ARTICLE

Health Care Measurements that Improve Patient Outcomes

Robert S. Kaplan, PhD, MS, Lara Jehi, MD, MHCDS, Clifford Y. Ko, MD, MS, MSHS, FACS, FASCRS, Andrea Pusic, MD, MHS, FACS, FRCSC, Mary Witkowski, MD, MBA

Vol. 2 No. 2 | February 2021

DOI: 10.1056/CAT.20.0527

This article describes the challenges and solutions in determining whether a patient's treatment has been successful. Such an assessment depends on multiple factors, including the patient's pretreatment status; the qualifications of personnel performing the treatment; the treating facility's infrastructure and culture; the use of evidence-based clinical processes; the baseline incidence of treatment complications; and, most challenging, the ability to measure the outcomes that matter to patients. Recent advances in IT and the development of validated measurement instruments now enable consistent collection and analysis of metrics that capture all the relevant dimensions of a patient's treatment. These data can be mobilized for learning that improves clinical and administrative processes, optimization of care pathways, shared decision-making, accountability, and payment contracting. Providers can now have access to the tools and technology that allow them to be transparent about and accountable for the outcomes that their patients experience.

Introduction

Measurement in health care delivery has made slow progress during the 100-plus years since Dr. Ernest Amory Codman urged every hospital to "follow every patient it treats, long enough to determine whether or not the treatment has been successful." In this article, we describe recent developments in measuring patient pretreatment conditions, which include social determinants, adherence to clinical processes, incidence of treatment complications, and patient-reported outcomes that now enable universal measurement of whether a treatment has been successful to become an achievable goal.

At first glance, measuring the outcomes of health care treatments should not have taken more than a century to realize. Medicine is a science, and science requires measurement. As noted by 19th century British physicist Lord Kelvin, "When you can measure what you are speaking about and express it in numbers, you know something about it. When you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind." Measurement is also foundational for improving the management of health care delivery, as expressed in a popular quote, "If you can't measure it, you can't manage it."

Of course, health care already collects many measurements. The National Quality Forum (NQF) has helped to create "a portfolio of endorsed performance measures that can be used to measure and quantify health care processes, outcomes, patient perceptions, and organizational structure and/or systems that are associated with the ability to provide high-quality care." Yet a careful examination of the NQF metrics reveals that few meet Codman's standard of informing whether a patient's treatment has been successful. Most of the NQF metrics relate to compliance, to the processes used to treat patients and to complications arising from the treatment, or to whether the treatment had to be repeated. With the dominance of compliance and process metrics, payers and regulators hold health care professionals accountable only to surrogate metrics rather than to the results that a properly functioning health care system should deliver: better outcomes for patients achieved at lower total cost.

Admittedly, measuring health care outcomes is much more difficult than measuring the outcomes from manufacturing and most other service delivery processes. Many outcomes that matter to patients are subjective and dependent on their specific circumstances. But the increased availability of electronic health records (EHRs), sophisticated data analytics, and patient-friendly technologies now empowers clinicians to measure condition-specific outcomes, adjusted for patient-specific circumstances.

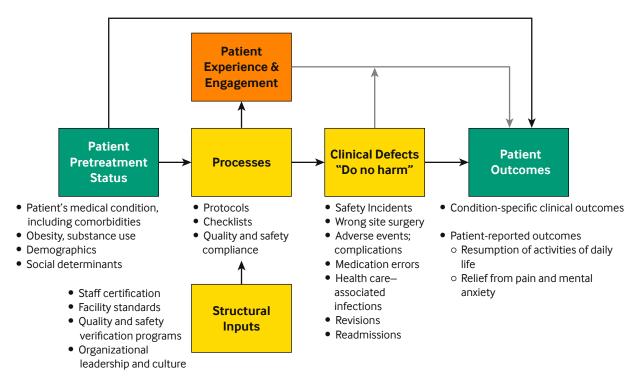
Health Care Measurement Across a Patient's Treatment Cycle

Measurement can occur at all stages of a patient's treatment cycle (Figure 1).

The treatment cycle starts with information about the patients themselves, including their medical, demographic, and socioeconomic conditions. The other inputs to the treatment process are the health care facility's clinical and management personnel, physical assets, and infrastructure, all of which are measurable. The cycle continues with metrics about the clinical and administrative processes used to treat patients, followed by clinical outcomes, usually measured by defects and adverse events from the treatment. The treatment cycle concludes with metrics reported by patients about the outcomes they experienced. These outcomes, as shown in Figure 1, are influenced by both the patients' pretreatment status and the processes used to treat them.

Measuring Performance over a Patient's Treatment Care Cycle

The patient, with multiple measures of pretreatment status, is treated by providers, with measurable characteristics of staff, equipment, facility, organizational culture, and quality and safety capabilities. The treatment consists of clinical processes and pathways, yielding both clinical outcomes, quantified by the incidence of defects and complications, and patient outcomes. Patient outcomes are also influenced by the patient's experiences during pretreatment, treatment, and recovery processes.



Source: Adapted from Porter ME. Value-based health care delivery: Core concepts. Presentation slide 12: Measure outcomes for every patient. January 15, 2020. Accessed December 19, 2020. https://www.isc.hbs.edu/Documents/pdf/2020-intro-vbhc-porter.pdf.

NEJM Catalyst (catalyst.nejm.org) $\ensuremath{\mathbb{C}}$ Massachusetts Medical Society

Measures of a Patient's Pretreatment Status

Manufacturing companies can specify the characteristics of the materials they use as production inputs. In contrast, health care professionals treat patients who have widely different characteristics, including age, obesity, substance use, disease severity, comorbidities, and demographics, including social determinants. These patient-specific characteristics, which describe the initial conditions for any health care delivery process, lead to variability in the potential outcomes that can be achieved. To properly interpret health care outcomes, leading health care institutions have begun to measure the specific characteristics of the patients they treat.



With the dominance of compliance and process metrics, payers and regulators hold health care professionals accountable only to surrogate metrics rather than to the results that a properly functioning health care system should deliver: better outcomes for patients achieved at lower total cost."

For example, new primary care patients at Cleveland Clinic complete an online survey ahead of the clinic visit or in an outpatient waiting room. In addition to standard data from the Patient-Reported Outcomes Measurement Information System (PROMIS)-10 survey instrument⁶ on patient demographics, family history, medications, and medical comorbidities, the survey includes depression screening measures from Patient Health Questionnaire-9⁷ and metrics on social determinants of health, such as financial resource strain, transportation needs, stress, intimate partner violence, housing stability, alcohol use, food insecurity, and physical activity. Cleveland Clinic integrates these patient-reported data into the patient's EHR, where they are curated by a dedicated scientific team, supported by an extensive analytics infrastructure, to risk adjust the patient's outcome measures collected during and after treatment.

Measures of Personnel and Organizational Inputs

People, equipment, and facilities are structural inputs used to treat patients. Donabedian's structure-process-outcome model⁸ asserts that a necessary condition for consistently delivering superior patient outcomes is having patients treated by qualified clinicians, using appropriate equipment, in facilities with a strong culture and infrastructure. Initially, hospital quality was measured by the qualifications of its clinical personnel, such as licensure and board certification of physicians, and nurse qualifications.

The American College of Surgeons' (ACS) Quality Verification Program introduced measurements of the structural inputs at a hospital for implementing and sustaining facility-wide surgical quality programs. Examples of such structures and capabilities are organizational management and communication, effective leadership at multiple levels, a culture of safety and high reliability, organization of care delivery through integrated practice units, standardized pathways across surgical care cycles, and the use of actionable and reliable quality metrics. The ACS developed metrics of these structural inputs for specific surgical programs, including trauma, cancer surgery, metabolic bariatric surgery, and geriatric and children's surgery. An ACS team assesses a hospital's performance on these multiple dimensions during a verification site visit to hospitals. Research has found that hospitals with structural components meeting the ACS standards are associated with higher-quality patient care. 10-12

The Center for Medicare and Medicaid Innovation, in collaboration with the ACS, is now developing a structural quality metric for the U.S. Centers for Medicare & Medicaid Services' (CMS) Bundled Payments for Care Improvement (BPCI) initiative. This would be a composite measure, calculated by aggregating six structural standard metrics.

Measuring patient-specific initial conditions and providers' structural capabilities are important advances in health care measurement. They are not, however, by themselves the answer to Codman's request for patient-centered outcome metrics.

Measures of Compliance, Conformance, and Process Excellence

U.S. manufacturing companies began to adopt a total quality management (TQM) approach in the 1970s as a competitive response to the process innovations of Japanese companies. TQM eventually reached the health care sector in the late 1980s. 13 The underlying principle of TOM was continuous improvement of processes (often described by the Japanese term kaizen). This pursuit required measurement, both to assess the current performance of processes and to quantify whether, and by how much, performance improved over time. In principle, applying the TQM approach to health care, based on satisfying the needs of customers, would have included patient outcome measures, such as restored functionality and lower levels of pain. But, as many clinicians immediately pointed out, such outcome measures failed to account for underlying differences in patients, such as their illness severity and health and socioeconomic status. At the time, the state of IT and data analytics limited both the measurement of patient outcomes and the causal linkages between patient-specific factors and patient outcomes. But rather than allowing these limitations to force abandonment of the TQM approach, health care professionals compromised by choosing to measure what they could, which was compliance and process quality. In effect, they adopted a principle often used by accountants: if you can't measure what you want, want what you can measure.

As observed by Brent James, MD,¹⁴ a leader of the quality movement in health care, a consistent and reliable care process is essential to improve quality.¹⁵ The CMS introduced 10 metrics of process quality in 2005 as part of Hospital Compare.¹⁶ The medical literature had shown that these process measures were associated with better patient outcomes. The CMS metrics became the standard for quality measurement throughout the industry, leading to widespread adoption of process metrics such as care coordination, compliance with evidence-based assessments and practices, efficiency of resource usage, readmission rates, patient and family engagement, and patient satisfaction.



Measuring patient-specific initial conditions and providers' structural capabilities are important advances in health care measurement. They are not, however, by themselves the answer to [Ernest Amory] Codman's request for patient-centered outcome metrics."

A principal goal of TQM was to reduce variation that did not lead to improved patient outcomes. Such undesirable variation occurs when clinicians do not consistently follow practices known to reduce the likelihood of adverse patient events. Peter Pronovost, MD, showed that rigorous use of a checklist reduced central line infections, then associated with more than 30,000 U.S. deaths a year, to near-zero incidence. None of the items on the checklist were new; the items simply codified procedures that doctors and nurses already knew they should be doing but were not performing

reliably on each occasion. Atul Gawande, MD,¹⁹ helped spread the message about checklists to a wide audience, including the World Health Organization.²⁰ To ensure that procedures specified in checklists were followed consistently, new conformance metrics related to checklist use were promulgated and widely adopted.

Brent James, as head of quality at Intermountain Healthcare, formed committees of doctors, nurses, and administrators to identify treatment variations for high-volume conditions. Each committee used quality metrics to detect which variations were not working. Intermountain could then measure compliance with the preferred evidence-based pathways, while allowing individual physicians to depart from the guidelines on the basis of specific patient circumstances. The teams modified the pathway when better variations emerged. Intermountain's demonstrated success led to the use, by other institutions and the CMS, of metrics on the incidence of noncompliance to evidence-based pathways. While these metrics were all desirable for improving consistency in care delivery, several studies found that they had only weak linkages to better patient outcomes. Such work would be improved by evaluating treatment variations against patient outcomes to determine which variation is actually clinically superior.

A specific variant of TQM was Lean management, based on the Toyota Production System.²¹ Lean's explicit goal was to eliminate waste, defined as any activity that did not add to the value experienced by the customer. Typical sources of waste in manufacturing were overproduction, defects, and high levels of inventory. When applied to health care, wasteful activities included defects, patient waiting times, and underutilization of personnel and equipment. Advocates of Lean management, such as Virginia Mason Hospital, introduced metrics of defect rates and cycle times for clinical and administrative processes, patient wait times, idle times for hospital personnel and equipment, and travel distance.

Lean management's disciplined focus on process improvement and process excellence certainly improved the reliability and safety of health care processes. Unlike in the production of automobiles, cameras, and appliances, however, eliminating waste and improving operational processes have been difficult to translate into stable or decreasing prices for health care services or into demonstrable improvements in outcomes that matter to patients.

Measures of Clinical Defects

Every physician takes the Hippocratic Oath to "first, do no harm." More than 2,300 years after the origin of this oath, the incidence of patient harm, unfortunately, remains far from zero. Consistent with the mantra that "if you don't measure, you can't manage," health care has introduced many metrics of the incidence of patient harm, such as safety incidents, wrong-site surgery, adverse event complications, medication errors, and health care–associated infections.

Health care has also introduced metrics such as the incidences of readmissions, re-treatments, and revision surgeries that signaled when the initial (index) treatment had been wasted by failing to adequately address the patient's condition. In these cases, the clinical defect represents both a negative outcome for the patient and an increase in treatment costs to remedy it.

The National Surgical Quality Improvement Program (NSQIP) from the ACS introduced systematic measurement and benchmarking of the risk-adjusted incidence of quality defects in surgical care. NSQIP, a nationally validated multispecialty program, started in the Veterans Affairs (VA) health system in response to a government mandate that surgical outcomes in the VA system be compared with those in the private sector for quality assurance. NSQIP helped the VA system realize significant improvements in risk-adjusted complication and mortality rates.²²



A patient seeking care for breast cancer cares not only about avoiding complications and her survival and duration of remission, but also about how she looks and feels, her level of pain, the length of the treatment cycle, and the empathy shown her during the treatment process."

The ACS subsequently modified the measurement of complication and mortality and extended it to the private sector. The program now has strict, clinically relevant, and meaningful definitions of surgical complication metrics such as thromboembolisms; reintubation; pneumonia; infections of wound, urinary tract, and central line; bleeding; and anastomotic leak. The metrics are collected for thousands of operative procedures across most surgical specialties, including general surgery, orthopedics, urology, and neurosurgery. Data collection is performed through multiple means, including direct uploads from the electronic record and by trained and annually certified abstractors. The NSQIP program is currently implemented in more than 800 hospitals across the United States and in 10 other countries.²³ Risk and reliability adjustments are performed to produce longitudinal, national benchmarking reports.

The accurate measurement of clinical defects reveals the problems that exist and whether improvement initiatives have succeeded in reducing their incidence and sustaining the reductions over time.

Measures of Patient Outcomes

Metrics on the quality of structure and inputs, compliance and conformance, process quality, and clinical defects are useful and necessary. But, collectively, they still fall short of providing evidence about the effectiveness of medical treatments. After all, patients do not seek treatment to have qualified physicians deliver standardized care with checklists and evidence-based pathways, while avoiding medical errors, defects, complications, and readmissions; patients seek care to have their condition cured or, at least, alleviated. For example, a patient seeking care for breast cancer cares not only about avoiding complications and her survival and duration of remission, but also about how she looks and feels, her level of pain, the length of the treatment cycle, and the empathy shown her during the treatment process.

We can view the differences among inputs, processes, and defect metrics versus outcome metrics through the lens of automobile design and production, the origin of the quality and Lean management movements. Let us stipulate that Toyotas today are produced perfectly, with no defects, low

maintenance and repair costs, high gasoline mileage, and exceptional longevity. The company has learned how to produce vehicles that conform to all its design specifications. Yet even a perfectly built Toyota Camry does not deliver the high-speed, comfortable, and responsive ride of an elegantly designed BMW, Audi, Mercedes, or Porsche. These vehicles deliver superior performance in addition to a reliable and consistent driving experience. Like Toyota, medicine must continue to use checklists and standardized pathways around evidence-based treatments so it can achieve better quality and avoid inflicting harm on patients. But providers need to supplement process excellence and "do no harm" metrics with outcomes metrics that quantify whether interventions have addressed patients' goals for their conditions, which include enabling them to live well without pain and limitations.

Patient-Reported Outcomes (PROs)

Historically, medical and surgical care focused only on saving lives, a relatively straightforward outcome to measure. Today, however, many health care interventions, such as knee and hip replacements, are performed not to extend survival but to improve the quality of patients' lives. The focus of measurement should similarly shift to assess whether and by how much the treatments, techniques, and technology have accomplished this goal. For example, physicians treating patients with cancer want to not only decrease mortality, but also achieve a longer life span with the least negative impact on the patient's quality of life. As an example, surgeons performing nerve-sparing prostatectomies on patients with prostate cancer strive to eliminate all malignant tissue while also preserving sexual function and urinary and bowel continence. The successful outcome to plastic surgery for breast reconstruction after a mastectomy includes a patient's confidence in her body image. Outcomes such as posttreatment physical function, relief from pain, and improved mental health are known only to patients and, hence, must be reported only by them. Achieving success with PROs should drive many treatment decisions and become an important aspect of accountability for the physician's performance.

Despite the importance of a patient's posttreatment quality of life, health care providers have been slow to adopt PROs because they considered the metrics too subjective or too "soft" to be reliable, especially when compared with clinical metrics used in the NSQIP program. But major advancements have been made over the past two decades in the field of psychometrics, the scientific discipline underlying the development and measurement of psychological constructs such as an individual's satisfaction and quality of life.²⁴ PRO measures (PROMs), assessment tools that use a combination of qualitative and quantitative methods, have been developed to provide rigorous, reliable metrics of patient outcomes. An example is PROMIS, a set of person-centered measures for the evaluation and monitoring of the physical, mental, and social health of adults and children.²⁵



Outcomes such as posttreatment physical function, relief from pain, and improved mental health are known only to patients and, hence, must be reported only by them."

A Cleveland Clinic team of orthopedic surgeons, statisticians, and database/web programmers has built OrthoMiDaS Episode of Care (OME), a comprehensive, research-grade clinical outcomes

evaluation system.²⁶ OME measures outcomes and develops predictive models, with risk adjustments, for orthopedic surgeries ranging from arthroscopy to every knee, hip, and shoulder arthroplasty performed at Cleveland Clinic facilities in northeast Ohio and Florida. OME collects data at three time points. Prior to surgery, patients complete a health assessment comprising general and joint-specific validated PROMs totaling approximately 50 questions. Following the procedure, surgeons document disease severity, accurate diagnosis, major risk factors for outcomes, implants, and other technical details. One year after surgery, patients complete a follow-up assessment comprising the same PROMs issued at baseline plus a few follow-up items.

For breast cancer treatment, one of this paper's coauthors (A.P.), while working at Memorial Sloan Kettering Cancer Center (MSK), introduced the BREAST-Q PROs to personalize and optimize the care of each patient.²⁷ The BREAST-Q PROM was created by a team of surgeons, quality-of-life researchers, and experts in psychometrics to evaluate the outcomes of women undergoing breast cancer treatment. The clinical teams at MSK achieved response rates of nearly 95% by collecting PROs via cell phones, text messaging, on-site iPads, and email. This PROM is now available in more than 30 languages and used in routine clinical care, published studies, and national quality improvement initiatives involving more than 100,000 women.

The Patient-Reported Outcomes, Value, and Experience (PROVE) Center at Brigham Health has created a platform to systemically measure and analyze PROMs for the full cycle of breast cancer care. The platform contains specific patient-facing reports as well as IT interfaces for providers and department leaders to chart, display, and compare PROMs. A team of surgeons, medical and radiation oncologists, personnel from other clinical services (such as social work), and patient representatives selected the set of PROMs, together with appropriate risk adjustments, to use for all patients with breast cancer. The PROVE center is now expanding the platform to measure PROMs for other medical conditions.

Outcomes Metrics Serve Multiple Purposes

Outcomes measurement enables health care organizations to perform five essential functions:

1. Learn and Improve

The introduction of NSQIP's clinical defect metrics illustrates how consistent and widespread measurement can drive improvements in patient care. Ranking hospitals by valid outcome metrics identified the opportunities to learn from hospitals that have the lowest rate of surgical defects: 80% of hospitals that were early adopters of NSQIP significantly lowered their complication rates, and two-thirds significantly reduced their mortality.²⁸ Later studies documented fewer complications, shorter hospital stays, and reduced readmissions.^{29,30}

Aggregate-level PROs inform quality improvement initiatives, provide system-level feedback for health care management, and identify best practices across organizations. In a large, randomized trial, patients with cancer who provided PROMs while undergoing chemotherapy had better outcomes than did patients treated with only routine collection of clinical data.³¹

Systematic measurement of the outcomes from kidney transplant surgery in the United States has been required since 1990. With the higher accountability for performance because of mandatory reporting of outcomes, physicians performing kidney transplants adopted new technologies and formed multidisciplinary care teams that helped increase 1-year graft survival rates from 80% to 95%.³²

2. Research to Optimize Care

As clinical teams experiment with different treatment methods, care pathways, personnel teams, diagnostic tests, drugs, and devices, they can assess the benefits by measuring improvements in their patients' quality of life. A surgeon may believe that a new procedure is superior; its effectiveness, however, should be assessed by whether patients believe that the outcomes from the procedure are superior in meeting their unique needs. PROMs for individual patients enable providers to benchmark treatment progress and, when various treatment options exist, to use aggregate PROMs data to guide decisions and select the pathway most likely to lead to outcomes preferred by the patient.



In a large, randomized trial, patients with cancer who provided PROMs while undergoing chemotherapy had better outcomes than did patients treated with only routine collection of clinical data."

For example, MSK clinicians used the BREAST-Q PROM and associated cost data to compare silicone with own-tissue reconstruction. Silicone implants offered lower initial surgical costs and shorter operating room times. But own-tissue reconstruction led to superior PROMs, including better sexual well-being; higher psychosocial outcomes; greater patient satisfaction; and fewer long-term complications, office visits, and reoperations.³³ The superior patient outcomes, even with a more resource-intensive surgical process, led MSK to introduce own-tissue reconstruction as an optimal treatment pathway for many patients.

Cleveland Clinic's Quantitative Health Sciences (QHS) department team leveraged the institution's extensive database on patients' EHRs and clinical outcomes to generate a patient-specific prediction of the risk of hospital readmission within 30 days of discharge after pneumonia.³⁴ The risk calculator displayed better discrimination when compared with a standard CMS model between the higher- and lower-risk patients over 65 years of age. With this initial success, the QHS team helped other departments develop readmissions risk calculators for diseases. The prediction for each patient, together with its main risk drivers, is embedded in the patient's EHR, allowing the inpatient treating team to customize care according to the patient's specific and evolving needs.

3. Shared Decision-Making

Physicians need to get beyond telling patients only about the range of adverse events associated with their treatment. Patients should know the likely outcomes on quality-of-life issues that matter to them if they are to be truly informed and capable of making shared decisions with their physicians about treatment options. PROMs have allowed improved communication between patients and providers, leading to better clinical outcomes, self-management (e.g., symptom monitoring), patient engagement, resource utilization, and referrals to supportive care services.

Additionally, the collection of PROMs prior to a clinic encounter enables the clinicians to focus on issues of most concern to the patient. Cleveland Clinic, which has been routinely collecting PROMs with every outpatient visit since 2007, has found that provider review of the PROMs with the patient during the clinic visit was associated with significantly higher positive responses to all questions pertaining to ease of use of the PROM collection system, appropriateness of survey length, and patient-perceived benefit to their care overall.³⁵

4. Accountability

Patients, their families, employers, policy makers, regulators, and society are entitled to information on the specific outcomes that health care providers achieve. Currently, large variations in patient outcomes are invisible not only to the treating clinicians, but also to those receiving the treatment and those paying for it. While patients are generally unaware of differential health outcomes scores across providers, the lack of transparency may be especially burdensome to vulnerable populations and to those with socioeconomic disparities if they also lack access to informed referrals to the high-quality providers for their medical condition.^{36,37} Therefore, accountability and transparency of outcomes should also help reduce current inequities in health care delivery.

Cleveland Clinic began reporting the clinical outcomes of its cardiac patients to referring physicians in 1989. In 2004, CEO Toby Cosgrove, MD, expanded outcomes reporting to other clinical areas, and now outcomes books for 14 clinical areas are published annually.³⁸ The organization finds that the transparency and accountability of outcomes reporting are a catalyst for improving patient care and outcomes (essential functions 1 and 2, as mentioned previously). Since 2012, Hoag Orthopedic Institute has published an annual outcomes book that includes process metrics, clinical defects, and PROs.³⁹

5. Contracting

Value-based payment models, such as bundled payments, enable payments to be linked to outcomes achieved. Fee-for-service (FFS) payments, in contrast, are made independently of outcomes achieved. In fact, FFS rewards poor outcomes when hospitals get paid again for revision treatments and certain types of readmissions. Initial bundled payment contracts, such as the CMS BPCI program and many commercial contracts, could not make the payments contingent, in some way, on patient outcomes because providers were not collecting and reporting risk-adjusted outcome metrics for the procedures they performed. As noted in the multiple examples in this paper, we now have the capability of measuring outcomes that patients care about and adjusting the outcomes for patients' pretreatment status, setting the stage for the introduction of outcomescontingent bundled payment contracts.

Standardizing Outcome Measures

It would be extraordinarily inefficient for each provider organization, payer, and regulator to devise its own set of clinical and patient-reported outcomes for each medical condition. Additionally, when different outcome metrics are used for the same medical condition, meaningful comparisons among providers are impossible and patients are unable to determine their best treatment options. To maximize the benefits of measuring and reporting outcomes, standardization of metrics by medical condition must occur.



Physicians need to get beyond telling patients only about the range of adverse events associated with their treatment. Patients should know the likely effect of outcomes on quality-of-life issues that matter to them if they are to be truly informed and capable of making shared decisions with their physicians about treatment options."

The ACS NSQIP system has now standardized the definitions of surgical defects across a range of procedures. The nonprofit International Consortium for Health Outcomes Measurement (ICHOM), founded in 2011, has published (as of mid-2020) standard sets for 27 medical conditions encompassing more than 60% of the global disease burden. ICHOM convenes international multidisciplinary working groups of clinical experts and patients in an iterative process to reach consensus on the 10 to 15 outcomes that matter most to patients for each condition. ICHOM's outcome metrics are available online and have been published in peer-reviewed medical journals. Standard sets include PROMs, and case-mix variables, and they measure definitions and specify different collection points in time throughout the condition's cycle of care. More than 650 organizations in 32 countries, including Boston Children's Hospital, the Mayo Clinic, and the United Kingdom's National Health Service, have adopted ICHOM standard sets. For the collaborative ACS-Harvard Business School THRIVE project, colorectal surgeons drew upon the ICHOM standard set and peer-reviewed literature to select the MSK Bowel Function Index⁴² and the PROMIS Global Health-10 survey for measuring colon cancer surgery outcomes.

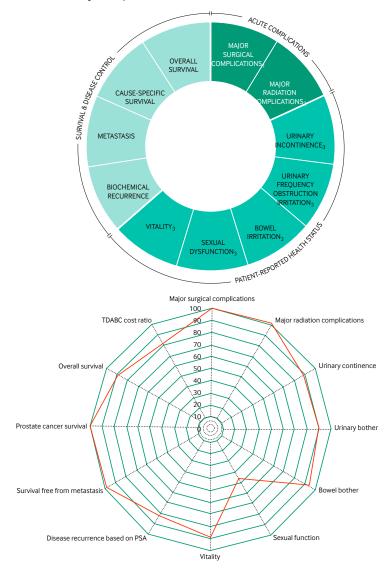
Displaying Outcome Measures

A challenge is how to communicate multidimensional outcome measures in a high-impact, accessible way. We have found that radar charts provide an excellent visualization tool (Figure 2).⁴³

Constructing a radar chart starts by selecting the relevant outcome dimensions for the medical condition. The top diagram in Figure 2 shows the ICHOM metrics for prostate cancer. The data for the radar chart on the bottom of Figure 2 come from the Expanded Prostate Cancer Index Composite survey instrument. The outcome data points are graphed on separate axes, all of which are scaled from 0 to 100, in which o represents the worst national or international performance and 100 represents perfect or ideal performance. The radar chart also contains an axis to measure the cost, as calculated by time-driven activity-based costing, of treating the condition. ⁴⁴ The radar chart enables clinicians, patients, and payers to compare the outcomes from and cost for alternative treatments for the same condition. In the case of low-risk prostate cancer, these would be brachytherapy, robotic prostatectomy, intensity-modulated radiation therapy, and proton therapy. ⁴⁵ The chart can also be used to compare the outcomes from different hospitals or clinicians treating the same condition using the same clinical pathway. The multiaxis radar chart enables outcomes data to be accessible to all stakeholders, promoting learning, improvement, and better decision-making in the continual effort to deliver better outcomes to patients.

Plotting Prostate Cancer Outcomes on a Radar Chart

The International Consortium for Health Outcomes Measurement Standard Set for localized prostate cancer (LPC) (top image) includes the outcomes that matter most to patients with LPC. It was produced by an international group of leading physicians, measurement experts, and patients. Below, the radar chart shows 6-month complications, 4-year patient-reported outcomes, 10-year survival, and provider costs for brachytherapy treatment of LPC. A score of 100 represents ideal performance on the measure. PSA = prostate-specific antigen, TDABC = time-driven activity-based costing (a measure of the cost incurred by the provider of the treatment).



Sources: Top chart based on information from the International Consortium for Health Outcomes Measurement. Localized Prostate Cancer. The Standard Set. Accessed November 11, 2020. https://www.ichom.org/portfolio/localized-prostate-cancer/. Bottom chart based on information from Thaker NG, Ali TN, Porter ME, Feeley TW, Kaplan RS, Frank SJ. Communicating value in health care using radar charts: A case study of prostate cancer. J Oncol Pract 2016;12:813-20 https://ascopubs.org/doi/10.1200/JOP.2016.011320. **
NEJM Catalyst (catalyst.nejm.org) © Massachusetts Medical Society

Barriers to Adoption of Outcome Measures

Despite the many benefits of health care outcome measures, several barriers limit their adoption and use. Technical challenges need to be overcome to allow outcome metrics to be collected from providers and patients and integrated into the EHR in a user-friendly, low-cost way. For PROMs, clinics must invest in hardware such as iPads in the clinic or rely on patients' use of a home computer or smartphone. This will also require administrative support to assist patients, hand out tablets, train patients on smartphones, answer questions, and support data submission. The data from these new platforms are often incompatible and difficult to embed into legacy IT systems and EHRs. These technical challenges are well known and are not insurmountable, but low-cost solutions to data collection and integration have been elusive. As telehealth (virtual visits and home monitoring) gains in utilization and achieves a greater share of health care delivery platforms and as innovative IT vendors expand capabilities, we should solve the technical issues and begin to integrate PROMs into the digital care delivery model.

Despite the efforts of organizations such as ICHOM and the ACS, many medical conditions still lack scientifically validated and risk-adjusted measure sets and measurement instruments. This leads to the perception, among many clinicians, that valid outcomes measurement remains infeasible. Medical education at all levels does not emphasize clinical outcomes and PROMs. Much work remains to be done to educate clinicians about the core concepts of PROMs. Studies must continue to demonstrate the importance and validity of PROMs.

Additional barriers to PROM collection are the burden on patients to complete surveys and the additional work required of providers. These barriers, however, can be lowered. We have found that when patients are asked questions that are relevant to their care, easy to understand, and not repetitive, even long surveys get high response rates. Clinicians respond well when they can easily access and interpret PROMs that are well targeted to the clinical condition, with data easy to access and interpret, to significantly improve the effectiveness and efficiency of their clinical treatments.



We have reached the point at which being accountable for patient outcomes must be seen as the table stakes for continuing to deliver care."

The cost of initiating and sustaining PROM data collection may be a higher barrier for smaller organizations that do not have the necessary personnel, infrastructure, and financial resources to maintain PROM collection. National programs, such as NSQIP and THRIVE, have the potential to lower these barriers by providing turnkey solutions that can be used by even small-practice organizations.

PROMs must be integrated into EHRs and clinical pathways so that clinicians can respond quickly to anomalies or emerging problems. For example, a clinical team asking a patient to submit a score on a suicidal ideation scale should be able to mobilize rapidly for a patient submitting an extreme

response. Collaborations with other departments may be necessary for patients who identify a concern that is not best addressed by the core team. Patients need to easily access and understand PROs if they are to participate in the decision about which clinician, primary care, or specialty center can best meet their specific health care needs. For shared decision-making, patients will likely need clinician guidance on how to interpret PROMs and the range of potential outcomes from treatment. Many national disease registries lack PROMs, have restricted access to data, and have limited or nonexistent patient involvement. Willingness to bring patients into the definition and coproduction of their outcomes remains a novel concept. But the potential benefits are great because introducing PROMs into a national registry, which several countries have already begun to do, allows for early detection of adverse events and device problems; this practice has been implemented, with text messaging, by the Australian Breast Device Registry. 47,48

While many clinicians and health care administrators may agree, in principle, with all the benefits from outcome measures, they may be reluctant to be the pioneers in their collection, use, and public reporting. A provider may be wary of being the first to report full sets of validated outcome measures, including PROMs, because users will not have benchmarks to determine whether the reported metrics are strong or weak. Government and private payers can help overcome this barrier by mandating that those they pay to treat patients must submit timely and accurate outcome metrics on their performance.⁴⁹

Looking Ahead

Achieving Dr. Codman's goal of measuring whether a patient's treatment has been successful, an apparently simple and straightforward target, has been challenging for health care providers to implement. Whether a treatment can be considered successful depends on multiple factors, including the patient's pretreatment status; the qualifications of the personnel performing the treatment; the facility's infrastructure and culture; the use of evidence-based clinical processes; the baseline incidence of treatment complications; and, finally, the challenges of measuring outcomes that matter to patients. Yet the multiple benefits from valid outcomes measurements — learning, improvement, shared decision-making, accountability, and payment contracting — have stimulated many important measurement advances during the past 25 years. These measurement advances have benefited from parallel advances in IT and artificial intelligence. We have reached the point at which being accountable for patient outcomes must be seen as the table stakes for continuing to deliver care. While continued development and utilization is needed, providers now have all the tools required to complete the journey of offering transparency in and accountability for their patients' outcomes.

Robert S. Kaplan, PhD, MS

Senior Fellow and Marvin Bower Professor of Leadership Development, Emeritus, Harvard Business School, Boston, Massachusetts, USA

Lara Jehi, MD, MHCDS

Chief Research Information Officer, Cleveland Clinic, Cleveland, Ohio, USA

Associate Professor of Neurology, Cleveland Clinic Lerner College of Medicine, Cleveland, Ohio, USA

Clifford Y. Ko, MD, MS, MSHS, FACS, FASCRS

Director, Division of Research and Optimal Patient Care, American College of Surgeons, Chicago, Illinois, USA

Robert and Kelly Day Professor of Surgery, UCLA David Geffen School of Medicine and Fielding School of Public Health, Los Angeles, California, USA

Andrea Pusic, MD, MHS, FACS, FRCSC

Chief, Division of Plastic and Reconstructive Surgery, and Director of Patient-Reported Outcomes, Value, and Experience (PROVE) Center, Brigham and Women's Hospital, Boston, Massachusetts, USA

Joseph E. Murray Professor of Surgery, Harvard Medical School, Boston, Massachusetts, USA

Mary Witkowski, MD, MBA

Fellow, Institute for Strategy and Competitiveness, Harvard Business School, Boston, Massachusetts, USA

Disclosures: Robert S. Kaplan is a paid advisor and has equity holdings in Avant-garde Health. Andrea Pusic is a codeveloper of the BREAST-Q and may receive royalties when it is used in for-profit, industry-sponsored clinical trials. Lara Jehi, Clifford Y. Ko, and Mary Witkowski have nothing to disclose.

References

- 1. Brand RA. Ernest Amory Codman, MD, 1869-1940. Clin Orthop Relat Res 2009;467:2763-5 https://springer.com/article/10.1007/s11999-009-1047-8 https://doi.org/10.1007/s11999-009-1047-8.
- 2. Ratcliffe S, ed. Lord Kelvin 1824–1907. British scientist. In Oxford essential quotations. 4th ed. Oxford, England: Oxford University Press, 2016. https://www.oxfordreference.com/view/10.1093/acref/9780191826719.001.0001/q-oro-ed4-00006236.
- 3. National Quality Forum. Measuring Performance. 2020. Accessed August 17, 2020. https://www.qualityforum.org/Measuring_Performance/Measuring_Performance.aspx.
- 4. Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. N Engl J Med 2016;374: 504-6 https://www.nejm.org/doi/full/10.1056/nejmp1511701 https://doi.org/10.1056/NEJMp1511701.
- 5. Porter ME. Value-Based Health Care Delivery: Core Concepts. Presentation Slide 12, Measure Outcomes for Every Patient. Harvard Business School. January 15, 2020. Accessed December 19, 2020. https://www.isc.hbs.edu/Documents/pdf/2020-intro-vbhc-porter.pdf.

- 6. Cella D, Yount S, Rothrock N, et al.; PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007;45(Suppl 1):S3-11 https://www.ncbi.nlm.nih.gov/pubmed/17443116 https://www.ncbi.nlm.nih.gov/pubmed/17443116 https://doi.org/10.1097/01.mlr.0000258615.42478.55.
- 7. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001;16:606-13 https://link.springer.com/article/10.1046/j.1525-1497.2001.016009606.x https://doi.org/10.1046/j.1525-1497.2001 https://doi.org/10.1046/j.1525-1497.2001 https://doi.org/10.1046/j.1525-1497.2001
- 8. Donabedian A. Evaluating the quality of medical care. Milbank Q. 2005;83:691-729 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2690293/https://doi.org/10.1111/j.1468-0009.2005.00397.x.
- 9. Hoyt DB, Ko CY, eds. Optimal Resources for Surgical Quality and Safety. Chicago: American College of Surgeons, 2017. https://www.facs.org/quality-programs/about/optimal-resources-manual.
- 10. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. A national evaluation of the effect of trauma-center care on mortality. N Engl J Med 2006;354:366-78 https://www.nejm.org/doi/full/10.1056/NEJMsa052049 https://www.ncbi.nlm.nih.gov/pubmed/16436768 https://doi.org/10.1056/NEJMsa052049.
- 11. Miller ME, Bleicher RJ, Kaufman CS, et al. Impact of Breast Center accreditation on compliance with breast quality performance measures at Commission on Cancer-accredited centers. Ann Surg Oncol 2019;26:1202-11 https://link.springer.com/article/10.1245/s10434-018-07108-7 https://www.ncbi.nlm.nih.gov/pubmed/30684159 https://doi.org/10.1245/s10434-018-07108-7.
- 13. Berwick DM. Continuous improvement as an ideal in health care. N Engl J Med 1989;320:53-6 https://doi.org/10.1056/NEJM198901053200110 https://doi.org/10.1056/NEJM198901053200110.
- 14. James BC. My Measured Experience. Excellence in Trauma and Critical Care. Intermountain Healthcare. September 19, 2014. Accessed December 19, 2020. https://intermountainphysician.org/ intermountaincme/Documents/17 Fri James-Measured%20Experience.pdf.
- 15. James BC, Savitz LA. How Intermountain trimmed health care costs through robust quality improvement efforts. Health Aff (Millwood) 2011;30:1185-91 https://www.healthaffairs.org/doi/full/10.1377/hlthaff. 2011.0358 https://doi.org/10.1377/hlthaff.2011.0358.

- 16. Hospital Compare. Centers for Medicare & Medicaid Services. Updated October 1, 2020. Accessed October 29, 2020. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare.
- 17. Berenholtz SM, Pronovost PJ, Lipsett PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit. Crit Care Med 2004;32:2014-20 https://journals.lww.com/ccmjournal/Abstract/2004/10000/Eliminating_catheter_related_bloodstream.6.aspx https://www.ncbi.nlm.nih.gov/pubmed/15483409 https://doi.org/10.1097/01.CCM.0000142399.70913.2F.
- 18. Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med 2006;355:2725-32 https://www.nejm.org/doi/full/10.1056/ https://www.nejm.org/doi.org/10.1056/ https://www.nejm.org/
- 19. Gawande A. The Checklist Manifesto: How to Get Things Right. New York: Picador, 2010.
- 20. World Health Organization. Checklists Save Lives. Bulletin of the World Health Organization. July 2008. Accessed December 19, 2020. https://www.who.int/bulletin/volumes/86/7/08-010708/en/.
- 21. Womack JP, Jones DT, Roos D. The Machine That Changed the World: The Story of Lean Production. New York: Rawson Associates, 1990.
- 22. Khuri SF, Daley J, Henderson W, et al.; National VA Surgical Quality Improvement Program. The Department of Veterans Affairs' NSQIP: The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. Ann Surg 1998;228:491-507 https://journals.lww.com/annalsofsurgery/Abstract/1998/10000/The_Department_of_Veterans_Affairs_NSQIP_The.6.aspx https://doi.org/10.1097/00000658-199810000-00006.
- 23. American College of Surgeons. ACS NSQIP History. Accessed October 30, 2020. https://www.facs.org/quality-programs/acs-nsqip/about/history.
- 24. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: A review of the options and considerations. Qual Life Res 2012;21:1305-14 https://springer.com/article/10.1007/s11136-011-0054-x https://doi.org/10.1007/s11136-011-0054-x.
- 25. National Institutes of Health, Office of Strategic Coordination–The Common Fund. Patient-Reported Outcomes Measurement Information System (PROMIS): Program Snapshot. January 29, 2019. Accessed November 16, 2020. https://commonfund.nih.gov/promis/index.

- 26. Piuzzi NS, Strnad G, Brooks P, et al.; OME Cleveland Clinic Orthopaedics. Implementing a scientifically valid, cost-effective, and scalable data collection system at point of care: The Cleveland Clinic OME Cohort. J Bone Joint Surg Am 2019;101:458-64 https://journals.lww.com/jbjsjournal/Abstract/2019/03060/Implementing_a_Scientifically_Valid,.10.aspx https://doi.org/10.2106/JBJS.18.00767.
- 27. Nelson JA, Allen RJ Jr, Polanco T, et al. Long-term patient-reported outcomes following postmastectomy breast reconstruction: An 8-year examination of 3268 patients. Ann Surg 2019;270:473-83 https://doi.org/10.1097/SLA.0000000000000003467.
- 28. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: An evaluation of all participating hospitals. Ann Surg 2009;250:363-76 https://journals.lww.com/pages/default.aspx https://journals.lww.com/pages/default.aspx https://doi.org/10.1097/SLA.obo13e3181b4148f.
- 29. Guillamondegui OD, Gunter OL, Hines L, et al. Using the National Surgical Quality Improvement Program and the Tennessee Surgical Quality Collaborative to improve surgical outcomes. J Am Coll Surg 2012;214:709-14, discussion 714-6 https://www.ncbi.nlm.nih.gov/pubmed/22265639 https://doi.org/10.1016/j.jamcollsurg.2011.12.012.
- 30. Cohen ME, Liu Y, Ko CY, Hall BL. Improved surgical outcomes for ACS NSQIP hospitals over time: Evaluation of hospital cohorts with up to 8 years of participation. Ann Surg 2016;263:267-73 https://doi.org/ACS_NSQIP_Hospitals.12.aspx https://doi.org/10.1097/SLA.00000000000001192.
- 31. Basch E, Artz D, Dulko D, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. J Clin Oncol 2005;23:3552-61 https://ascopubs.org/doi/10.1200/JCO.2005.04.275 https://ascopubs.org/doi/10.1200/JCO.2005.04.275 https://doi.org/10.1200/JCO.2005.04.275.
- 32. Porter ME, Baron JF, Chacko JM, Tang RJ. The UCLA Medical Center: Kidney Transplantation. Harvard Business School Case 711-410, August 2010. Revised March 2012. Accessed December 19, 2020. https://www.hbs.edu/faculty/Pages/item.aspx?num=39195.
- 33. Santosa KB, Qi J, Kim HM, Hamill JB, Wilkins EG, Pusic AL. Long-term patient-reported outcomes in postmastectomy breast reconstruction. JAMA Surg 2018;153:891-9 https://jamanetwork.com/journals/jamasurgery/fullarticle/2685263 https://doi.org/10.1001/jamasurg.2018.1677.
- 34. Hatipoğlu U, Wells BJ, Chagin K, Joshi D, Milinovich A, Rothberg MB. Predicting 30-day all-cause readmission risk for subjects admitted with pneumonia at the point of care. Respir Care 2018;63:43-9 https://www.ncbi.nlm.nih.gov/pubmed/29066584 https://www.ncbi.nlm.nih.gov/pubmed/29066584 https://www.ncbi.nlm.nih.g

- 35. Recinos PF, Dunphy CJ, Thompson N, Schuschu J, Urchek JL III, Katzan IL. Patient satisfaction with collection of patient-reported outcome measures in routine care. Adv Ther 2017;34:452-65 https://struck.springer.com/article/10.1007/s12325-016-0463-x https://www.ncbi.nlm.nih.gov/pubmed/28000165 https://doi.org/10.1007/s12325-016-0463-x.
- 36. Sequist TD. The Disproportionate Impact of Covid-19 on Communities of Color. NEJM Catalyst. July 6, 2020. Accessed November 16, 2020. https://catalyst.nejm.org/doi/full/10.1056/CAT.20.0370.
- 37. Yancy CW. COVID-19 and African Americans. JAMA 2020;323:1891-2 https://doi.org/10.1001/jama.2020.6548.
- 38. Kattan MW. Why Cleveland Clinic Shares Its Outcomes Data with the World. Harvard Business Review. September 22, 2015. Accessed December 19, 2020. https://hbr.org/2015/09/why-cleveland-clinic-shares-its-outcomes-data-with-the-world.
- 39. Hoag Orthopedic Institute. HOI Outcomes Reports. Accessed August 21, 2020. http://www.hoioutcomes.com/.
- 40. Porter ME, Kaplan RS. How to pay for health care. Harv Bus Rev 2016;94:88-98, 100, 134 https://www.ncbi.nlm.nih.gov/pubmed/27526565.
- 41. International Consortium for Health Outcomes Measurement. Standard Sets. Accessed August 21, 2020. https://www.ichom.org/standard-sets/.
- 42. Temple LK, Bacik J, Savatta SG, et al. The development of a validated instrument to evaluate bowel function after sphincter-preserving surgery for rectal cancer. Dis Colon Rectum 2005;48:1353-65 https://journals.lww.com/dcrjournal/Abstract/2005/48070/The_Development_of_a_Validated_Instrument_to.3.aspx https://doi.org/10.1007/s10350-004-0942-z.
- 43. Kaplan RS, Blackstone RP, Haas DA, Thaker NG, Frank SJ. Measuring and Communicating Health Care Value with Charts. Harvard Business Review. October 26, 2015. Accessed December 19, 2020. https://hbr.org/2015/10/measuring-and-communicating-health-care-value-with-charts.
- 44. Kaplan RS, Porter ME. How to solve the cost crisis in health care. Harv Bus Rev 2011;89:46-52, 54, 56-61 passim https://hbr.org/2011/09/how-to-solve-the-cost-crisis-in-health-care https://www.ncbi.nlm.nih.gov/pubmed/21939127.
- 45. Thaker NG, Ali TN, Porter ME, Feeley TW, Kaplan RS, Frank SJ. Communicating value in health care using radar charts: A case study of prostate cancer. J Oncol Pract 2016;12:813-20 https://ascopubs.org/doi/10.1200/JOP.2016.011320 https://doi.org/10.1200/JOP.2016.011320.
- 46. Katzan I, Speck M, Dopler C, et al. The Knowledge Program: An innovative, comprehensive electronic data capture system and warehouse. AMIA Annu Symp Proc 2011;2011:683-92 https://www.ncbi.nlm.nih.gov/pubmed/22195124.

- 47. Nelson EC, Dixon-Woods M, Batalden PB, et al. Patient focused registries can improve health, care, and science. BMJ 2016;354:i3319 https://www.ncbi.nlm.nih.gov/pubmed/27370543 https://doi.org/10.1136/bmj.i3319.
- 48. Ng S, Pusic A, Parker E, et al. Patient-reported outcome measures for breast implant surgery: A pilot study. Aesthet Surg J 2019;39:NP314-21 https://academic.oup.com/asj/article-abstract/39/8/NP314/5306134?redirectedFrom=fulltext https://doi.org/10.1093/asj/sjz023.
- 49. Kaplan RS, Porter ME. Mandate Outcomes Reporting. Health Management, Policy and Innovation. 2019. Accessed December 19, 2020. https://hmpi.org/2019/11/25/mandate-outcomes-reporting/.