This document contains objective evaluations of various report cards. HANYS neither endorses, nor should be taken to endorse, any particular report card. HANYS has no financial or other interest in any report card it evaluates, other than providing informed educational material in reviewing the report cards. Each hospital is encouraged to make independent conclusions about the various report cards, including whether or not to use the report card information to drive quality improvement, and whether or not to respond to a request to participate in any quality reporting initiative.
HANYS supports the availability of hospital quality and safety information to help patients make choices and assist providers in improving care. However, the information must be based on a standard set of measures that have been proven to be valid, reliable and evidence-based to ensure a more accurate representation of the quality of care delivered.

HANYS envisions a future in which consumers have access to information that is meaningful, accurate, reliable and relevant to their unique healthcare circumstances, so they feel empowered to make informed choices about their care. We also support reliable, nationally vetted measures of patient-reported outcomes to better engage consumers in their care. To achieve this vision, HANYS urges all stakeholders to reduce the number of report cards so that information is less confusing for consumers and providers can focus on improving patient care.

The proliferation of hospital report cards has not achieved their stated goal of helping consumers understand the quality of care offered at hospitals. Consumers looking to make informed healthcare decisions end up confused by multiple, contradictory reports. While well intentioned, these reports provide conflicting information and produce dramatically different ratings. Likewise, policymakers and healthcare providers looking to bolster quality improvement efforts derive limited value from these reports. In addition, responding to multiple requests for data unfairly burdens providers.

HANYS developed Report on Report Cards as an educational resource for hospital leaders and their boards; it serves as a primer for evaluating and responding to publicly available consumer report cards.

HANYS identified a set of guiding principles to which report cards should adhere. The data shared in these reports must be:

- transparent;
- evidence-based;
- aligned with national measures;
- rooted in clinical care;
- timely;
- risk-adjusted;
- valid and reliable;
- from consistent timeframes and sources;
- available to providers for review before publication to correct errors;
- published under a business model free from conflicts of interest; and
- representative of the population.
# TABLE OF CONTENTS

1 MEASUREMENT ENVIRONMENT
   Report card variation puts an undue burden on patients 1
   Measure madness 1
   Insufficient progress has been made 3
   Limitations of EHR technology 4
   New challenges 5
   Shift to eCQMs 5
   Proliferation of social media ratings 6
   The composite craze 7
   Measuring quality across the continuum 8
   Incorporating population health 9
   The challenge of commercial payer quality incentives 10
   Scarce data for the broader population 10

11 CALL TO ACTION

13 GUIDING PRINCIPLES FOR EVALUATION

15 HOW HANYS SCORED THE REPORT CARDS

17 HANYS REPORT CARD RATINGS

18 KEY FINDINGS

20 DISCUSSION OF KEY FINDINGS

22 LIST OF QUALITY REPORT CARDS

25 HANYS’ STATEWIDE STEERING COMMITTEE ON QUALITY INITIATIVES

26 ENDNOTES
Patients and healthcare providers face a proliferation of publicly available reports rating the quality of care provided in hospitals. Supporters of hospital report cards promote them as a means to improve the quality of care and help consumers make better-informed healthcare choices. However, these goals are thwarted by multiple reports that use conflicting information and produce dramatically different ratings. Despite the confusion that contradictory reports create, hospital report cards continue to garner attention from consumers, hospitals engaged in quality improvement efforts and policymakers seeking to make improvements to the healthcare delivery system.

The healthcare industry is at a crossroads — the call for less measurement is amplifying, but new challenges and the increased (and anticipated continued increase) in consumer interest in this type of information means that more reports could proliferate. Without action, the burden on providers, societal cost and lack of value to patients will worsen.

In 2008 and 2013, HANYS published previous versions of this report, which garnered national attention, amplifying the call for reducing measurement. However, this rapidly changing environment has resulted in a continued evolution of “measure madness.” More needs to be done to promote value in the nation’s healthcare system and provide consumers with meaningful information to make decisions about their care.

**REPORT CARD VARIATION PUTS AN UNDUE BURDEN ON PATIENTS**

Consumer engagement in healthcare continues to grow, from searching for care and accessing new channels of care to tracking and sharing health data. A 2018 survey by Deloitte Center for Health Solutions found that more consumers are using quality ratings and other tools than in previous years. When searching for a new doctor or medical professional, consumers are most concerned about convenience, cost and reputation.

Consumers are seeking clarity. The variation in hospital ratings makes it challenging to determine the most accurate representation of hospital quality; consumers are asked to decipher the difference between stars, letter grades, points and national rankings. To further complicate matters, most of the hospital report cards rely heavily on publicly available data sets, which skew toward the Medicare population. These often outdated data are frequently not generalizable, particularly for patients looking for maternity care or other services that lend themselves to prior decision-making.

**MEASURE MADNESS**

Quality measurement and reporting are critical to improving patient care, outcomes and experience; however, every measure requires an investment of resources. Hospitals and other providers face a staggering number of demands for data from a growing number of stakeholders. Government and commercial payers, accreditation agencies, professional societies, registries and other organizations require hundreds of measures. These measures have different purposes and audiences, including process/quality improvement, data for consumers and billing.

Often, different measures are intended to evaluate the same focus area or population but have different specifications. Even the smallest specification changes often require providers to implement new approaches for data collection, abstracting, electronic health record configuration and reporting. This constant tweaking diverts significant resources, with little or no added value.

Important opportunities to make meaningful enhancements in quality and patient safety may be lost as a result of the proliferation of measurement and the limitations of current EHR technology. Undue time spent on measurement takes caregivers away from patient care and addressing more meaningful quality and patient safety priorities.

As healthcare continues to move toward value-based payment and away from payment for volume, measure madness also brings financial risk. Provider performance is being evaluated in the public and private sectors, with significant consequences for the organization’s
bottom line. This scrutiny drives efforts to improve report card grades, rather than performance improvement. Providers must now consider their ability to improve performance when selecting measures for payer contracts.

Organizations of all types and sizes are impacted by measure madness. Just as large hospitals are challenged by many competing demands, smaller health systems face similar difficulties, often with fewer supports and infrastructure to accommodate extensive quality reporting obligations. Clinicians in these health systems frequently serve in multiple roles, including data collector, reporter, analyzer, information technology specialist and improvement coordinator, and often have additional administrative or clinical responsibilities.
INSUFFICIENT PROGRESS HAS BEEN MADE

Since the publication of HANYS’ 2013 Report on Report Cards, some progress has been made to address the measure madness that plagues hospitals and other healthcare providers. However, more must be done. HANYS continues to advocate for measurement that is streamlined, aligned and focused on the measures that matter most for patient care. In recent years, HANYS has been a strong voice on the Medicare inpatient, outpatient, physician and post-acute quality reporting programs; the Centers for Medicare and Medicaid Services’ Overall Star Ratings; Hospital Consumer Assessment of Healthcare Providers and Systems modernization; Medicare social risk factors; and new directions for the Center for Medicare and Medicaid Innovation.

• In 2016, HANYS published Measures that Matter, which detailed the chaotic state of healthcare reporting and measurement, provided a clear call to action and outlined a vision for the future of quality measurement where stakeholders work together to streamline, align and focus on the measures that matter most for improving patient care and outcomes. This comprehensive report also offered multiple measure management strategies and tools for healthcare providers.

Soon after HANYS published Measures that Matter, the Core Quality Measure Collaborative, led by the America’s Health Insurance Plans and its member plans’ chief medical officers; leaders from CMS, National Quality Forum and national physician organizations; employers and consumers worked hard to reach consensus on core performance measures. Using a multi-stakeholder process, the collaborative promotes alignment and harmonization of measure use and collection across payers in both the public and private sectors. Designed to be meaningful to patients, consumers and physicians, the collaborative developed core measures in eight areas. Unfortunately, these consensus measure sets aren’t being used and new measures continue to be developed and adopted. The group is expected to reconvene, but may not be able to stem the tide of measurement.

UNFORTUNATELY, THESE CONSENSUS MEASURE SETS Aren’T BEING USED AND NEW MEASURES CONTINUE TO BE DEVELOPED AND ADOPTED.

• In 2018, CMS launched Patients Over Paperwork, directing federal agencies to “cut the red tape” to reduce burdensome regulations. Through this initiative, CMS established an internal process to evaluate and streamline regulations with a goal to reduce unnecessary burden, increase efficiency and improve the beneficiary experience. In carrying out this internal process, CMS hopes to diminish regulatory obstacles that divert providers away from patient care.

• Also in 2018, CMS launched its Meaningful Measures initiative, which identifies the highest priorities for quality measurement and improvement. The goal is to assess those core issues that are the most critical to providing high-quality care and improving individual outcomes. The Meaningful Measures areas connect CMS’ strategic goals and individual measures/initiatives.

• National associations and specialty societies have also weighed in. The American Hospital Association urged CMS to use only measures that truly matter and “to work with a variety of stakeholders to identify what the critical indicators of quality and safety are that would be useful in giving patients an accurate sense of the
The American Academy of Hospice and Palliative Medicine released *Measuring What Matters*, a consensus recommendation for a portfolio of performance measures for all hospice and palliative care programs to use for program improvement.

- In 2016, the National Quality Forum issued a report that identified 1,367 quality measures used by 48 state and regional programs. Only 509 were distinct; the remaining 800 measures overlapped or had similar focus or variations in specifications. NQF noted that “slightly different versions of the same measure contribute to waste through reporting burden for providers and make performance comparisons more difficult.”
- In 2019, the *NEJM Catalyst* published “Rating the Raters,” which noted that “the numerous currently available public hospital quality rating systems frequently offer conflicting results, which may mislead stakeholders relying on the ratings to identify top-performing hospitals.” The article includes an evaluation of four nationally-known hospital report cards, assigning letter grades to each. Like HANYS’ previous studies, the authors found significant limitations in the reports’ methodologies and results.

**LIMITATIONS OF EHR TECHNOLOGY**

Nationwide, the EHR and health data infrastructure is characterized by a variety of “different systems with limited interoperability, disparate levels of use and approaches to use based on local factors and needs.” Many of these problems stem from vendors’ attempts to develop customer-friendly products by allowing each facility significant customization. However, customization inhibits interoperability and can exacerbate the problem of fragmented and conflicting measures across organizations. EHR documentation to meet measure requirements also adds additional fields, checklists and alerts that providers need to complete, cluttering the medical record and preventing ready access to the most important information needed for patient care. This level of variation and additional administrative burden increases delays and introduces significant patient safety risks.

Recent EHR enhancements have begun to support real-time measurement, but these systems fall woefully short in meeting the needs of providers. Many systems cannot generate simple, reliable and actionable reports. Many measures continue to require meticulous reviews of medical records by trained professionals who otherwise would be directing their expertise to providing and improving patient care processes and outcomes. In addition, providers must dedicate resources to the validation of data for claims-based measures. These coding checks add no value to the care already delivered but are an essential step in ensuring accurate public reporting and payment for services rendered. In short, providers are still spending too much time addressing EHR requirements at the expense of more time with patients.

**“SIGNIFICANT OPPORTUNITY COSTS ARE ENTAILED IN DEVOTING RESOURCES TO INEFFICIENT, REDUNDANT OR POORLY SPECIFIED MEASUREMENT ACTIVITIES, WHICH CAN DISPLACE OTHER VALUABLE OPPORTUNITIES TO IMPROVE HEALTH AND HEALTHCARE.”**

The 2015 National Academies report, *Vital Signs*, which aimed to target and align measurement efforts in the United States, recognized that EHRs are a critical part of the solution to reduce the burden on providers and help measurement systems become more effective. The report also states that more changes are needed to move toward complete interoperability among providers. Until then, staff will continue the important but arduous process of manually pulling data from medical charts, consuming and diverting an organization’s critical clinical resources. Importantly, “significant opportunity costs are entailed in devoting resources to inefficient, redundant or poorly specified measurement activities, which can displace other valuable opportunities to improve health and healthcare.”

**MEASUREMENT ENVIRONMENT**

The *Report on Report Cards* provides an overview of measurement efforts in the United States and examines how these efforts can improve health and healthcare.
Healthcare leaders have called for a fundamental rethinking of the purpose of EHRs to create the dynamic, patient-focused and patient-facing tool needed in the current healthcare landscape. Michael Dowling, president and CEO of Northwell Health, urged the industry to think bigger. “Beyond giving nurses and physicians information about medical diagnoses, it’s critically important that whatever tools we rely on shed light on patients’ personal circumstances. By illuminating social, environmental and lifestyle factors that are influencing patients’ overall health, a more-comprehensive EHR would arm clinicians with the knowledge to better address some of the underlying issues that are contributing to chronic medical conditions,” Mr. Dowling wrote in a Becker’s Hospital Review essay.15

Also, the lack of a national health information exchange limits the ability of hospitals and health systems to have real-time data for improvement.

**NEW CHALLENGES**

While progress is being made, new challenges have arisen.

**Shift to eCQMs**

Electronic clinical quality measures use data extracted from EHRs and/or health information systems to measure the quality of healthcare provided. CMS uses eCQMs in a variety of quality reporting and value-based purchasing programs. eCQMs are also used in reporting to accrediting bodies and commercial insurance payers in programs that reimburse providers based on quality reporting.19

Theoretically, the shift to eCQMs is directionally correct; however, the technology has not yet realized the vision of seamless quality reporting. eCQMs are challenging because submissions are gathered only through what the clinician electronically documents within the EHR’s structured data field at the point of care. Extensive resources are spent on getting providers to document in the structured fields to pass a measure. This also adds to increased clicks and checklists that take away from an accurate narrative of the patient story. Passing or failing a core measure is based purely on automation — a human-free, technology-to-technology submission without the benefit of chart abstractors for testing.
As noted in a Nuance white paper, “Although EHR vendors are required by law to meet eCQM requirements, it’s common to find that these systems lack proper formatting capabilities. Most EHRs were developed to capture documentation for reimbursement and, in many instances, physician satisfaction. There’s great risk that these systems are not sufficient for the full automation of the eCQM data extraction process. Without proper testing, it is difficult to substantiate that the data submission is meeting requirements.”20

In short, eCQM data may not accurately reflect the quality of care delivered. In addition, eCQM data are costly because hospitals must implement processes to continuously validate and improve the quality of eCQM data.

**Proliferation of social media ratings**

In recent years, social media ratings of hospitals and healthcare providers have exploded. Consumers are increasingly seeking feedback through Facebook, Yelp, Angie’s List and other sites. Hospitals and physicians often struggle to monitor and respond to this online feedback in a meaningful way. Still, many hospitals and health systems have dedicated resources to monitoring and responding to social media posts because of the strong reputational impact.

The healthcare community has serious concerns about the proliferation of social media ratings. There is little detail available about each site’s methodology, making it difficult to determine how to improve. Users can be anonymous. There is a strong potential for selection bias — those consumers with the best and the worst experiences are more likely to post their reviews.

Most importantly, healthcare visits are much more complex and patient-specific than a visit to a restaurant or hotel. Pre-existing conditions, personal relationships with physicians and hospital staff, other patient emergencies and myriad other factors can impact the patient experience of care. Yet, high-level, subjective social media reviews aren’t designed to deliver this level of detail. Alarmingly, there have been reports of physicians being targeted with hundreds of unfavorable ratings for their personal stances on issues such as vaccination.21

Some research suggests that consumers are sharing valid insights on social media. Studies have found that patients’ informal comments on social media sites help predict a hospital’s formal measures of patient experience on the HCAHPS survey’s “Overall Hospital Rating and Willingness to Recommend the Hospital.”22 Comments on social media may also serve as an early snapshot of patient-reported experiences, alerting administrators to problems that may appear in subsequent HCAHPS survey results.23

However, other studies have found that crowdsourcing/social media sites provide less consistent information when it comes to risk-adjusted measures of patient safety and clinical quality.24 There are also concerns about trolling in social media, which can produce inaccurate or false information.
The composite craze

In an effort to provide simple information to consumers, many report card organizations use composite measures—a single measure that combines information on individual measures into a single score. As an example, CMS now distills up to 60 measures of inpatient and outpatient hospital care into a single Overall Star Rating. However, these measures reflect very different populations, time periods and focus areas.

The composite craze has gone so far as to result in a composite of report cards—Becker’s Healthcare publishes a list of 100 great hospitals that is based on an analysis of other ranking and award agencies, including U.S. News & World Report’s rankings, CareChex, Healthgrades, CMS Star Ratings, Leapfrog grades and IBM Watson Health top hospitals.

Despite their now widespread use, composite measures are controversial and often problematic. A recent paper in BMJ Quality and Safety identified six common problems associated with composite indicators that seek to summarize hospital quality or safety. The authors noted that many current composite indicators “suffer from statistical flaws that greatly limit their usefulness...” and that “much greater transparency is needed.”

The researchers urge measure developers to ensure that composite indicators be “designed in accordance with good clinical practices. Underlying measures should, at a minimum, be appropriately adjusted for case mix, assessed for possible sources of bias and meet basic standards of interunit reliability. The reasons for missing data should be explored and principled approaches should be adopted to address missing data.”

The hospital industry finds Patient Safety Indicator 90 to be especially problematic. The measure is a patient safety and adverse event composite designed to provide an overview of hospital-level quality as it relates to a set of potentially preventable hospital-related events associated with harmful outcomes for patients. PSI 90 is claims-based—meaning the data are derived from billing records and have limited utility for quality improvement. It is also highly focused on surgical measures, which do not apply to many hospital patients. In addition, many concerns have been raised about the measure, given the lack of validity of the underlying components and high rates of misclassification compared to chart review.
These concerns are rooted in evidence. For example, Hefner et. al. conducted a retrospective analysis of all PSIs flagged in a fiscal year at a six-hospital academic medical center and found that of 657 PSI flags, 185 were reversed because of algorithm limitations, coding misinterpretations, present upon admission and documentation insufficiency. If, despite poor validity, U.S. policy continues to rely on PSIs for incentive and penalty programs, improvements are needed in the quality of administrative data and the standardization of PSI algorithms,” Hefner et. al. concluded.

In creating composite measures, research by Shwartz et. al. finds the major issues are as follows:

1. whether to aggregate measures at the patient level through all-or-none approaches or the facility level, using one of the several possible weighting schemes;
2. when combining measures on different scales, how to rescale measures (using z scores, range percentages, ranks or 5-star categorizations); and
3. whether to use shrinkage estimators, which increase precision by smoothing rates from smaller facilities but also decrease transparency.

The decisions made in each of these areas have a strong impact on the overall rating earned by each organization.

Measuring quality across the continuum
Healthcare organizations are growing larger by the day as the result of mergers, acquisitions and other arrangements. Consolidation has increased in an effort to achieve economies of scale, reduce clinical variation, manage population health and increase access to capital.

In the early stages of these new arrangements, measurement can be a challenge. To drive results in an effective and efficient way, the newly formed partners must negotiate and reach agreement on a standard set of measures that will apply to most (or all) facilities in the system. However, sometimes a health system may want to maintain some level of autonomy for the individual facilities to stimulate innovation and best practice sharing. Empowering local institutions while also maintaining alignment with larger health system strategic goals can drain limited resources.

Managing care across the continuum means commitment to the Triple Aim of better care, better health and lower cost. Quality professionals have a keen focus on access to care in the appropriate setting and reducing unnecessary care. Federally funded innovation models support new payment models that health systems are learning and trying to master. The models rely on primary care providers integrated with specialties such as behavioral health and strong social services. Healthcare leader training on population health is becoming the norm and systems are attempting to integrate quality efforts across the continuum to align performance incentives.
Incorporating population health

When it comes to population health, measurement can take many forms. An ideal population health outcome metric should reflect a population’s dynamic state of physical, mental and social well-being. Positive health outcomes include being alive; functioning well mentally, physically and socially; and having a sense of well-being. Negative outcomes include death, loss of function and lack of well-being. In contrast to these health outcomes, diseases and injuries are intermediate factors that influence the likelihood of achieving a state of health.35

Rather than simple mortality measures, a more appropriate indicator may be something like “quality-adjusted life years.” Patient-reported outcome measure tools such as the Patient-Reported Outcomes Measurement Information System (PROMIS)36 gather various aspects of quality of life directly from patients. Report cards should better emphasize broader outcomes, shifting their focus to how a clinical population might sensibly do over time, rather than a more complex and detailed, measure-by-measure approach to quality reporting that may not be pertinent to the patient seeking care.

In New York, healthcare organizations engaged in population health efforts reference the Prevention Agenda, the state's health improvement plan, which serves as a blueprint for improving the health and well-being of all New Yorkers and promoting health equity in all populations who experience disparities. Priorities include preventing chronic diseases, healthy eating and food security, physical activity, indoor and outdoor environmental safety, maternal and women’s health, child and adolescent health and preventing mental and substance use disorders.37

Also in New York, hospitals, health systems and community-based organizations have been engaged in the Delivery System Reform Incentive Program, which is focused on transforming the way care is delivered to Medicaid patients.38 As part of DSRIP, Performing Provider Systems are working together to improve population health through integration of physical and behavioral health services, care management and other strategies. PPSs and the state face significant financial consequences related to performance on a defined set of measures.

The list of potential measures is endless; it requires discipline to choose a limited set of indicators to best determine the impact on the health of the broader community. This measure chaos intensifies because while policymakers are clamoring about social determinants of health,39 healthcare leaders and policymakers are still building their expertise in this area. Together, we are working to identify interventions that have a positive impact on patient outcomes, the appropriate measurement strategies and effective payment structures to propel these efforts forward.
The challenge of commercial payer quality incentives

In addition to publicly reported quality measures, hospitals and other healthcare organizations are held to performance standards through contracts with commercial payers. These programs vary by payer and can include more than 50 measures and standards extracted from other groups, including CMS, NQF, The Joint Commission and other specialty accreditation bodies. Many of these measures are drawn from the Healthcare Effectiveness Data and Information Set, which includes more than 90 measures across six domains:

- effectiveness of care;
- access/availability of care;
- experience of care;
- utilization and risk-adjusted utilization;
- health plan descriptive information; and
- measures collected using electronic clinical data systems.

The HEDIS measures are generally aligned with the required measures under the federal Hospital Inpatient and Outpatient Quality Reporting Programs. However, a hospital must choose measures wisely when drafting contracts and must continuously monitor performance to ensure it meets its goals, earns incentives or avoids penalties. An organization may focus on areas that provide the best opportunity to improve and earn the financial bonus. These choices may or may not match current state and federal reporting mandates or quality improvement priorities.

As noted by the Hospital Financial Management Association, “Years ago, contracting primarily was left to the contracting department. Yet, since the advent of value-based payment contracts, healthcare organizations increasingly have relied on multidisciplinary teams to help shape the design of these programs, including how quality and financial incentives are aligned.”

To be successful, organizations must make contracts actionable by having a strong sense of: the proposed attribution model, opportunities to mitigate risk, measurement periods and how baselines and targets are calculated. From there, organizations should involve physician leaders, develop a preferred list of measures and appeal to providers through incentives. This is very difficult work and the stakes are high.

Scarce data for the broader population

Despite these new demands for measurement and growing financial incentives, comprehensive utilization and outcome data are lacking for most patients.

Medicare data, largely reflecting patients over age 65, are the most widely available. Medicare accounts for 20% of healthcare spending and the federal government has a record of every payment made through the program. By contrast, payments under the other major public healthcare program, Medicaid, are made by individual states, so there is no equivalent central database. The remaining source of information about spending is from private insurance claims, which comprise 33% of U.S. spending. Private insurance data have historically been the most challenging to access because insurers consider the information proprietary and, like Medicaid data, they aren’t centrally collected.

Research shows that Medicare and private healthcare spending are very different; Medicare and private spending are not highly correlated and spending drivers are very different in the two markets. Commercial payers, report card agencies, researchers and others rely on the convenience of the Medicare data set. As a result of mixing clean measures with this not-always-representative population, the conclusions about quality, cost and utilization remain muddy at best.
HANYS supports the availability of hospital quality and safety information to help patients make choices and assist providers in improving care. However, the information must be based on a standard set of measures that have been proven to be valid, reliable and evidence-based.

We envision a future in which consumers have access to information that is meaningful, accurate, reliable and relevant to their unique healthcare circumstances and feel empowered to make informed choices about their care. We also support reliable, nationally vetted measures of patient-reported outcomes to better engage consumers in their care.

To achieve this vision, HANYS strongly calls on all stakeholders to reduce the number of report cards.

We also urge all publicly available report cards to provide an accurate representation of the quality of hospital care, using the best available data. The data shared in these reports should be:

- transparent;
- evidence-based;
- aligned with national measures;
- rooted in clinical care;
- timely;
- risk-adjusted;
- valid and reliable;
- from consistent timeframes and sources;
- available to providers before publication to correct errors;
- free from conflicts of interest; and
- representative of the population.
INFORMATION MUST BE BASED ON A STANDARD SET OF MEASURES THAT HAVE BEEN PROVEN TO ACHIEVE THIS VISION.

HANYS and our members are committed to collaborating with consumers, the healthcare field, payers and government to make this vision a reality, knowing that patients across the country depend on providers to use metrics that drive excellence, innovation, quality improvement and patient safety.

Doing so will enable healthcare providers to achieve the future state outlined in HANYS' Measures that Matter:

- Measures will reflect “clinical reality” by accurately measuring the intended target and be actionable by providers who can use the data to implement evidence-based practices to improve care.
- The number of reported measures required of providers by payers (government and commercial) and other entities will be consistent, align with one another using standardized definitions and represent only the most important health priorities.
- The data acquisition and reporting process will “no longer [distract] from the process of care nor [require] extra effort” and will be embedded seamlessly in integrated, interoperable EHRs, allowing for more comprehensive measurement.
- Providers will focus their quality and patient safety efforts on their most serious safety concerns and prioritize time and resources to improve care with a goal of zero harm.
HANYS URGES THAT PUBLICLY AVAILABLE CONSUMER REPORT CARDS ADHERE TO THE FOLLOWING GUIDING PRINCIPLES:

1 TRANSPARENT METHODOLOGY
The complete methodology is available, enabling hospitals to replicate the results and analyze the data. The methodology also clarifies the circumstances under which hospitals are excluded from the report card. Report cards that are generated from proprietary blinded calculations, commonly known as “black box” methodologies, limit the degree to which hospitals or others can use the information or ensure that it is a fair representation of practices. The methodology should also clarify the circumstances under which hospitals are excluded from the report card.

2 EVIDENCE-BASED MEASURES
Measures must be rooted in science and supported by peer-reviewed literature. Measures must be evidence-based and accurately reflect the quality of healthcare delivered.

3 MEASURE ALIGNMENT
The quality measures are endorsed by NQF and the Measure Application Partnership and/or aligned with CMS or other national government-based or accrediting organizations.

4 APPROPRIATE DATA SOURCE
Evidence-based clinical data obtained through medical chart abstraction or from a national quality performance registry are used. The report is not based on administrative data.

Administrative data are collected for billing purposes, rather than for the evaluation of performance, and have significant limitations. While administrative data are considered an inexpensive and easy-to-access alternative for certain outcome measures such as mortality, for which the coding patterns are relatively consistent across healthcare providers, other measures drawn from administrative data have significant limitations and are susceptible to variations in hospital or regional coding practices. HANYS is particularly concerned about measures that come from voluntarily reported survey data that have not undergone appropriate validity testing.

5 MOST CURRENT DATA
The data used to generate the report are no more than one year old from the release of the report. Unfortunately, the current state of the quality measurement infrastructure typically results in a one-year lag or more for the public release of data.

As hospitals are engaged in aggressive quality campaigns, including programs such as the federal Partnership for Patients, their performance is continually improving. Report cards that use data that are more than one year old do not provide an up-to-date picture of the care delivered at a particular hospital. In the future, as EHRs evolve and become more prevalent, HANYS anticipates that more current data will be available.

HANYS IS PARTICULARLY CONCERNED ABOUT MEASURES THAT COME FROM VOLUNTARILY REPORTED SURVEY DATA THAT HAVE NOT UNDERGONE APPROPRIATE VALIDITY TESTING.
6 RISK-ADJUSTED DATA
A statistical model is applied to the data that adjusts for significant differences in patient severity of illness, demographic status and other factors that impact patient outcomes. The risk adjustment must be transparent. HANYS urges report cards to incorporate an adjustment for socioeconomic factors. Research has demonstrated that these factors impact outcomes. It is essential to make every attempt to account statistically for the wide variation among populations served by hospitals.

7 DATA VALIDITY AND RELIABILITY
The data have undergone quality and integrity edits to correct for errors in the source file and eliminate outliers that can skew the data results. Hospitals with incomplete data should be eliminated from model building and reporting.

8 CONSISTENT DATA
Comparative data points are gathered from the same sources and timeframes. Some report cards incorrectly compare data from sources with different populations and different reporting periods to generate a composite score or ranking.

9 PROVIDER ENGAGEMENT
The report card organization allows hospitals to review the report prior to its release to correct potential errors. The report card organization also gathers input from the provider community about how to improve the measures, identify unintended consequences and continuously improve the methodology.

10 CONFLICT-FREE BUSINESS MODEL
The organization publishing the ratings does not stand to profit from the release of the ratings through the sale of subscriptions, marketing fees or consulting services.

11 REPRESENTATIVE POPULATION
The report card uses data from a representative population, rather than relying solely on Medicare data, which are more widely available but do not capture information about children and most adults under age 65. We urge states to make Medicaid and commercial data more widely available for the purposes of identifying best practices and driving quality improvement.
1. A panel of eight judges completed an independent evaluation of 12 report cards. The judges assessed whether report cards met 11 agreed-upon standards using a Likert scale.

2. The Likert scale was then converted to a numerical score:
   - STRONGLY AGREE = 15 points
   - AGREE = 10 points
   - DISAGREE = 5 points
   - STRONGLY DISAGREE = 0 points
   - MISSING OR “DON’T KNOW” = not included

3. The mean score of all judge assessments for each of the 11 evaluation criteria was then averaged to create a summary score for each report card.

4. The summary score was then grouped according to the scores closest to one of the five star categories:
   - ★ 1 star = 3 points
   - ★★ 2 stars = 6 points
   - ★★★ 3 stars = 9 points
   - ★★★★ 4 stars = 12 points
   - ★★★★★ 5 stars = 15 points

Information about the methodologies was obtained from public websites and is the most current available at the time of the analysis (May 2019). Unlike the recent NEJM Catalyst report card analysis, HANYS did not engage the report card organizations in the course of our review. Similarly, HANYS did not provide the report card organizations with advance notice of their scores.

It is important to note that while many of these report card organizations generate several different reports, HANYS did not evaluate every report from each organization. The specific reports included in our evaluation are delineated in the HANYS Report Card.
Note: There are many other hospital quality report cards. Some are limited in value, reliability and reach. For the purposes of this report, HANYS focused only on the report cards listed above. New York reports may not be applicable to readers outside of the state.

Report cards are displayed in descending order within each star category, according to HANYS’ evaluation.

Ratings are based on information available as of May 15, 2019.

<table>
<thead>
<tr>
<th>STARS</th>
<th>REPORT CARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>★★★★★</td>
<td>NYSDOH Clinical Reports</td>
</tr>
<tr>
<td>★★★★★</td>
<td>NYSDOH Hospital Profiles</td>
</tr>
<tr>
<td>★★★★★</td>
<td>CMS Hospital Compare</td>
</tr>
<tr>
<td>★★★★★</td>
<td>The Joint Commission Quality Check</td>
</tr>
<tr>
<td>★★★★★</td>
<td>CMS Hospital Compare overall hospital ratings</td>
</tr>
<tr>
<td>★★★★★</td>
<td>Vizient Quality Leadership Award</td>
</tr>
<tr>
<td>★★★★</td>
<td>The Leapfrog Group Hospital Safety Grade</td>
</tr>
<tr>
<td>★★★★</td>
<td>Healthgrades America’s Best Hospitals</td>
</tr>
<tr>
<td>★★★★</td>
<td>IBM Watson Health 100 Top Hospitals</td>
</tr>
<tr>
<td>★★★★</td>
<td><em>U.S. News and World Report</em> Best Hospitals by Specialty</td>
</tr>
<tr>
<td>★★★★</td>
<td><em>U.S. News and World Report</em> Best Hospitals for Procedures and Conditions</td>
</tr>
<tr>
<td>★★★</td>
<td><em>Newsweek’s World’s Best Hospitals</em></td>
</tr>
</tbody>
</table>
HANYS’ evaluation of 12 popular hospital report cards, based on our guiding principles, revealed wide variation in the methodologies and results.

In general, government and accrediting organizations are more successful than other report card organizations in meeting HANYS’ criteria for evaluating hospital performance. The report cards receiving lower scores generally relied heavily on administrative claims data and/or unvalidated survey data; gathered comparative data points from different sources and timeframes to generate a composite score or ranking; and/or did not use measures aligned with NQF, CMS or national accrediting organizations.

HANYS hopes that as more clinical data are made available through EHRs and other means, and as widely-used risk-adjustment methodologies are further refined to include socioeconomic factors, the report card organizations will update their methodologies to generate a more accurate evaluation of hospital quality.

Importantly, HANYS strongly urges the healthcare field to reduce the number of report cards so that information is less confusing for consumers and providers can focus on improving patient care.
## DISCUSSION OF KEY FINDINGS

### 1. Transparent Methodology
Almost all of the report cards posted information regarding their methodologies on public websites. However, some report cards provided more details regarding their methodology than others.

### 2. Evidence-Based Measures
The report cards use a combination of structure, process and outcome measures. Many of these measures are evidence-based and reflect the quality of healthcare delivered. However, some of the report cards use surveys, which measure a subjective perception and/or collect self-reported information that is not validated.

### 3. Measure Alignment
Many of the quality measures used in the report cards are endorsed by NQF and/or are aligned with CMS. However, some report cards use these measures to generate new composite scores.

### 4. Appropriate Data Source
Few report cards use evidence-based clinical data obtained through medical chart abstraction or from a national quality performance registry. Instead, they rely heavily on administrative data that are collected for billing purposes and have significant limitations when trying to evaluate performance. Others rely on reputation surveys, which are a lagging indicator and lack reliability and evidence-base.

### 5. Most Current Data
There is variation in the timeliness of data. Some measures are from within the past year; other measures use data that are between one and two years old at the time of the report card’s release. Some measures, by their design, use data that span several years (e.g., readmissions).

### 6. Risk-Adjusted Data
The majority of the report cards satisfied this criterion by using risk-adjusted data from CMS or the Agency for Healthcare Research and Quality or by conducting their own risk adjustment. However, these data are often combined with data that are not risk-adjusted. Additional opportunities exist to improve risk adjustment to address sociodemographic factors that impact patient outcomes.

### 7. Data Validity and Reliability
All of the report cards use data that are edited before publication to correct for errors and remove outliers that may skew the results. Report cards also often exclude hospitals that do not meet minimum thresholds for measures.
8 **CONSISTENT DATA**
Comparative data points are not always gathered from the same sources and timeframes. Some report cards incorrectly compare data from sources with different populations and different reporting periods to generate a composite score or ranking.

9 **PROVIDER ENGAGEMENT**
Many report card organizations allow hospitals to review the report prior to its release to correct potential errors. Some gather input from the provider community about how to improve their measures, identify unintended consequences and continuously improve their methodology.

10 **CONFLICT-FREE BUSINESS MODEL**
Some organizations stand to profit from the release of the ratings through the sale of subscriptions, marketing fees, licenses and/or consulting services, which creates a significant conflict of interest. The profit motive also creates an incentive for the publishing organization to identify more differences as meaningful than may actually exist. Other organizations use pressure tactics to encourage participation.

11 **REPRESENTATIVE POPULATION**
Most report cards use data solely from Medicare or Medicaid, which is more widely available than data that reflect the broader healthcare population.
CMS Hospital Compare

Hospital Compare has information about the quality of care at more than 4,000 Medicare-certified hospitals across the country, including over 130 Veterans Administration medical centers. Hospital Compare was created through the efforts of CMS in collaboration with organizations representing consumers, hospitals, doctors, employers, accrediting organizations and other federal agencies.

CMS Hospital Compare Overall Hospital Ratings

CMS also publishes Overall Star Ratings. The overall hospital rating summarizes a variety of inpatient and outpatient measures on Hospital Compare reflecting common conditions that hospitals treat. Hospitals may perform more complex services or procedures not reflected in the measures on Hospital Compare. The overall hospital rating shows how well each hospital performed, on average, compared to other hospitals in the U.S. The overall hospital rating ranges from 1 to 5 stars. The more stars, the better a hospital performed on the available quality measures. The most common overall hospital rating is 3 stars.

The Joint Commission Quality Check

The Joint Commission is a nonprofit organization that accredits and certifies more than 21,000 healthcare organizations and programs in the United States. Quality Check allows consumers to search for accredited and certified organizations and download free hospital performance measure results.

NYSDOH Hospital Quality Profile

NYSDOH publishes a set of metrics that look at quality, safety and inspections. Data are drawn from a number of sources, including NYSDOH clinical programs, the Statewide Planning and Research Cooperative System and the Quality Improvement Organization Clinical Warehouse, the national data repository for private healthcare data.

NYSDOH Clinical Reports

NYSDOH publishes a number of statewide reports focused on key clinical issues. The data are clinically abstracted and reported to NYSDOH by New York state hospitals. Reports are published on an annual basis and include comparisons with the state average.

The Leapfrog Group Hospital Safety Grade

The Leapfrog Group is a nonprofit organization that represents employers and insurance purchasers. The Leapfrog Hospital Safety Grade uses performance measures from CMS, the voluntary Leapfrog Hospital Survey, AHRQ, the Centers for Disease Control and Prevention and the American Hospital Association’s Annual Survey and Health Information Technology Supplement. Nearly 2,000 hospitals voluntarily participate in the survey; however, Leapfrog publishes letter grades (ranging from A to F) for more than 2,600 general acute-care hospitals across the country twice annually.
HEALTHGRADES AMERICA’S BEST HOSPITALS
healthgrades.com/quality/hospital-ratings-awards
The Healthgrades America’s Best Hospitals™ achievements are based on a review of clinical outcomes across multiple conditions and procedures, analyzing the performance of 4,500 hospitals nationwide. Healthgrades uses Medicare inpatient data from the Medicare Provider Analysis and Review (MedPAR) file purchased from CMS. Patient outcomes data for 32 conditions or procedures are analyzed; hospitals that perform in the top 5% of all facilities receive the distinction of America’s 250 Best Hospitals.

U.S. NEWS AND WORLD REPORT BEST HOSPITALS BY SPECIALTY
health.usnews.com/best-hospitals/rankings
U.S. News and World Report analyzes data from nearly 5,000 medical centers and survey responses from more than 30,000 physicians to rank hospitals in 16 adult specialties, including cancer, diabetes, rheumatology and more. Survival rates, patient safety, specialized staff and hospital reputation are among the factors weighed. Nationally, only 158 hospitals ranked in at least one of the specialties in 2018-2019.

U.S. NEWS AND WORLD REPORT BEST HOSPITALS FOR PROCEDURES AND CONDITIONS
health.usnews.com/best-hospitals
U.S. News and World Report Best Hospitals for Procedures and Conditions (previously called “Best Hospitals for Common Care”) evaluates hospital performance for nine procedures and conditions: abdominal aortic aneurysm repair, aortic valve surgery, chronic obstructive pulmonary disease, colon cancer surgery, congestive heart failure, heart bypass surgery, hip replacement, knee replacement and lung cancer surgery. Data are obtained from the Standard Analytic File, publicly reported data from CMS, the American Hospital Association’s annual survey, the American Nurses Credentialing Center and the Society of Thoracic Surgeons.

IBM WATSON HEALTH 100 TOP HOSPITALS
ibm.com/watson-health/services/100-top
The Watson Health 100 Hospitals study, formerly the Truven Health Analytics® study, identifies 100 top-performing hospitals based on publicly available data and an examination of clinical, operational, financial and patient perception of care metrics. The 100 Top Hospitals study categorizes the nation’s hospitals into five groups: major teaching, teaching, large community, medium community and small community hospitals.

VIZIENT QUALITY LEADERSHIP AWARD
vizientinc.com/Members/Member-awards
The Bernard A. Birnbaum, MD, Quality Leadership Award is given to members from top-performing academic medical centers, complex teaching medical centers and community hospitals that demonstrate excellence in delivering high-quality care based on the measures in the Vizient Quality and Accountability study. The measures include safety, mortality, clinical effectiveness, efficiency, patient centeredness and equity of care. The award also considers performance in core measures data, the HCAHPS survey and the CDC’s National Healthcare Safety Network.

NEWSWEEK’S WORLD’S BEST HOSPITALS
newsweek.com/best-hospitals-2019
Newsweek partners with Statista, Inc., a global market research and consumer data company, to rank hospitals around the world. In 2019, 1,000 hospitals were selected based on recommendations from medical professionals, patient survey results and medical performance indicators. For the recommendations, Statista collaborated with GeoBlue and invited tens of thousands of doctors, hospital managers and other healthcare professionals to an online survey. The World’s Best Hospitals list can be filtered by country.
HANYS extends our sincere appreciation to the members of our Statewide Steering Committee on Quality Initiatives and our Report on Report Cards Working Group for their input and guidance during the development of this document. We thank them for their leadership and participation in the national dialogue on improving quality and patient safety through reasonable healthcare measurement.

Report on Report Cards Working Group

Robert J. Panzer, MD  
Associate Vice President, Patient Care Quality and Safety and Chief Quality Officer  
University of Rochester Medical Center

Alison L. Hong, MD  
Chief Quality Officer  
St. Peter’s Health Partners

Mary Ellen Crittenden, RN, MS, CPHQ, CPPS  
Vice President, Quality and Patient Safety Officer  
Greater Hudson Valley Health System

Mark Jarrett, MD, MBA, MS  
Senior Vice President, Chief Quality Officer and Associate Chief Medical Officer  
Northwell Health

Brian J. D’Arcy, MD  
Senior Vice President, Medical Affairs  
Catholic Health System

Martha J. Radford, MD, FACC, FAHA  
Senior Clinical Quality Advisor  
NYU Langone Health

David Evelyn, MD, MPH  
Vice President, Medical Affairs  
Cayuga Medical Center at Ithaca

Brian K. Regan, PhD  
Vice President, Quality and Patient Safety  
NewYork-Presbyterian

Jason M. Golbin, DO, MBA, MS, FCCP, FACP, FACHE, FACMQ  
System Chief Quality Officer Senior Vice President  
Catholic Health Services of Long Island
<table>
<thead>
<tr>
<th>Name</th>
<th>Title and Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anu Banerjee, PhD, MS, MHIM, FACHE</td>
<td>System Vice President and Chief Quality and Innovation Officer, Arnot Health</td>
</tr>
<tr>
<td>Ashima Butler</td>
<td>Vice President and Chief Operating Officer, Ellenville Regional Hospital</td>
</tr>
<tr>
<td>Hans Cassagnol, MD</td>
<td>Executive Vice President and Chief Physician, Medical Administration, Catholic Health System</td>
</tr>
<tr>
<td>Mary Ellen Crittenden, RN, MS, CPHQ, CPPS</td>
<td>Vice President, Quality and Patient Safety Officer, Greater Hudson Valley Health System</td>
</tr>
<tr>
<td>Christine Cunningham, RN, BSN, MPA</td>
<td>Director, Clinical Quality, NYU Langone Health</td>
</tr>
<tr>
<td>Brian J. D’Arcy, MD</td>
<td>Senior Vice President, Medical Affairs, Catholic Health System</td>
</tr>
<tr>
<td>Rajesh Dave, Sr., MD</td>
<td>Executive Vice President and Chief Medical Officer, United Health Services Hospitals, Inc.</td>
</tr>
<tr>
<td>Yosef Dlugacz, PhD</td>
<td>Senior Vice President and Chief, Clinical Quality, Education and Research, Northwell Health</td>
</tr>
<tr>
<td>David Evelyn, MD, MPH</td>
<td>Vice President, Medical Affairs, Cayuga Medical Center at Ithaca</td>
</tr>
<tr>
<td>Lawrence Faltz, MD</td>
<td>Chief Medical Officer, Phelps Hospital, United Health Services Hospitals, Inc.</td>
</tr>
<tr>
<td>Candace Finkelstein</td>
<td>Vice President, Quality and Regulatory Officer, Mount Sinai Health System</td>
</tr>
<tr>
<td>Anna Gaeta, RN, BSN, MS, CPHQ, CPPS</td>
<td>Associate Vice President, Quality Support Services, Saratoga Hospital</td>
</tr>
<tr>
<td>Renee Garrick, MD, FACT</td>
<td>Chief Medical Officer, Westchester Medical Center</td>
</tr>
<tr>
<td>Paul A. Gitman, MD, MACP</td>
<td>Quality Advisor, Westchester Medical Center</td>
</tr>
<tr>
<td>Jason M. Golbin, DO, MBA, MS, FCCP, FACP,</td>
<td>System Chief Quality Officer and Senior Vice President, Catholic Health Services of Long Island</td>
</tr>
<tr>
<td>Frederick Goldberg, MD, MHC&amp;M</td>
<td>Vice President, Medical Affairs and Chief Medical Officer, Nathan Littauer Hospital</td>
</tr>
<tr>
<td>William H. Greene, MD</td>
<td>Senior Quality Consultant, OR Dx + Rx: Solutions for Surgical Safety, Northwell Health</td>
</tr>
<tr>
<td>Kent N. Hall, MD</td>
<td>Vice President, Medical Affairs, Cheion Medical Officer, The University of Vermont Health Network–</td>
</tr>
<tr>
<td>Alison L. Hong, MD</td>
<td>Chief Quality Officer, St. Peter’s Health Partners</td>
</tr>
<tr>
<td>Mark Jarrett, MD, MBA, MS</td>
<td>Senior Vice President, Chief Quality Officer, Northwell Health</td>
</tr>
<tr>
<td>Wolf-Dieter Krahn, MD</td>
<td>Medical Director, UPMC Chautauqua</td>
</tr>
<tr>
<td>William W. LeCates, MD</td>
<td>President, Bassett Medical Center</td>
</tr>
<tr>
<td>Daniel Lombardi, DO, MBA, FACOEP</td>
<td>Vice President, Chief Quality Officer, Northwell Health</td>
</tr>
<tr>
<td>Charlene J. Ludlow, BSN, MHA, CIC</td>
<td>Chief Quality and Safety Officer, Erie County Medical Center</td>
</tr>
<tr>
<td>William Lynch</td>
<td>Executive Vice President and Chief Operating Officer, Jamaica Hospital Medical Center</td>
</tr>
<tr>
<td>Marie Maxson, RN, BS, CHCQM</td>
<td>Director, Quality Management and Clinical Effectiveness, Bassett Medical Center</td>
</tr>
<tr>
<td>Robert R. Mayo, MD</td>
<td>Chief Medical Officer and Executive Vice President, Rochester Regional Health</td>
</tr>
<tr>
<td>Ann Marie McDonald, EDD, RN, FACHE</td>
<td>Senior Director, Regulatory Affairs, Mount Sinai Health System</td>
</tr>
<tr>
<td>Maralyn Militello, RN, BSN, MPA, CPHQ</td>
<td>Senior Director, Nursing, Eastern Niagara Hospital</td>
</tr>
<tr>
<td>Marc L. Napp, MD, MS, FACS</td>
<td>Senior Vice President, Medical Affairs and Deputy Chief Medical Officer, Mount Sinai Health System</td>
</tr>
<tr>
<td>Robert J. Panzer, MD</td>
<td>Associate Vice President, Patient Care Quality and Safety and Chief Quality Officer, University of Rochester Medical Center</td>
</tr>
</tbody>
</table>
ENDNOTES


2 Ibid.

3 HANYS. 2016. Measures that Matter. hany.org/quality/clinical_operational_oversight/measures_that_matter/


13 Committee on Core Metrics for Better Health at Lower Cost; Institute of Medicine; Blumenthal D, Malphrus E, McGinnis JM, editors. Washington (DC): National Academies Press (US); 2015 April 28. nbci.nlm.nih.gov/books/NBK316121


23 Ibid.


27 Ibid.

28 Ibid.


This report is a publication of the Healthcare Association of New York State. HANYS’ focus on healthcare quality improvement helps New York’s hospitals and health systems prioritize, manage and sustain initiatives that improve patient care and outcomes.

HANYS is a leading voice with policymakers in New York state and Washington, D.C. Our quality advocacy, led by a team of clinical and policy experts in our Quality Advocacy, Research, and Innovation division, emphasizes rational, evidence-based and sustainable policies that enable providers to offer the highest quality care in the most efficient manner possible.

QARI LEADERSHIP

Loretta Willis, RN, BSN, CCM, CPHQ
Vice President

Nancy Landor, RN, MS, CPHQ
Senior Director, Strategic Quality Initiatives

Christina Miller-Foster, MPA
Senior Director, Quality Advocacy, Research, and Innovation

Kathleen Rauch, RN, MSHQS, BSN, CPHQ
Senior Director, Quality Advocacy, Research, and Innovation
If we can help you with your printing needs, call us at (855) 853-5234. Visit us at hanysprintingservices.com.