

Heatmap, a Visual Tool for Comparative Analytics

Jamie Palumbo MA; Randall Wetzel MD, MBA; Gerardo Soto-Campos PhD, MS



Virtual Pediatric Systems, LLC (myvps.org)
Los Angeles, CA, USA.



INTRODUCTION

Quality improvement¹ (QI) in healthcare relies, in part, on benchmark comparisons between different hospitals. Benchmarking compares a **target** hospital, the subject under study, against a peer group of hospitals called a **reference** group. Changes in specific metrics between the target and its reference provide important insight into QI.

The VPS, LLC² database collects high quality pediatric (PICU), cardiac (CICU), neurologic (NCC), and neonatal (NICU) critical care data. It contains more than one million cases with more than 300 fields ranging from clinical and demographic to operational data.

The most common clinical and operational metrics in the VPS, LLC database used to evaluate ICU quality are:

- Severity of illness through standardized mortality ratios³ (SMR)
- Standardized length of stay ratios⁴ (SLOS)
- Discharge delay⁵
- Readmissions within 24 hours
- Extubation failure within specific time thresholds^{6,7}
- Percent of patients receiving mechanical ventilation⁸
- Percent of patients discharged directly home.

This poster presents a heuristic method that combines the metrics listed above in such a way that one can rank hospital performance by metric, and determine total performance through a composite weighted score of the individual metrics. While the actual weights are proprietary IP of VPS, the general approach is discussed below.

METHODOLOGY AND RESULTS

For **each metric**, the percentiles of the distribution for all the hospitals (composite metric) are calculated. The **mean value** of each hospital metric naturally falls within a percentile range of the composite metric.

The heatmap uses the following percentile ranges for the composite metrics: 0-10, 10-20, 20-50, 50-80, 80-90, 90-100. Each range is assigned a color and a numerical *metric score*, which reflects a degree of clinical benefit. For example, since low mortality (low SMR) is a desirable outcome, the heatmap assigns it a large numerical metric score and color green. Similarly, high mortality (high SMR) receives a low metric score and a color red denoting hindrance. Figure 1 summarizes the color scale used in the heatmap.



Figure 1. Heatmap colors associated with percentile ranges of the composite metrics.

In order to build a score to rank hospital performance, each metric was assigned a specific weight, w . The weighted sum along all metrics gives a composite score, unique to each hospital, called the **reward score**:

$$Reward\ Score = \sum_i w_i * M_i$$

Each weight w_i is associated with a degree of **clinical importance**, defined by VPS, LLC medical experts. In the equation above, M_i stands for each of the metrics used in the heatmap.

The reward score was scaled to a maximum of 100.

The metrics readmissions, extubation failure, percent mechanical ventilation, and discharged home do not make sense as absolute values. The heatmap includes these metrics measured in relation to a benchmark, which is built as the five-year average from all the hospitals in the VPS registry.² For example, Figure 2 displays a histogram for percent of ICU stay on mechanical ventilation defined as ventilation days divided by physical length of stay. The VPS benchmark is 37.3%. The heatmap uses the absolute value of the difference between a hospital ICU ventilation percent and the VPS benchmark. In Figure 3, this metric is called “% Ventilation”. The rationale behind these types of transformations is that the benchmark provides a data-driven **standard of practice**. Similar transformations were done for readmissions, extubation failure, and percent of patients discharged home.

For the metrics readmission rate, mechanical ventilation rate, extubation failure within 6 and 24 hours, and percent discharged home, the heatmap shows whether each hospital fell above or below the standard of care benchmark, with the symbols “+” and “-” respectively.

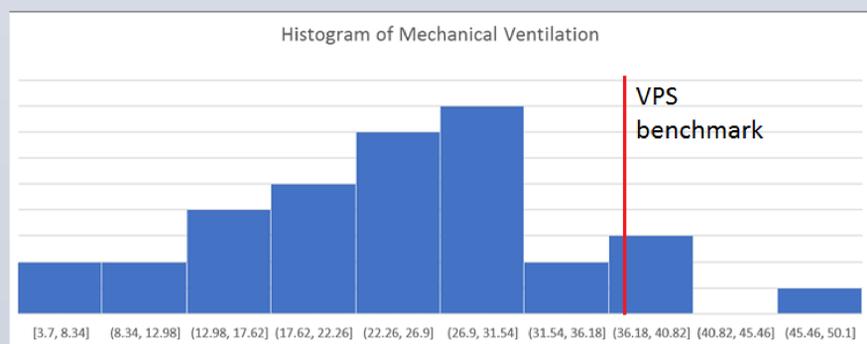


Figure 3. Histogram of percent mechanical ventilation, defined as ventilation days divided by physical length of stay for a group of hospitals from the VPS registry. The red vertical line represents the five-year average value for all the units from the VPS registry, 37.3%. The heatmap assumes that this benchmark is the standard of care, and the metrics entering the heatmap should be measured with respect to this standard. Similar assumptions were made for readmission, extubation failure, and percent of patients discharged home.

An example of the heatmap for sixteen hospitals is displayed in Figure 3. The reward score, 73, is shown below the target hospital.

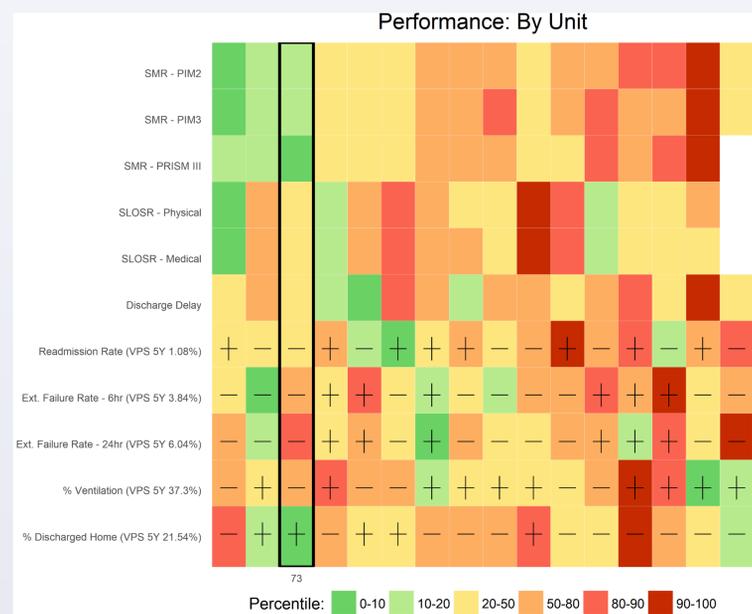


Figure 3. Comparison of a target of over 4,600 cases versus a reference group of 15 hospitals with more than 72,000 cases. Each column represents a hospital, and each row shows a metric. Therefore, each cell represents a hospital’s individual metric value. The black rectangle represents the target with its reward score (73) displayed at the bottom. SMR is displayed for three mortality models: PIM2⁹, PIM3-NA¹⁰, and PRISM III¹¹. Physical and medical SLOS are represented for the VPS PRISM III physical⁴ and medical¹² length of stays models. The “+” and “-” symbols denote whether each particular hospital falls above or below the VPS benchmarks, which are constructed from five years of data in the VPS registry (>150 hospitals in the US and abroad).

Summary statistics for the reward scores in the sixteen hospitals used in the previous example are represented in Table 1.

Table 1. Summary statistics for the sixteen reward scores in the heatmap comparison between the target and its comparative reference.

Minimum	1 st Quartile	Median	Mean	3 rd Quartile	Maximum
41	50	56	58.4	64.5	80

CONCLUSIONS AND FUTURE WORK

- The heatmap is a promising tool to rank hospital performance in comparative studies.
- The tool shows a visual comparison between multiple metrics and hospitals in one snapshot.
- In its current form, the heatmap ignores correlations between the metrics used to build it as well as overlaps between individual metric distributions.
- The set of metrics utilized to build the heatmap were chosen by an expert panel of VPS clinicians. Other clinical experts may choose a different set of metrics to rank performance.
- The placement of the cells along the percentile ranges used mean values or percentages of the metrics. Other central tendency measures could be used.
- The heatmap enables a relative comparison. As more hospitals are used in the reference, the reward score of the unit could change.
- Future work will apply the heatmap technology to all the hospitals in the VPS registry.

REFERENCES

1. Institute of Medicine 2001, <http://www.nationalacademies.org/hmd/Reports/2001/Crossing-the-Quality-Chasm-A-New-Health-System-for-the-21st-Century.aspx>.
2. VPS, LLC, www.myvps.org.
3. Gemre RJ and Vugh J, Scoring systems in pediatric intensive care: PRISM III versus PIM, Int Care Med (2002) 204-7.
4. VPS length of stay model, <https://s3.amazonaws.com/vpspublic/PRISM+LOS+brochure.pdf>.
5. Tirovapati R et al, Intensive care discharge delay is associated with increased hospital length of stay: A multicentre prospective observational study, PLoS ONE (2017).
6. Kurachek SC et al. Extubation failure in pediatric intensive care: A multiple-center study of risk factors and outcomes, Crit Care Med. (2003) 2657-64.
7. Kim F et al, Timing of Extubation Failure in Pediatric Intensive Care Units, Proc SCCM (2018) 645.
8. Morrison WE et al, Roger’s Handbook of Pediatric Intensive Care (Wolters Kluwer, 2017, 5th Edition).
9. Slater A et al, PIM2: a revised version of the Paediatric Index of Mortality, Int Care Med (2003) 278-285.
10. Soto-Campos G et al, PIM3 North American Development [calibration, validation and comparison to PIM3] http://docs.wixstatic.com/ugd/40378f_7d95638844c04c67abde9fed99f256b1.pdf.
11. Pollack M et al, PRISM III: An updated Pediatric Risk of Mortality score, Crit Care Med (1996) 743-52.
12. VPS medical length of stay model, results not published.

ACKNOWLEDGEMENTS

Chloe Gordon, Jill Hall, and Ashley Sweeney from VPS, LLC made important suggestions about this work.