burdens across hospitals within the index. We have, however, shared this information for individual hospitals so readers can contextualize a hospital's Index results with their level of COVID-19 hospitalizations through this time.

We used the COVID-19 facility level data from the Centers for Disease Control and Prevention's COVID-19 Reported Patient Impact and Hospital Capacity by Facility data set.[14] Hospitals reported the 7-day average of the number of patients currently hospitalized in an inpatient bed who have suspected or confirmed COVID-19. The CDC started collecting this information from most hospitals on 31st July, 2020.

We used this data and the Medicare FFS data from the CMS Chronic Conditions Warehouse to estimate the total number of patients with COVID-19 prior to the CDC's starting collection date. Patients were identified in the data set if they had a principal or admitting diagnosis of ICD-10 codes B97.29 or U07.1. We fitted a linear regression at the hospital level to the CDC data between July 31, 2020 to November 31, 2021 and the Medicare FFS COVID-19 patients. We did not include an intercept term in these regressions. We then found the predicted number of total patients based on the Medicare FFS patients with COVID-19 through March 1, 2020 to the CDC collection start date (31 July 2020 for most hospitals).

We found the number of inpatient beds (including intensive care units) from a hospital's 2019 Hospital Cost Report. Available hospital beds can fluctuate due to staffing changes. We decided to use a fixed number of hospital beds reported prior to the pandemic as a benchmark.

We calculated two metrics to display on the website and dashboard. The first was the *Maximum COVID-19 beds*. We found the three week rolling average of the number of COVID-19 hospitalized patients in the hospital, and then found the maximum of this value in the hospital up to March 1, 2021. We reported this value as a percentage of the hospital's 2019 inpatient bed count.

The second value we reported was *COVID-19 burden.* This was the number of weeks where the percentage of COVID-19 hospitalized patients were 10% or more of a hospital's 2019 inpatient bed count.

Limitations to COVID-19 burden

There is some missing information in the CDC facility data set. Some hospitals do not have reported values for every week. We have included the number of weeks on the website where a hospital has reported this information along with the COVID-19 burden value.

The relationship between the number of Medicare FFS COVID-19 patients and the total hospitalized COVID-19 patients varied over the course of the pandemic. Hospitalized patients at the start of the pandemic may have been more likely to be older and in the

[14] https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/anag-cw7u

Medicare population than they were later on, when vaccination rates in older groups were high. We did not include CDC data post-November 2021 due to the impact of the COVID-19 Omicron wave on hospitalization rates, which had a very clear difference to previously reported hospitalizations. We also did not have data on patients enrolled in Medicare Advantage, and the proportion of Medicare Advantage versus Fee-for-service will vary between hospitals.

For these reasons we caution against relying on these estimates as a true value of hospital burden during the initial COVID-19 wave in the US.