BMJ Open Quality

Hospital quality reporting in the pandemic era: to what extent did hospitals' COVID-19 census burdens impact 30-day mortality among non-COVID Medicare beneficiaries?

Benjamin D Pollock ,[1] Sean C Dowdy[1,2]

To cite: Pollock BD, Dowdy SC. Hospital quality reporting in the pandemic era: to what extent did hospitals' COVID-19 census burdens impact 30-day mortality among non-COVID Medicare beneficiaries? *BMJ Open Quality* 2023;**12**:e002269. doi:10.1136/bmjoq-2023-002269

Received 17 January 2023
Accepted 10 March 2023

[1]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, Minnesota, USA
[2]Department of Quality, Experience, and Affordability, Mayo Clinic, Rochester, Minnesota, USA

Correspondence to
Dr Benjamin D Pollock; Pollock.Benjamin@Mayo.Edu

ABSTRACT

Objectives Highly visible hospital quality reporting stakeholders in the USA such as the US News & World Report (USNWR) and the Centers for Medicare & Medicaid Services (CMS) play an important health systems role via their transparent public reporting of hospital outcomes and performance. However, during the pandemic, many such quality measurement stakeholders and pay-for-performance programmes in the USA and Europe have eschewed the traditional risk adjustment paradigm, instead choosing to pre-emptively exclude months or years of pandemic era performance data due largely to hospitals' perceived COVID-19 burdens. These data exclusions may lead patients to draw misleading conclusions about where to seek care, while also masking genuine improvements or deteriorations in hospital quality that may have occurred during the pandemic. Here, we assessed to what extent hospitals' COVID-19 burdens (proportion of hospitalised patients with COVID-19) were associated with their non-COVID 30-day mortality rates from March through November 2020 to inform whether inclusion of pandemic-era data may still be appropriate.

Design This was a retrospective cohort study using the 100% CMS Inpatient Standard Analytic File and Master Beneficiary Summary File to include all US Medicare inpatient encounters with admission dates from 1 April 2020 through 30 November 2020, excluding COVID-19 encounters. Using linear regression, we modelled the association between hospitals' COVID-19 proportions and observed/expected (O/E) ratios, testing whether the relationship was non-linear. We calculated alternative hospital O/E ratios after selective pandemic data exclusions mirroring the USNWR data exclusion methodology.

Setting and participants We analysed 4 182 226 consecutive Medicare inpatient encounters from across 2601 US hospitals.

Results The association between hospital COVID-19 proportion and non-COVID O/E 30-day mortality was statistically significant ($p<0.0001$), but weakly correlated ($r^2=0.06$). The median (IQR) pairwise relative difference in hospital O/E ratios comparing the alternative analysis with the original analysis was +3.7% (−2.5%, +6.7%), with 1908/2571 (74.2%) of hospitals having relative differences within ±10%.

Conclusions For non-COVID patient outcomes such as mortality, evidence-based inclusion of pandemic-era data is methodologically plausible and must be explored rather than exclusion of months or years of relevant patient outcomes data.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Many hospital quality reporting stakeholders in the USA and Europe have excluded pandemic-era data from rankings, ratings and pay-for-performance programmes, indicating they do not intend to hold hospitals accountable for pandemic-era outcomes.

WHAT THIS STUDY ADDS

⇒ US hospitals' COVID-19 patient burdens explained only a small amount of the variation in 30-day mortality among elderly non-COVID patients during the early months of the pandemic, and hospitals' risk-adjusted performances did not vary meaningfully on average when high COVID burden months were excluded from the analysis.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ For non-COVID patient outcomes, evidence-based inclusion of pandemic-era data in hospital quality reporting is methodologically plausible and must be explored more rigorously rather than exclusion of months or years of patient outcomes data which may mask deteriorations in hospital quality unrelated to the pandemic.

INTRODUCTION

The 2022 edition of the US News & World Report's (USNWR) Best Hospitals rankings used hospital outcomes data through the year 2020,[1] representing the first time that quality outcomes reported by USNWR have overlapped with the era of the COVID-19 pandemic. USNWR methodologists were confronted with a decision as to what extent hospitals should be held accountable for non-COVID patient outcomes during this period

of unprecedented stress to the healthcare system. Other quality reporting stakeholders faced a similar quandary, including the Centers for Medicare & Medicaid Services (CMS) which chose to exclude all data from 2020 in the Hospital-Acquired Condition Reduction Program penalty[2] and to exclude 6 months of outcomes data from the Overall Hospital Star Rating.[3] The Leapfrog Hospital Safety Grade has used a similar pandemic-era data period exclusion for several measures.[4] In the UK, the National Health Service suspended its pay-for-performance (P4P) programme, the Quality and Outcomes Framework, for the duration of the pandemic to allow hospitals to focus on COVID-19 care without the additional stress of meeting performance targets,[5] while several other nations across Europe saw changes in reimbursement for non-COVID care, though not necessarily based on performance.[6] Ostensibly, these decisions were made with the best of intentions to prevent potentially unfair comparisons in outcomes due to pandemic-era stresses on hospital operations such as high COVID-19 census, which was the key pandemic-era concern prioritised for prediction by many institutions, including ours, throughout the pandemic.[7]

Critically missing from these decisions, however, have been publicly available analyses to support or refute the pre-emptive methodological decisions to implement months-long or years-long data exclusions for non-COVID patient outcomes. On the contrary, the established paradigm in hospital quality reporting has been to use risk adjustment to reduce the substantial differences between hospitals' patient populations, providers and processes and allow for creation of meaningful performance benchmarks. The unprecedented nature of the pandemic appears to have steered quality measurement stakeholders away from this well-founded, data-driven approach, seemingly without any examination or presentation of evidence to support this paradigm shift. Inappropriate data exclusions could potentially limit the ability to accurately assess hospital quality, safety and experience, potentially mislead patients making decisions about where to seek care and mask improvements or deteriorations in quality unrelated to the pandemic that may have occurred. Here, we mirrored one of the most well-described and specific exclusion methodologies (put forth by USNWR) to assess to what extent hospitals' COVID-19 burdens were associated with their 30-day mortality rates among non-COVID patients during the early pandemic to inform whether widespread non-COVID outcomes data exclusion is necessary.

METHODS

We used the 100% CMS Inpatient Standard Analytic File and Master Beneficiary Summary File to include all Medicare inpatient encounters in the USA with admission dates from 1 April 2020 through 30 November 2020 (to allow for full 30-day mortality follow-up through the end of 2020). To reduce the possibility of results being impacted by inclusion of small hospitals with zero or few events, we restricted our analysis to hospitals with at least 25 non-COVID mortalities and tabulated the hospital-specific monthly proportion of encounters containing a COVID-19 diagnosis code (U07.1). Next, we excluded all encounters with a COVID-19 diagnosis in order to calculate hospital-specific observed/expected (O/E) 30-day mortality ratios among non-COVID patients using logistic regression to adjust for age, sex and individual comorbidities in the Elixhauser Mortality Index,[8] using a similar modelling approach to the USNWR methodology.[1] This approach allows for an approximately normal distribution of O/E hospital performance as well as controlling for any potential increases in non-COVID patient comorbidity burden during the pandemic. We further removed hospitals with O/E ratios above the 99th percentile and below the 1st percentile as outliers due to potentially incomplete or inaccurate claims data. Using linear regression, we modelled the association between hospitals' COVID-19 proportions and O/E ratios, testing whether the relationship was non-linear using a 5-knot cubic spline function and Wald t-test for the linear combination of spline terms.

Finally, we calculated the alternative hospital O/E ratios after selective data exclusions mirroring the USNWR data exclusion methodology.[1] Namely, we used the identical risk modelling approach described above, but for this alternative analysis we first excluded all monthly hospital-specific encounters from any month in which that hospital's proportion of inpatients with COVID-19 was either more than the national average or more than 15%.[1] Finally, we calculated and plotted the distribution of the pairwise relative difference of hospitals' O/E ratios comparing their performance under this alternative analysis (in which monthly data were excluded) versus the original analysis (in which no monthly data were excluded). These pairwise relative differences serve as an indicator of the extent to which intrahospital performance changed from the application of these widespread monthly data exclusions.

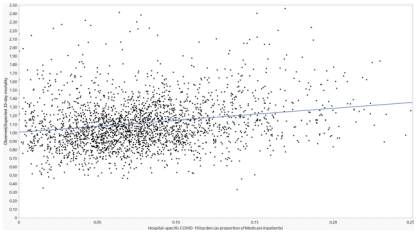

Figure 1 Relationship between hospital-specific COVID-19 burden and observed/expected (O/E) 30-day mortality among non-COVID Medicare patients hospitalised during the early COVID-19 pandemic. COVID-19 burden was defined as a hospital's proportion of Medicare patients who were hospitalised with COVID-19 during the time period. Expected 30-day mortality was adjusted for age, sex and Elixhauser comorbidities. The relationship between COVID-19 burden and non-COVID O/E 30-day mortality was approximately linear and statistically significant ($p<0.0001$), but weakly correlated ($r^2=0.06$).

RESULTS

We analysed 4 182 226 encounters from across 2601 US hospitals with a mean (SD) age of 72.3 (13.0). The c-statistic for the 30-day mortality model was 0.76. The association between hospital COVID-19 proportion and non-COVID O/E 30-day mortality was statistically significant (p<0.0001) with r^2=0.06. A 1% increase in COVID-19 inpatient proportion (eg, 5%–6%) was linearly associated (figure 1) with a 0.014 (95% CI 0.012, 0.016) increase in O/E mortality (eg, 0.900 to 0.914). The test for non-linearity was not significant (p=0.73).

In the alternative analysis in which hospital-specific months with greater-than-national-average COVID-19 proportion were excluded, there were 2 750 794 encounters (65.8% of original cohort) across 2571 hospitals (table 1). The c-statistic for the 30-day mortality model in this cohort was also 0.76. The median (IQR) pairwise relative difference in hospital O/E ratios comparing the

Table 1 Comparison of hospital-level characteristics when including all data versus excluding high-COVID burden months

Characteristic	All non-COVID inpatient encounters (n=2601 hospitals)	Non-COVID inpatient encounters with hospital-months excluded due to COVID-19 inpatient burden (n=2571 hospitals)*
Total encounters (median, IQR)	4 182 226	2 750 794
Age (median, IQR)	73 (71, 74)	73 (71, 74)
Female (%), (median, IQR)	53.1% (50.7%, 55.4%)	53.0% (50.4%, 55.6%)
Elixhauser comorbidities (%), (median, IQR)		
ALCOHOL	3.1% (2.2%, 4.1%)	3.1% (2.2%, 4.2%)
ANEMDEF	24.4% (19.8%, 29.3%)	24.5% (19.7%, 29.6%)
BLDLOSS	0.9% (0.6%, 1.3%)	0.9% (0.6%, 1.3%)
CHF	18.5% (15.6%, 21.5%)	18.4% (15.6%, 21.7%)
CHRNLUNG	26.1% (22.1%, 30.1%)	26.1% (22.2%, 30.5%)
COAG	6.8% (5.2%, 8.5%)	6.9% (5.2%, 8.7%)
DEPRESS	14.6% (10.9%, 18.5%)	14.6% (11.0%, 18.6%)
DMCX	24.2% (20.4%, 28.1%)	24.3% (20.2%, 27.9%)
HTN_C	64.8% (61.2%, 67.8%)	64.7% (61.0%, 68.0%)
LIVER	4.6% (3.6%, 5.9%)	4.7% (3.6%, 6.0%)
LYMPH	1.1% (0.7%, 1.4%)	1.1% (0.7%, 1.5%)
LYTES	40.0% (34.9%, 44.9%)	39.9% (34.9%, 45.2%)
METS	3.0% (2.2%, 3.9%)	3.0% (2.2%, 4.0%)
NEURO	12.7% (11.0%, 14.7%)	12.6% (10.9%, 15.0%)
OBESE	16.1% (12.5%, 20.5%)	16.1% (12.4%, 20.6%)
PARA	4.6% (3.5%, 5.9%)	4.5% (3.4%, 5.9%)
PERIVASC	7.8% (6.2%, 9.5%)	7.8% (6.1%, 9.6%)
PSYCH	4.3% (3.2%, 5.7%)	4.3% (3.1%, 5.8%)
PULMCIRC	0.8% (0.6%, 1.1%)	0.8% (0.5%, 1.2%)
RENLFAIL	23.7% (20.0%, 27.2%)	24.0% (20.2%, 28.4%)
TUMOUR	3.3% (2.6%, 4.0%)	3.3% (2.6%, 4.1%)
WGHTLOSS	7.9% (5.4%, 10.9%)	8.0% (5.5%, 11.1%)
Observed mortality (%), (median, IQR)	9.9% (8.5%, 11.5%)	9.6% (8.1%, 11.3%)
Expected mortality (%), (median, IQR)	9.2% (8.4%, 10.1%)	8.8% (8.1%, 9.7%)
Hospital-specific pairwise relative difference in O/E (%), (median, IQR)	Reference	+3.7% (–2.5%, +6.7%)

*n=30 hospitals had all months excluded due to greater-than-average monthly COVID-19 patient burden.
ALCOHOL, alcohol abuse; ANEMDEF, anaemia deficiency; BLDLOSS, blood loss anaemia; CHF, congestive heart failure; CHRNLUNG, chronic pulmonary disease; COAG, coagulopathy; DEPRESS, depression; DMCX, diabetes with chronic complications; HTN_C, hypertension; LIVER, liver disease; LYMPH, lymphoma; LYTES, fluid and electrolyte disorders; METS, metastatic cancer; NEURO, other neurological disorders; OBESE, obesity; PARA, paralysis; PERIVASC, peripheral vascular disease; PSYCH, psychoses; PULMCIRC, pulmonary circulation disorder; RENLFAIL, renal failure; TUMOUR, solid tumour without metastasis; WGHTLOSS, weight loss.

alternative analysis with the original analysis was +3.7% (−2.5%, +6.7%), with 1908/2571 (74.2%) of hospitals having relative differences within ±10% (eg, an O/E=1.10 or 0.90 vs an original O/E=1.00 when no months were excluded) and 2427 (94.4%) within ±25% (figure 2).

DISCUSSION

Our analysis indicated that during the initial period of the pandemic (April through November 2020), only 6% of the variation in US hospitals' risk-adjusted non-COVID 30-day mortality could be explained by their COVID-19 census burdens. Further, this relationship was approximately linear, making it potentially amenable for adaptation as a post hoc hospital-level 'COVID-census-adjustment' to outcome rates used in quality reporting. We found that roughly three-fourths of hospitals performed within ±10% relative difference in observed versus expected 30-day mortality among non-COVID beneficiaries regardless of whether we included or excluded hospital-specific months of data with greater than the national average COVID-19 burden. Excluding hospital-months based on proportion of inpatients with COVID-19 resulted in minimal change to hospitals' benchmarked performances at the expense of discarding the patient outcomes assessment for more than one-third of non-COVID encounters during this time frame.

Limitations include our usage of American data only, which was necessitated by our access to US Medicare claims in the absence of a centralised European or global claims data set. Similarly, the decision to include only hospitals with least 25 non-COVID mortalities during the study period due to statistical df considerations resulted in the exclusion of small US hospitals which may have shouldered a greater proportionate COVID-19 burden. However, our final cohort of 2601 hospitals is comparable to the CMS Overall Hospital Star Rating in which approximately 3000 hospitals are ultimately eligible for a rating each year and represents a majority of US hospitals. Second, our analysis relied on claims data and was therefore unable to adequately distinguish between exclusion of COVID-19 cases which may have been healthcare acquired rather than the main cause of admission, though there is evidence that hospital-acquired COVID-19 infections were not highly prevalent[9] in the early stages of the pandemic. Third, a more thorough economic impact analysis of the small proportion of hospitals that had significant differences in performance based on the methodology applied would be needed prior to implementation of any new methodological approaches in specific P4P programmes. For example, in the USA it would be prudent to investigate specific Hospital Readmission Reduction penalty implications for all hospitals, or likewise in the UK to investigate the hospital-specific implications if the Quality and Outcomes Framework P4P was not suspended during the pandemic. In fact, it may be more appropriate for these stakeholders themselves to conduct and publish these analyses publicly in the interest of transparency. Here, though, we specifically mirrored the USNWR exclusion methodology which is a hospital quality ranking system which does not have implications for hospital payment in order to first share a broad example of how pandemic-era data could plausibly be analysed and how results may or may not vary regardless of data exclusion methodologies. Similarly, most P4P programmes focus on specific conditions and procedures (such as hip and knee replacement or heart failure), and further investigation or alternative methodological approaches in those specific cohorts may yield different results than in our all-non-COVID patient 30-day mortality analysis. Because of sample size considerations, the early pandemic months may not be sufficiently powered in some conditions and thus be limited to large academic medical centres, whereas our approach to include all non-COVID conditions allowed for inclusion of most US hospitals. Finally, our analysis relies on a framework which assumes that hospitals' COVID-19 census burdens were a key influencer of hospital operations and quality outcomes during the pandemic, though other factors could be explored such as staffing, burnout and turnover.

A recent *JAMA* article showed that leading up to the pandemic, the USA experienced a decade of improvement in adverse patient events from 2010 to 2019.[10] The decade after the pandemic will bring improvements in electronic health record sharing and standardised collection of clinical data, and hospitals must find innovative ways to improve quality outcomes and increase accountability to their patients through transparent public reporting, rather than exclusion, of such data. In fact, such accountability may be of greater importance during times of uncertainty, such as the pandemic, which led to a decreasing amount of public trust in medical and scientific leaders.[11] It is imperative in the years ahead to extend analyses to understand hospital quality outcomes and performance, and hold hospitals accountable for the quality of care provided during the pandemic.

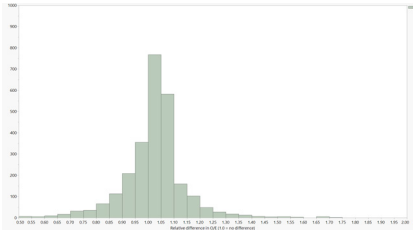

Figure 2 Histogram of relative difference in hospital-specific 30-day mortality observed/expected (O/E) comparing all data versus data excluding months of non-COVID patient data based on hospital-specific COVID burden. Y-axis is the total number of eligible hospitals (n=2571); x-axis is the relative difference in hospital-specific 30-day mortality O/E ratio comparing inclusion of all pandemic-era data for non-COVID patients versus a methodology of excluding hospital-months of non-COVID patient data based on the hospital-specific COVID burden (1.0=no difference).

The observed versus expected 30-day mortality rates among non-COVID Medicare beneficiaries during the first year of the COVID-19 pandemic in the USA varied only slightly in correlation with their burden of hospitalised COVID-19 beneficiaries, and exclusion of data during high-COVID census months did not meaningfully change the majority of hospitals' benchmarked mortality performance. For non-COVID patient outcomes such as mortality, evidence-based inclusion and adjustment for pandemic-era data is methodologically plausible and must be explored more rigorously prior to the pre-emptive and paradigm shifting exclusion of months or years of patient outcomes data.

Twitter Benjamin D Pollock @Pollock_BD

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; internally peer reviewed.

ORCID iD

Benjamin D Pollock http://orcid.org/0000-0003-4544-3158

REFERENCES

1 Methodology: U.S. News & World Report 2022-23 Best Hospitals: Specialty Rankings. *RTI international*. 2022.
2 Hospital-Acquired Condition Reduction Program. Table 1. key program dates for FY 2021 to FY 2023. In: *Centers for Medicare and Medicaid Services*. Available: https://www.cms.gov/files/document/fy-2022-hac-reduction-program-key-dates-matrix.pdf
3 Overall Hospita lQuality Star Rating Hospital-Specific Report User Guide. *Centers for medicare and medicaid services*. 2022. Available: https://mctools.sharepoint.com/teams/QNet-Reporting/Shared%20Documents/Forms/AllItems.aspx?FolderCTID=0x012000AE6904D45F49DF4ABEEE9E8B2F2CE6AA&id=%2Fteams%2FQNet%2DReporting%2FShared%20Documents%2FStar%20Ratings%2F2022%2FJULY%202022%2FJuly2022%2DOvrll%2DHsp%2DRtng%2DHUG%2Epdf&parent=%2Fteams%2FQNet%2DReporting%2FShared%20Documents%2FStar%20Ratings%2F2022%2FJULY%202022
4 Leapfrog Hospital Safety Grade Scoring Methodology. Spring 2022 safety grade. Available: https://www.hospitalsafetygrade.org/media/file/Safety-Grade-Methodology-Spring-20222.pdf [Accessed 19].2022
5 Department of Health and Social Care. Quality and outcomes framework (QOF) suspended for 2020-21. In: *GOV.UK*. 2020. Available: https://www.gov.uk/government/news/quality-and-outcomes-framework-qof-suspended-for-2020-21
6 Waitzberg R, Gerkens S, Dimova A, *et al*. Balancing financial incentives during COVID-19: a comparison of provider payment adjustments across 20 countries. *Health Policy* 2022;126:398–407.
7 Mayo Clinic COVID-19 Predictive Analytics Task Force, Pollock BD, Carter RE, *et al*. Deployment of an interdisciplinary predictive analytics Task force to inform Hospital operational decision-making during the COVID-19 pandemic. *Mayo Clin Proc* 2021;96:S0025-6196(20)31482-8:690–8.:.
8 Comorbidity Index Creation - Version 1.0. HCUP elixhauser comorbidity software. In: *Healthcare Cost and Utilization Project (HCUP)*. Rockville, MD: Agency for Healthcare Research and Quality, Available: https://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comindex2012-2015.txt [accessed 22 Apr 2019].
9 Habermann EB, Tande AJ, Pollock BD, *et al*. Providing safe care for patients in the coronavirus disease 2019 (COVID-19) era: a case series evaluating risk for hospital-associated COVID-19. *Infect Control Hosp Epidemiol* 2021;42:1479–85.
10 Eldridge N, Wang Y, Metersky M, *et al*. Trends in adverse event rates in hospitalized patients, 2010-2019. *JAMA* 2022;328:173–83.
11 Kennedy B, Tyson A, Funk C. Pew research center, feb 2022, americans' trust in scientists, other groups declines. 2022. Available: https://www.pewresearch.org/science/2022/02/15/trust-in-scientists-declines-methodology