BMJ Health & Care Informatics

From measures to action: can integrating quality measures provide system-wide insights for quality improvement decision making?

Inas S Khayal ,[1,2,3,4] Jordan T. Sanz[4]

To cite: Khayal IS, Sanz JT. From measures to action: can integrating quality measures provide system-wide insights for quality improvement decision making? *BMJ Health Care Inform* 2023;**30**:e100792. doi:10.1136/bmjhci-2023-100792

Received 03 May 2023
Accepted 21 June 2023

[1]The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA
[2]Biomedical Data Science, Geisel School of Medicine, Lebanon, NH, USA
[3]Cancer Population Sciences Program, Norris Cotton Cancer Center, Lebanon, NH, USA
[4]Department of Computer Science, Dartmouth College, Hanover, NH, USA

Correspondence to
Professor Inas S Khayal; inas.khayal@dartmouth.edu

ABSTRACT

Background Quality improvement decision makers are left to develop an understanding of quality within their healthcare system from a deluge of narrowly focused measures that reflect existing fragmentation in care and lack a clear method for triggering improvement. A one-to-one metric-to-improvement strategy is intractable and leads to unintended consequences. Although composite measures have been used and their limitations noted in the literature, what remains unknown is 'Can integrating multiple quality measures provide a systemic understanding of care quality across a healthcare system?'

Methods We devised a four-part data-driven analytic strategy to determine if consistent insights exist about the differential utilisation of end-of-life care using up to eight publicly available end-of-life cancer care quality measures across National Cancer Institute and National Comprehensive Cancer Network-designated cancer hospitals/centres. We performed 92 experiments that included 28 correlation analyses, 4 principal component analyses, 6 parallel coordinate analyses with agglomerative hierarchical clustering across hospitals and 54 parallel coordinate analyses with agglomerative hierarchical clustering within each hospital.

Results Across 54 centres, integrating quality measures provided no consistent insights across different integration analyses. In other words, we could not integrate quality measures to describe how the underlying quality constructs of interest—intensive care unit (ICU) visits, emergency department (ED) visits, palliative care use, lack of hospice, recent hospice, use of life-sustaining therapy, chemotherapy and advance care planning—are used relative to each other across patients. Quality measure calculations lack interconnection information to construct a story that provides insights about where, when or what care is provided to which patients. And yet, we posit and discuss why administrative claims data—used to calculate quality measures—do contain such interconnection information.

Conclusion While integrating quality measures does not provide systemic information, new systemic mathematical constructs designed to convey interconnection information can be developed from the same administrative claims data to support quality improvement decision making.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Quality improvement decision makers are left to develop improvement intervention decisions from a deluge of quality measures that are siloed, non-actionable and for which it is unclear how they relate to each other.

WHAT THIS STUDY ADDS

⇒ Our findings show that quality measures cannot be integrated to provide insights to support a system-wide understanding of care delivery and lack any interconnection information about other quality constructs. Fundamentally, this is likely because quality measure calculations do not include any interconnection information.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Despite the limitations in quality measure calculations, administrative claims data contain interconnection information that can support system-wide quality improvement decision making.

INTRODUCTION

Over two decades ago, two Institute of Medicine (IOM) reports, *To Err Is Human* and *Crossing the Quality Chasm*, focused our attention on the need for substantial improvement in the quality of healthcare.[1 2] In addition, two more IOM reports[3 4] identified concerning end-of-life cancer care patterns, such as very late chemotherapy use, very short hospice enrolment, repeated hospitalisation during patients' last month of life and insufficient use of palliative care.[5 6]

In response, healthcare stakeholders produced a deluge of quality measures. More recent IOM reports stressed that measures are narrowly focused, poorly delineated, reflect existing fragmentation in care and lack a clear method for triggering improvement.[7 8] Cancer care delivery serves as an exemplar of a complex, multidimensional system of providers from multiple specialties providing care across many settings,[7] where it is not clear how fragmented quality measures account for interactions between interdependent siloed entities. Shwartz *et al* and Levesque *et al* suggest that while individual quality measures identify a single aspect of care, they do not provide an accessible whole-system overview of quality.[9 10]

Quality improvement decision makers are left to develop an understanding of quality within their healthcare system from many measures that lack actionable information.[11–14] A one-to-one metric-to-improvement strategy is (1) intractable and (2) incorrectly assumes independence between service utilisation. The literature includes examples of actions to improve one aspect of care that leads to unintended consequences—negatively affecting other care aspects. Teno *et al* showed that in an effort to provide patient-concordant wishes of dying at home, actions resulted in fewer decedents dying in hospitals but also unintended consequences of increases in intensive care and care transitions.[15] Consequently, since actions are interdependent (one measure likely to impact others), it is necessary to develop a system-wide understanding of how different elements of quality underlying quality measures are interconnected and impact each other.

A popular approach to developing an overall understanding of quality has been to combine quality measures into a composite measure—as endorsed by two IOM reports[8 16]—to summarise information and reduce cognitive load.[17 18] The value of composite measures has been mixed, some reporting substantive interpretability, internal consistency and accounting for the majority of variation,[19–22] while others report significant limitations of producing widely varying results with poor face validity and lacking the ability to signal specific changes to be targeted for improvement.[14 23–25]

Consequently, quality improvement decision makers require system-wide insights about where, when or what care is provided to which patients to construct a story about how differential aspects of care are interconnected and impact each other in order to determine what systemic actions can potentially improve the overall quality of care delivered in a manner that is efficient and minimises unintended consequences. To provide quality improvement decision makers with system-wide insights based on multiple quality measures, we devised a data-driven strategy that includes machine learning to determine if consistent insights exist about the differential utilisation of end-of-life care in order to answer the question, 'Can integrating multiple quality measures provide a systemic (interconnected) understanding of care quality across a healthcare system for end-of-life cancer care'?

METHODS

Study population and quality measure data

We downloaded hospital-level, end-of-life cancer care quality measures from the Dartmouth Atlas from a replication repository[26] (https://doi.org/10.21989/D9/BWKLG5). The data represent an exemplary set of quality measures typically reported for hospitals. The data come from the Center for Medicare and Medicaid Services' (CMS) Medicare program—the only US national health insurance for people 65 years and older. Briefly, the data were based on a 100% sample of Medicare fee-for-service (FFS) beneficiaries drawn from 2015 to 2016 Centers for Medicare and Medicaid Services files that included the (a) master beneficiary summary file, (b) Medicare analysis and review (MedPAR) file, (c) physician/supplier carrier file, (d) outpatient file and (e) hospice file. We used a 40% subsample of these FFS beneficiaries drawn from (f) drug coverage/Part D files.

The cohort included Medicare FFS beneficiaries with poor prognosis cancers (metastatic cancers and primary cancers associated with high risk of mortality) with the following inclusion criteria: died between 1 April 2016 and 31 December 2016, between the ages 66 and 99, and for whom there existed a complete 6-month look-back period between 1 October 2015 and 31 March 2016. With a focus on cancer, we included hospitals with a cancer designation of National Cancer Institute (NCI) or National Comprehensive Cancer Network (NCCN).

The data included eight quality measures, including five National Quality Forum (NQF)-endorsed end-of-life cancer care quality metrics: (a) receipt of chemotherapy in the last 14 days of life (NQF#0210) (named *Chemo*); (b) more than one emergency department (ED) visit in the last 30 days of life (NQF#0211) (*ED*); (c) intensive care unit (ICU) admission in the last 30 days of life (NQF#0213) (*ICU*); (d) non-referral to hospice (NQF#0215) (*No Hospice*) and (e) late (NQF #0216) referral to hospice, defined as within 3 days of death (*Recent Hospice*) and three other measures (f) life-sustaining treatment rates in the last 30 days of life (eg, mechanical ventilation, hemodialysis, feeding tubes or cardiopulmonary resuscitation) (*Life Sustaining*); (g) palliative care claims in the last 6 months of life, using International Classification of Diseases, Tenth Revision (ICD-10) diagnosis codes Z51.5 (*Palliative Care*) and (h) advance care planning claims in the last 6 months of life, based on billing codes G9054, S0257, 99497 and 99498 (*Advanced Care*). Further calculation details can be found in Wasp *et al*.[27]

Data-driven analysis strategy

We normalised data variables and performed four classes of analyses. We performed a unit-based normalisation to scale each variable to the range of 0–1. For each measure, we subtracted each value from the minimum and divided by the range (max–min). After normalisation, we reversed the scales for variables *Palliative Care* and *Recent Hospice* to ensure that higher values represented poorer quality for all variables.

First, we determined the count and distribution for each variable, calculated the Pearson correlation coefficients for every variable combination and calculated histograms, continuous probability density lines, linear regression and kernel density estimate plots. Second, we performed a principal component analysis (PCA) to evaluate the variance in the quality measures. We calculated the cumulative percentage of the variance explained by each additional principal component. Third, we plotted quality measures on a parallel coordinate plot and performed a clustering analysis for multidimensional pattern recognition. Specifically, we organised the quality measures into a vector per hospital and applied a bottom-up agglomerative hierarchical clustering using the Wards minimum variance method. We first performed this analysis with the two most available variables. Next, we added to the parallel coordinate system one variable at a time, including the hospitals that had all the variables available for each analysis. We then ranked the clusters by summing the normalised quality values and dividing by the number of hospitals in the cluster. Fourth, and for each hospital, we plotted a line for every combination of two measures available for a hospital. For every combination of two measures, we performed an agglomerative hierarchical clustering (as described above) to group similar lines into eight groups, chosen based on visual inspection and a dendrogram analysis. We visualised each group with a different colour (red, orange, yellow, green, turquoise, light blue, blue or purple). Red indicates the average fell into the highest and worst-performing cluster, orange indicates the average fell into the second highest and second worst-performing cluster and so on. Purple indicates the average fell into the lowest and best-performing cluster.

Analyses were performed using Python V.3.7.

RESULTS

There were 54 NCI and NCCN-designated cancer hospitals/centres in our dataset. Each hospital had the potential to include up to eight cancer quality measures. To assess the correlation and independence of these quality measures, we visualised the correlation, distribution, actual points and kernel density estimation for every combination of the eight cancer quality measures in figure 1. Because of a low number of hospitals with the variables, *Chemo* and *Advanced Care*, we excluded them from the remainder of the analysis, as they showed the least normal Gaussian distributions and existed for only 18.5% and 27.8% of the hospitals, respectively. Four combinations of a possible 36 showed significantly correlated measures, shown with a shaded red background in figure 1. These include a weak but significant (1) positive correlation between *ICU* and *Life Sustaining* (r=0.47, p=0.0007), *ICU* and *Recent Hospice* (r=0.46, p=0.0039) and *Life Sustaining* and *No Hospice* (r=0.51, p=0.0002) and (2) negative correlation between *Recent Hospice* and *No Hospice* (r=−0.52, p=0.0007). These figures also highlight that these quality measures are not normally distributed.

Does the number of quality measures affect a systemic understanding of a hospital?

To evaluate the variance and determine how the number of used quality measures affects identified orthogonal principal components, we conducted three principal component analyses with 4, 5 and 6 quality measure variables. Figure 2A shows a PCA biplot with four quality measures (n=47 hospitals), consisting of the variables *ICU*, *Life Sustaining*, *Palliative*

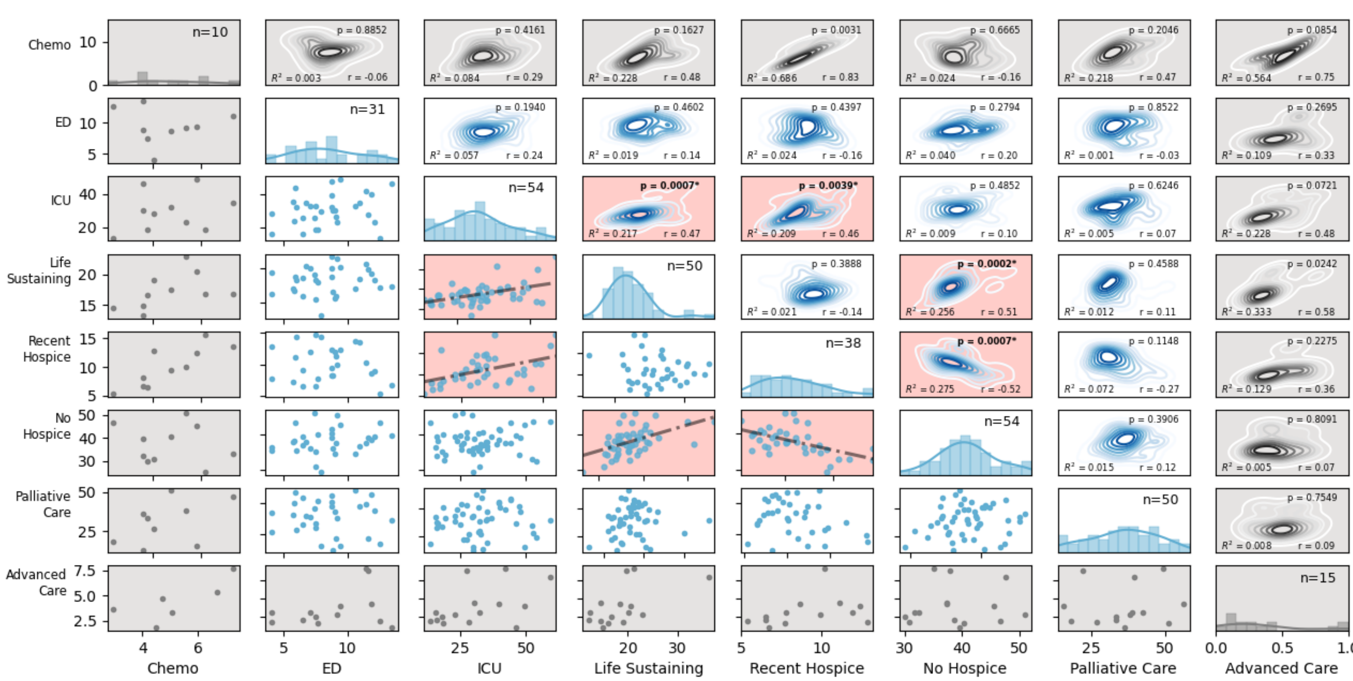

Figure 1 A conglomerate of figures showing (1) normalised histograms for each quality measure (on the diagonal), (2) kernel density estimation (upper right quadrant) and (3) individual hospital points for every combination of the eight cancer quality measures (lower left quadrant). Significantly correlated variables are shaded in red. Variables available for a very low number of hospitals were excluded from further analysis and shaded in grey. ED refers to emergency department and ICU refers to intensive care unit.

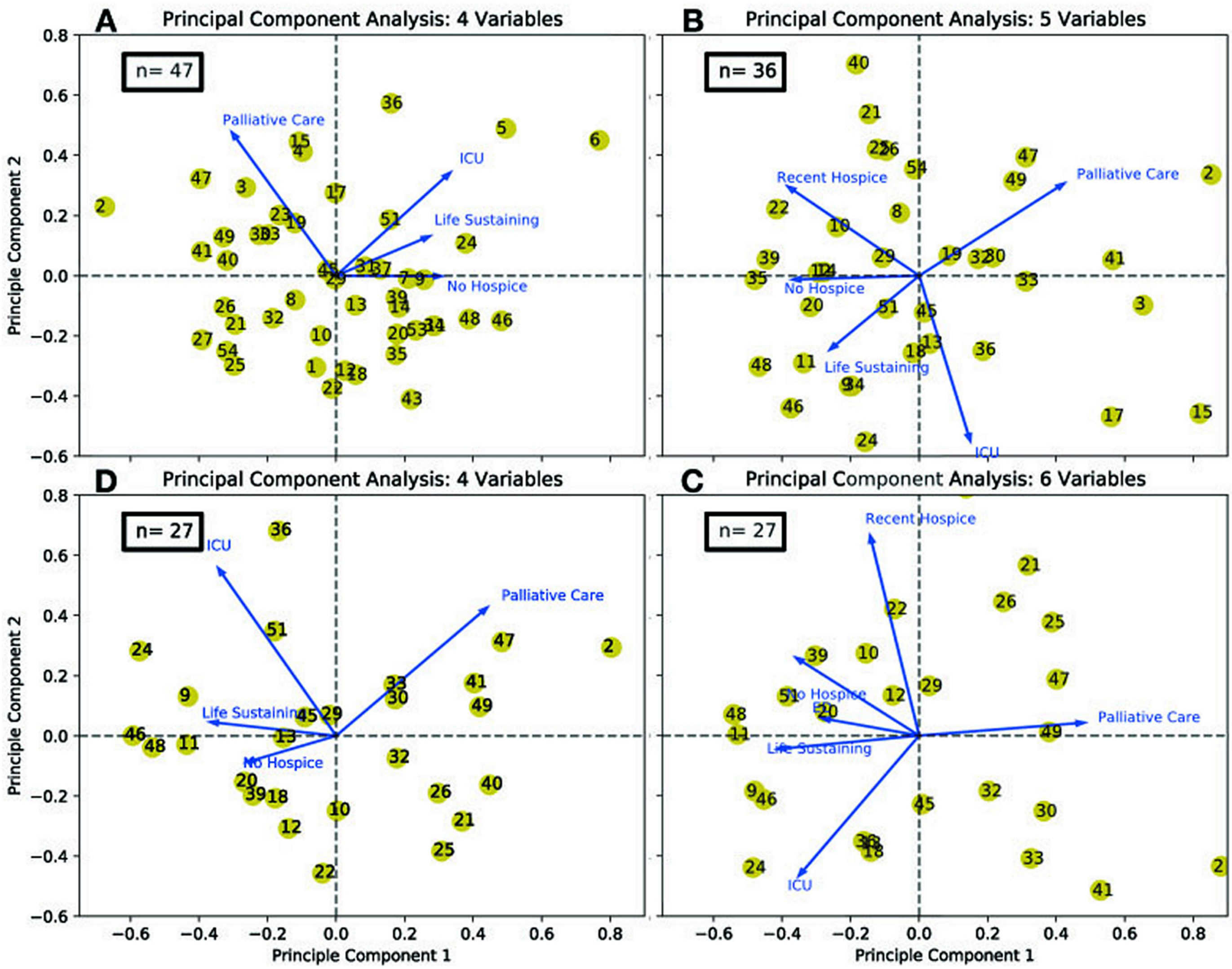

Figure 2 Principal component analysis (PCA) biplot for (A) 4, (B) 5, (C) 6 variables and (D) 4 variables with a reduced hospital set found in 6 variables. The contribution of each variable to principal components 1 and 2 is shown in blue arrows. The numbered yellow points represent the hospital number with transformed quality measures into principal component coordinates. Adding additional quality measure variables to the PCA leads to differences in the contributions to the principal components and, therefore, a different understanding of the components. ED refers to emergency department and ICU refers to intensive care unit.

Care and *No Hospice*. Figure 2B shows a PCA with five quality measures (n=36 hospitals), consisting of the variables from the previous analysis, with the addition of *Recent Hospice*. Figure 2C shows a PCA with six quality measures (n=27 hospitals) consisting of the variables from the previous analysis and *ED*. These biplots reveal how each variable contributes to each principal component. The blue arrows highlight the direction and magnitude of the contribution of each variable to the first and second principal components. To more precisely investigate the effects of changing the number of variables and samples, figure 2D shows a PCA biplot with four quality measures that only includes the (n=27) hospitals with all six quality measure variables. Additional variables incorporated additional information relative to each other. Despite the correlation among the quality measures, a change in the number of quality measures led to very different contributions to the first and second principal components, as the number of quality measures was increased from 4 to 6. This was also evident in the remainder of the principal components. Furthermore, *changing the number of hospitals* in the four-quality measure analysis led to somewhat different contributions (both in terms of eigenvalues and eigenvectors) from each variable towards the principal components, as illustrated by the visual difference between figure 2A (n=47) and figure 2D (n=27). In addition, and much more strongly, *changing the number of quality measures* led to different contributions from each variable towards the principal components, as illustrated by the difference between figure 2D and figure 2C, likely due to the correlation between the variables. This leads to components with different meanings and contributions from each quality measure component.

To further investigate how the number of quality measures affects a systemic understanding of care delivery, we conducted multiple parallel coordinate analyses. We applied

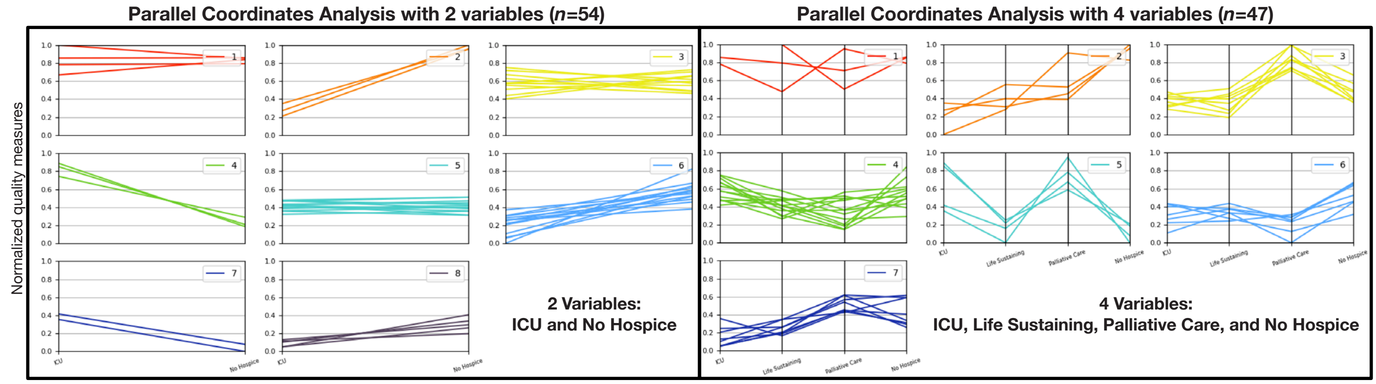

Figure 3 Parallel coordinates analyses with 2 variables (left) and 4 variables (right). Hospitals with a similar set of quality measures (ie, similar pattern) are clustered together. Clusters are coloured and numbered for each analysis. Cluster 1 signifies the highest averages and worst performing set of quality measures, while cluster 8 for two variables, or 7 for four variables, signifies the lowest averages and best performing set of quality measures. ED refers to emergency department and ICU refers to intensive care unit.

agglomerative hierarchical clustering to a set of quality measures organised in a parallel coordinate system. We analysed two, three, four, five and six quality measure combinations. First, we clustered each hospital by only considering its *ICU* and *No Hospice* measures with n=54 hospitals, as shown in figure 3A. We also visualised the four-variable analysis in figure 3B, using the same four variables from the PCA, which included: *ICU*, *Life Sustaining*, *Palliative Care* and *No Hospice*. We aligned the clusters to most closely match across analyses. The multiple parallel coordinate cluster analyses indicate that the addition of multiple variables affects the majority of the hospitals that do not include the most extreme quality measures. This highlights that the number of quality measures used affects the group a hospital belongs to and, therefore, the overall systemic understanding of end-of-life care delivery.

Does the choice of a specific set of quality measures affect a systemic understanding of end-of-life care delivery?

To understand how the choice of a specific set of quality measures affects a systemic understanding, we analysed combinations of quality measures, two variables at a time, across all hospitals. We performed an agglomerative hierarchical clustering and identified eight clusters, as shown in figure 4 (left). Again, cluster 1 corresponded to the highest average of the two variables (poor quality),

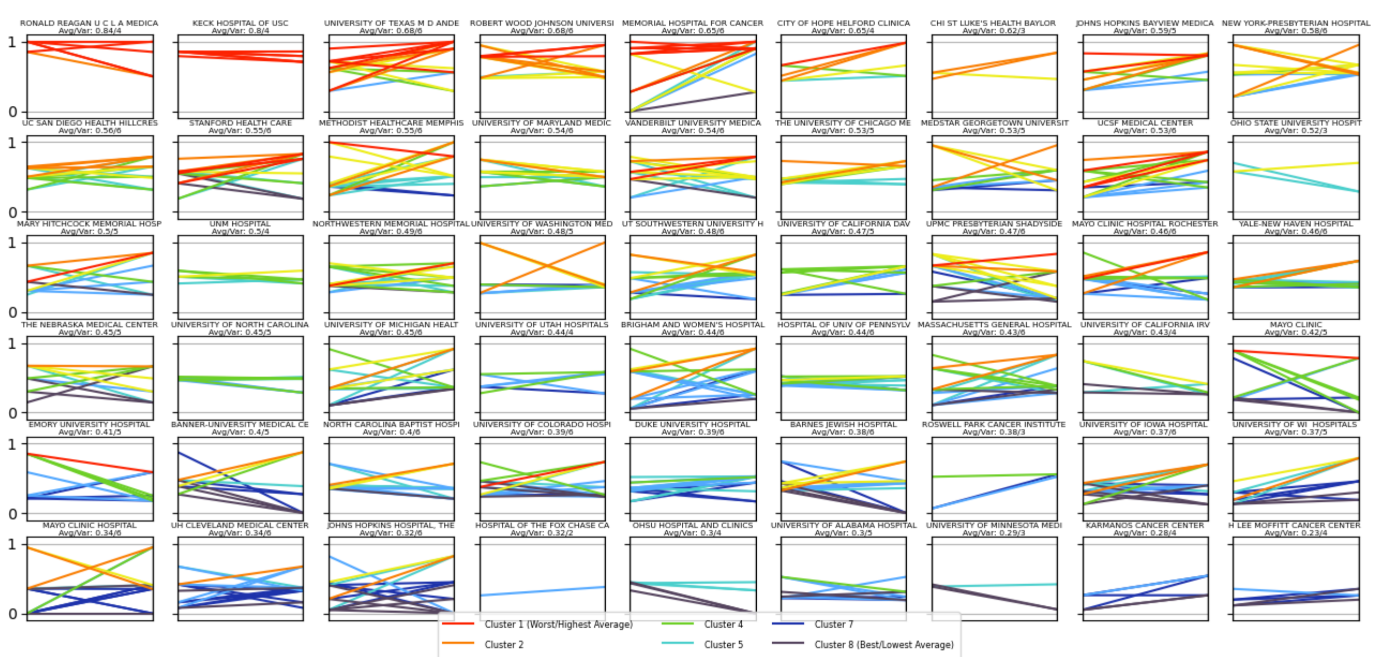

Figure 4 All 54 hospitals, showing the clusters corresponding to every combination of two quality measures available for each hospital. A red line colour indicates the highest quality average cluster ranking (the poorest quality) for a two-variable combination, while a purple line indicates the lowest quality average cluster ranking (the highest quality). The hospitals are ordered from the worst overall average (top left) to the best overall average (bottom right).

and cluster 8 corresponded to the lowest average (high quality). The hospital plots are ordered from the worst overall average (top left) to the best overall average (bottom right). The corresponding cluster numbers are indicated by the colour of the line, as described above. Hospitals towards the extreme average values have relatively similar cluster rankings for all the combinations of variables, while most hospitals in the middle have a wide variety of cluster rankings, shown by the changes in line colour. This clearly shows that even within a hospital, the choice of quality measures used can lead to a completely different view of the hospital. Only in a very limited number of hospitals (eg, Keck Hospital of USC) do all the quality value combinations fall on the extreme end and lead to the same (red) colour categorisation for all combinations of quality measures. Thus, similar to the parallel coordinate analysis, only the most extreme values, in this case, one of 54 hospitals, can be reliably categorised from any set of quality measures. This highlights that the strategy of how quality measures are combined affects the analysis.

Does the strategy of combining quality measures affect a systemic understanding of end-of-life care delivery?

We analysed a combination of quality measures using the four techniques described in the methods for a combination of 92 views, including 28 correlation analyses, 4 principal component analyses, 6 parallel coordinates analyses with agglomerative hierarchical clustering across hospitals and 54 parallel coordinate analyses with agglomerative hierarchical clustering within each hospital. These experiments demonstrate an iterative strategy to combine quality measures. Only in the case of a parallel coordinate analysis with agglomerative hierarchical clustering *across* and *within* hospitals lead to the same conclusion for Keck Hospital of USC. Otherwise, the remaining experiments, using different strategies, led to different insights and a different view of the system depending on the strategy used.

DISCUSSION

In this data-driven quantitative analysis, quality measures varied in their existence, value and distribution, even after normalisation, and provided no consistent insights across different integration analyses. In other words, we could not integrate quality measures to describe how the underlying quality constructs of interest—ICU visits, ED visits, palliative care use, lack of hospice, recent hospice, use of life sustaining therapy, chemotherapy and advance care planning—are used relative to each other across patients. Our findings suggest that quality measure calculations lack any interconnection information that can be potentially used to construct a story to provide insights about where, when or what care is provided to patients that can be used to support quality-improvement decision making. This is possibly for two reasons, (1) quality measure calculations mathematically abstract out siloed aspects of care across a cohort of patients, and by doing so, (2) disregard patient-level healthcare utilisation trajectories that contain the longitudinal interconnection between different services delivered. And yet, we posit that the administrative claims data—used to generate these quality measures—actually *do* contain such interconnection information because claims data include patient-level healthcare trajectories that include sequence order and timing for the delivery of different services at different places. Viewing healthcare systems from a lens of 'general systems theory' helps to explain why examining the parts of the system independently does not provide a systemic view of the behaviour of care delivery since emergent behaviours arise from the system that do not appear in any individual component.

Interconnection information can be elucidated by understanding the sequential utilisation of care by patients. For example, the interconnection between ICU visits and Life Sustaining Therapy can be examined by exploring their sequential utilisation across patients. *Does Life Sustaining Therapy only occur following an ICU visit, or does it occur following an ED or inpatient (hospital) visit? Does Life Sustaining Therapy occur after all, most, some or no ICU visits?* Similarly, the interconnection between Life Sustaining Therapy and Palliative care can be elucidated by examining their sequential utilisation. Interconnection information from sequential utilisation across multiple sets of quality constructs can then be used to develop insights about the behaviour of care delivered within a healthcare organisation that can be used to provide more actionable information about potentially how, where, when and for which patients to intervene to improve quality for multiple quality measures at once. Quality improvement decision making that takes into account multiple aspects of care quality, may also help alleviate unintended consequences that may arise from focusing on a single measure for improvement without reflecting on its impact on other measures, that is, other aspects of care. For example, Sedhom *et al* describe that in an effort to provide patient-concordant wishes of dying at home and not in a hospital, actions taken resulted in more patients dying not only at home but also in nursing facilities. For older patients, this action to avert a hospital death has led to the unintended consequences of only a 25% chance of dying at home and a much higher chance of dying in a nursing facility, where less attention is provided to symptoms, existential distress and grief compared with patients that are able to remain in their home.[28]

In quality measurement calculations, patients' real-life sequential utilisation of care is ignored and *only* the relevant quality construct is extracted for each quality measure calculation independently. For each quality measure, the relevant quality construct (ICU use, palliative care use and so on) is aggregated across patients, essentially providing a siloed organisation perspective of quality that is disconnected from the *patient's* perspective of quality. This calculation has two critical limitations. First, it does not distinguish, how many and which patients contributed to each quality measure—information which is insightful for quality improvement decision making. For example, if quality measures for ICU visits, ED visits and palliative care were 'poor', it is unclear if that is a result of different patients contributing to each of these quality measures independently, or if it is the same patients contributing to all these quality measures. In the former, quality improvement efforts may target different patients with different interventions, whereas, in the latter, a single intervention for all patients may ameliorate poor quality care for all measures. Second, quality measure calculations fundamentally ignore the temporal contiguity of care that can provide insights as to how patients are trajecting through the healthcare system both in terms of where they receive care (ICU, ED and so on) and what care they receive (palliative, life-sustaining therapy and so on). Instead, by aggregating patient-level trajectory information, underlying care patterns can emerge if patterns exist across patients.[29] Hofstede *et al* have shown the importance of patient-level data, and how ecological fallacy potentially influences the interpretation of hospital performance when patient-level associations are not taken into account.[30] Indeed, temporal patterns of care across patients can provide contextual insights about how, when, where and what care is delivered and, consequently, provide potential information as to how, when or where to intervene. Khayal *et al* have developed new systemic mathematical constructs (images) designed to convey interconnection information from administrative claims data to provide quality information in a new mathematical construction that goes beyond instances of many single quality measures.[29]

General systems theory describes that for systems, such as healthcare systems, with interconnected components, emergent behaviours arise from the system that does not appear in any individual components.[31] For cancer care delivery, and in the typical case of a patient receiving care from multiple specialties across several settings, a change in a healthcare provider's decisions and actions can change the context for other healthcare providers. This phenomenon of complex systems explains why an examination of the parts, such as through quality measures, gives us no information about the coordination of parts and processes.[30] In other words, it is in the coordination and interconnection of information between the parts or processes that higher-level behaviours emerge. It is those behaviours that need to be understood to develop quality improvement interventions that can target them. While very few examples exist of using system-based methodologies such as discrete-event simulation, agent-based modelling and others to develop an understanding of the behaviour of a system from administrative claims data,[29] it is fundamentally in the aggregation of the sequence of delivered care (patient healthcare trajectory) where patterns of systemic interconnection can emerge. Consequently, a systems approach to quality measurement—with the creation of systemic mathematical constructs—is very likely required to make system-level improvement decisions that take into consideration upstream and downstream effects to minimise unintended consequences.

CONCLUSION

A system-wide understanding of a healthcare system would provide quality improvement decision makers with the needed systemic understanding to make decisions as to how to improve the system as a whole, especially when it is unclear as to which part to target for improvement. A systems framework helped to explain why system-wide healthcare delivery behaviour cannot be explained from the parts and suggests that a systems approach to quality measurement is likely required to make system-level improvement decisions that take into consideration upstream and downstream effects to minimise unintended consequences.

Acknowledgements The authors would like to acknowledge Amber Barnato, MD, MPH, MS for her early contributions and funding support by the American Cancer Society Award (RSG-18-017-01-CPHPS) and the Levy Cluster in Health Care Delivery Science at Dartmouth.

Contributors ISK conceived the study, performed the analysis and wrote the paper. JS performed the initial analyses and contributed to the writing. Both ISK and JS reviewed and approved the final manuscript. ISK is the guarantor.

Funding The authors would like to acknowledge the funding support from the American Cancer Society Awards (RSG-22-128-01-HOPS and RSG-18-017-01-CPHPS) and the Levy Cluster in Health Care Delivery Science at Dartmouth.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. G. Wasp, S. Alam, G. Brooks, I. Khayal, N. Kapadia, D. Carmichael, A. Austin, and A. Barnato, "Replication Data for: Quality of EOL Care for Medicare Decedents at Minority-Serving Cancer Centers: A Retrospective Study," https://doi.org/10.21989/D9/BWKLG5, Dartmouth Dataverse, V4, 2019.

ORCID iD

Inas S Khayal http://orcid.org/0000-0001-8354-6006

REFERENCES

1 Donaldson MS, Corrigan JM, Kohn LT, *et al*. *To err is human: building a safer health system*. 2000.
2 Corrigan JM. Crossing the quality chasm. 2005;89.
3 Simone JV, Hewitt M. *Ensuring quality cancer care*. 1999.
4 Foley KM, Gelband H. *Improving palliative care for cancer*. Washington, DC: National Academy Press, 2001.
5 Earle CC, Neville BA, Landrum MB, *et al*. Trends in the aggressiveness of cancer care near the end of life. *J Clin Oncol* 2004;22:315–21.
6 Care F-O, Planning AC. *Dying in america: Improving quality and honoring individual preferences near the end of life*. 2014.
7 Levit L, Balogh E, Nass S, *et al*. *Delivering high-quality cancer care: charting a new course for a system in crisis*. Washington, D.C: National Academies Press, 2013.
8 Blumenthal D, McGinnis JM. Vital signs: core Metrics for health and health care progress. *JAMA* 2015;313:1901–2.

9 Shwartz M, Restuccia JD, Rosen AK. Composite measures of health care provider performance: a description of approaches. *Milbank Q* 2015;93:788–825.
10 Levesque J-F, Sutherland K. Combining patient, clinical and system perspectives in assessing performance in Healthcare: an integrated measurement framework. *BMC Health Serv Res* 2020;20:23.
11 Shekelle P, Lim Y-W, Mattke S, *et al*. *Does public release of performance results improve quality of care. A systematic review*. London: The Health Foundation, 2008.
12 Werner RM, Konetzka RT, Stuart EA, *et al*. Impact of public reporting on quality of Postacute care. *Health Serv Res* 2009;44:1169–87.
13 McGinnis JM, Malphrus E, Blumenthal D, *et al*. *Vital signs: core Metrics for health and health care progress*. National Academies Press, 2015.
14 Barclay M, Dixon-Woods M, Lyratzopoulos G. The problem with composite indicators. *BMJ Qual Saf* 2019;28:338–44.
1 Teno JM, Gozalo PL, Bynum JPW, *et al*. Change in end-of-life care for Medicare beneficiaries: site of death, place of care, and health care transitions in 2000, 2005, and 2009. *JAMA* 2013;309:470–7.
16 Ring JC, Chao SM. Performance measurement: accelerating improvement (retrieved from Institute of medicine). 2006.
17 Cerully JL, Parker AM, Rybowski L, *et al*. Improving patients' choice of clinician by including roll-up measures in public Healthcare quality reports: an online experiment. *J Gen Intern Med* 2019;34:243–9.
18 Composite measure evaluation framework and national voluntary consensus standards for mortality and safety—composite measures: a consensus report. 2009.
19 Zaslavsky AM, Shaul JA, Zaborski LB, *et al*. Combining health plan performance indicators into simpler composite measures. *Health Care Financ Rev* 2002;23:101–15.
20 Dimick JB, Staiger DO, Baser O, *et al*. Composite measures for predicting surgical mortality in the hospital. *Health Affairs* 2009;28:1189–98.
21 Trzeciak S, Gaughan JP, Bosire J, *et al*. Association between Medicare summary star ratings for patient experience and clinical outcomes in us hospitals. *J Patient Exp* 2016;3:6–9.
22 Kolfschoten NE, Gooiker GA, Bastiaannet E, *et al*. Combining process indicators to evaluate quality of care for surgical patients with colorectal cancer: are scores consistent with short-term outcome *BMJ Qual Saf* 2012;21:481–9.
23 Friebel R, Steventon A. Composite measures of Healthcare quality: sensible in theory, problematic in practice. *BMJ Qual Saf* 2019;28:85–8.
24 Bilimoria KY, Barnard C. The new CMS hospital quality star ratings: the stars are not aligned. *JAMA* 2016;316:1761–2.
25 Chung JW, Dahlke AR, Barnard C, *et al*. The centers for Medicare and Medicaid services hospital ratings: pitfalls of grading on a single curve. *Health Affairs* 2019;38:1523–9.
26 Wasp G, Alam S, Brooks G, *et al*. n.d. Replication data for: quality of EOL care for Medicare Decedents at minority-serving cancer centers: a retrospective study. *Dartmouth Dataverse*
27 Wasp GT, Alam SS, Brooks GA, *et al*. Quality of end-of-life care for Medicare Decedents at minority-serving cancer centers: A retrospective study. *Cancer Med* 2020;9:1911–21.
28 Sedhom R, Kuo P-L, Gupta A, *et al*. Changes in the place of death for older adults with cancer: reason to celebrate or a risk for unintended disparities. *J Geriatr Oncol* 2021;12:361–7.
29 Khayal IS, Brooks GA, Barnato AE. Development of dynamic health care delivery Heatmaps for end-of-life cancer care: a cohort study. *BMJ Open* 2022;12:e056328.
30 Hofstede SN, van Bodegom-Vos L, Kringos DS, *et al*. Mortality, readmission and length of stay have different relationships using hospital-level versus patient-level data: an example of the ecological fallacy affecting hospital performance indicators. *BMJ Qual Saf* 2018;27:474–83.
31 Von Bertalanffy L. *The meaning of general system theory. General system theory: Foundations, development, applications 30–53*. 1973.