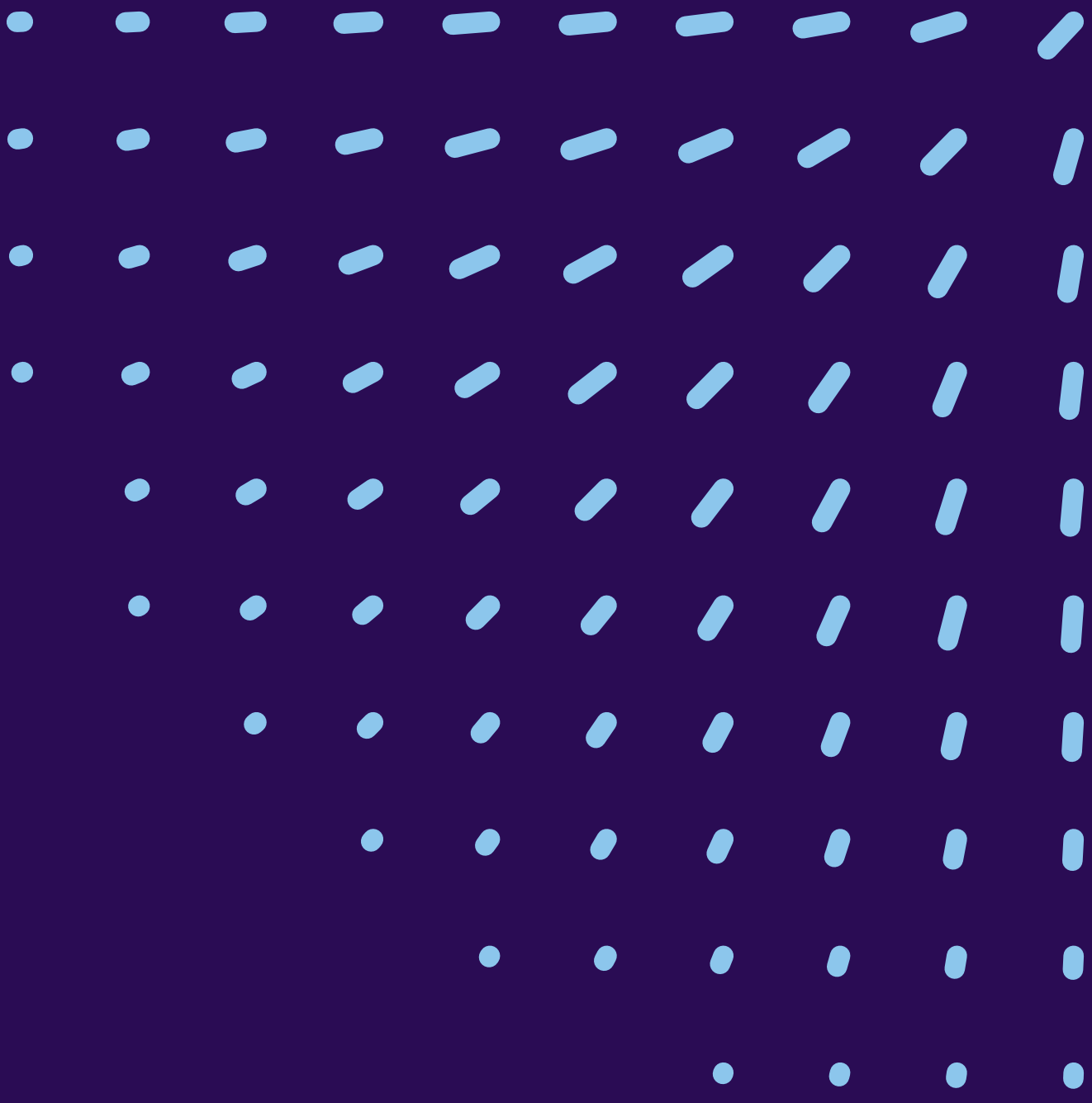
IBM Watson Health™

Key elements of the empirical operational benchmarking methodology



5 benefits of the operational benchmarking methodology

Adjusting performance measurements based on workloads has five advantages:

1. Benchmarks are based on a larger set of hospitals and several years of data.
2. The model is flexible and automatically adapts when additional factors are introduced.
3. The model enables staff to focus more time on performance improvement initiatives versus managing compare groups.
4. The model enables executives to engage directly in benchmarking analytics without the need for training about comparison group selection criteria.
5. The model enables direct performance comparisons of hospitals with different characteristics.

Developing sound comparison standards is a critical part of the process to identify performance improvement opportunities in health systems, hospitals and departments. Traditionally, health systems have employed a comparison group method to set operational benchmarks for performance evaluations. The criterion selected for the benchmarks is based on what the health system determines to be relevant.

This approach is valid but can result in narrower comparison groups that leave out evaluation data from facilities that offer similar services but do not match the self-identified set of criteria.

There is an alternative approach for health care systems that levels the playing field to ensure evidence-based data drives operational benchmarks. When used in addition to – or instead of – the comparison group method, health systems can employ an empirical method to uncover improvement opportunities across all hospitals and departments.

The empirical method addresses the challenge inherent in operational benchmarking to understand what factors have an impact on performance and how to apply those parameters to generate comparison groups. By using an empirical approach incorporated into analytic methods applied against data, health systems can access comparison groups that account for operational characteristics that impact performance without manual effort or limitations on the compare group size.

It is important to understand key elements of the empirical operational benchmarking methodology to appreciate the level of rigor required to make comparisons based on validated, multi-year data sets from a large number of facilities.

Large data sets drive decision-making

The basis of the empirical operational benchmarking methodology is an analysis of four years of annual data that includes data from the system, facility and department levels. Each level has attributes that require consideration in the development of benchmarks to produce comparisons on a fair basis that consider the varied needs of all facilities.

Facility level benchmarking

The first step at this level is to identify facility characteristics and factors that consistently influence hospital resource demand over multiple years of data. The multivariate model evaluates all identified factors, then retains only the factors that add significant predictive power.

The operational model simultaneously uses the retained hospital characteristics to adjust facility performance indicators to be comparable across hospitals with differing characteristics. The process is similar to clinical performance models that simultaneously use diagnoses, procedures and patient characteristics to adjust mortality, complication or readmission rates to be comparable across hospitals or clinical service lines.

The facility adjustment factor is updated for every period in which data are submitted to account for changes in facility volumes and characteristics. Then a workload adjustment factor is applied for the period against the volume of CMI (Case Mix Index) weighted discharges which is the default workload denominator for all facility level KPIs (Key Performance Indicators).

For example, an organization has major teaching status, operates a Level 1 trauma center and has other specific characteristics expected to increase resource demand. Together these factors might result in an adjustment factor of 1.44. The facility is expected to need 44% greater work effort than the average hospital to care for the same number of patients. To account for the meaningful difference in expected resource demand, the model calculates the CMI-weighted discharges in the selected period and multiplies it by the adjustment factor for the hospital for that time period. This step levels the playing field for performance in that period.

Department level benchmarking

The creation of department-level benchmarks starts with the same rigorous research on each department to determine which facility and departmental factors consistently influence resource demand. Most department types have between 1 to 5 meaningful, non-redundant factors.

The model considers key factors such as total operating beds, total patient discharges, hospital all-payer case-mix index and hospital teaching status. Other facility characteristics such as location and nurse magnet facility status also relate to expected work effort in some departments.

Many characteristics specific to particular departments impact labor demand, including types of patients served by department (pediatric, trauma, telemetry, etc.), hours staff on site, specific types of services performed, and whether staff of the unit is unionized.

A regression tree model is applied to department data to empirically generate comparison group criteria that define mutually exclusive comparison groups that are as homogenous as possible with respect to expected work effort. The method ensures that at least 10 department instances are assigned to each comparison group in each data reporting period.

System level benchmarking

Health systems have different and more complex characteristics than individual hospitals. For example, where a single hospital has one specific location, the facilities that comprise a system are often located in multiple settings such as urban, suburban and rural.

Summary characteristics can also be problematic. A 1,500 bed system may consist of ten 150-bed hospitals or three hospitals with 750, 500 and 250 beds. Resource demands are quite different in each scenario.

The complexity of system characteristics is compounded by the number of multiple hospital systems as compared to individual hospitals. The number of potentially important system factors is much larger than the number of systems for which there are data, making it difficult to identify factors that would explain meaningful differences in resource demand between systems. Therefore, systems are compared to all systems without applying adjustments or specifying separate comparison groups.

Quality data produces quality results

Implicit in the application of the empirical operational benchmarking methodology is that all analyses use data that passes strict consistency and completeness quality checks.

Key performance indicators and workload units
Worked hours are the purest indicator of facility and department resource demand. Analyses use worked hours non MD to determine facility adjustment factors and department comparison groups. Only the highest ranked workload unit of service in each department as determined by department-specific scales are used to determine comparison group criteria.

For a given hospital department, operational performance indicators are displayed based on the highest ranked workload unit for which the hospital department submitted data and for which there were at least ten observations in the comparison group for the data period.

Facility characteristics

The intent is to adjust workload or create comparison groups to take into account additional indications of intensity of services and work effort required, not to reflect management choices on operations.

– Some facility characteristics may affect many departments within a hospital, including factors like total capacity and facility patient volumes that contribute to economies of scale and result in the need for less labor and lower cost required per unit of service for larger capacities and service volumes. In addition, facility characteristics such as major teaching status, case mix index, disproportionate share status, and location may act as a proxy for the types or severity of patients and the level of care offered by departments within the hospital.

Department specific characteristics

Departments realize economies of scale related to their own capacity and department workload. Many departments have a capacity element such as numbers of beds, bassinets, operating rooms, procedure rooms, treatment rooms or examination spaces. Every department has one or more workload volume measures. Often larger and busier departments have lower expected worked hours per unit of service than smaller, lower volume departments of the same type.

Correlated characteristics and factors

Many factors and characteristics correlate highly with each another. For example, hospitals with larger numbers of beds tend to have higher volumes of discharges, are less likely to be rural, and are more likely to have major teaching status. Conversely, a smaller, non-teaching hospital is unlikely to have a high case-mix index, or have a Level 3 or 4 NICU, perform organ or bone marrow transplants, or maintain a Level 1 trauma center. Thus, many seemingly important criteria prove redundant in models. If a model already includes predictors such as teaching status and number of adjusted discharges, additional factors that highly correlate with teaching and discharges might not add any additional predictive power and may not be retained in a model.

Reliable and consistent characteristics

Facilities and departments have hundreds of characteristics that may distinguish them from other facilities and departments. In any given time period, many of these characteristics may appear related to required work effort but may only be temporary. Only factors that consistently relate to required work effort over multiple years are retained in the facility model and in the automatically generated department comparison group criteria.

Enabling fair comparisons

The goal of operational performance evaluations is to identify opportunities to improve quality of care, patient outcomes and costs. The empirical operational benchmarking method adds to health systems' abilities to equitably identify improvement opportunities with a complete picture of current performance and emerging trends.

About IBM Watson Health

Each day, professionals throughout the health ecosystem make powerful progress toward a healthier future. At IBM Watson Health, we help them remove obstacles, optimize efforts and reveal new insights to support the people they serve. Working across the landscape, from payers and providers to governments and life sciences, we bring together deep health expertise; proven innovation; and the power of artificial intelligence to enable our customers to uncover, connect and act — as they work to solve health challenges for people everywhere.

For more information on IBM Watson Health, visit ibm.com/watsonhealth.

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
May 2019

PROV 00201 0419

IBM Watson Health™