Published in final edited form as:

Am J Manag Care. ; 26(6): e172–e178. doi:10.37765/ajmc.2020.43492.

FASStR: A Framework for Ensuring High-Quality Operational Metrics in Healthcare

Elham Torabi, MSIE, PhD,

Department of Computer Information Systems and Business Analytics, College of Business,
James Madison University, 421 Bluestone Dr. Harrisonburg VA 22807

Tugba Cayirli, PhD,

Faculty of Business, Ozyegin University, Nisantepi Mah. Cekmekoy 34794 Istanbul Turkey

Craig M. Froehle, PhD,

Department of Operations, Business Analytics, and Information Systems, Carl H. Lindner College
of Business, University of Cincinnati, 2925 Campus Green Dr. Cincinnati OH 45221

Department of Emergency Medicine, College of Medicine, University of Cincinnati, CARE/Crawley
Building, Cincinnati, OH 45267

Kenneth J. Klassen, PhD,

Department of Finance, Operations & Information Systems, Goodman School of Business, Brock
University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S 3A1

Michael Magazine, PhD,

Department of Operations, Business Analytics and Information Systems, Carl H. Lindner College
of Business, University of Cincinnati, 2925 Campus Green Dr. Cincinnati OH 45221

Denise L. White, PhD,

Department of Operations, Business Analytics, and Information Systems, Carl H. Lindner College
of Business, University of Cincinnati, 2925 Campus Green Dr. Cincinnati OH 45221

James M. Anderson Center for Health Systems Excellence, Cincinnati Children's Hospital
Medical Center, 3333 Burnet Avenue, MLC 5040, Cincinnati, OH 45229-3039

Michael Ward, MD, PhD, MBA*

Department of Emergency Medicine, Vanderbilt University Medical Center, 1313 21st Ave. S, 703
Oxford House, Nashville, TN 37232

Abstract

Objectives—Poorly defined measurement impairs inter-institutional comparison, interpretation of results, and process improvement in healthcare operations. We sought to develop a unifying framework that could be used by administrators, practitioners, and researchers to help define and document operational performance measures that are comparable and reproducible.

Study Design—Retrospective Analysis

*Corresponding Author: michael.j.ward@vumc.org, (615) 936-8379.

Methods—Healthcare operations and clinical researchers used an iterative process consisting of: 1) literature review, 2) expert assessment and collaborative design, and 3) end-user feedback. We sampled the literature from the medical, health systems research, and healthcare operations (business and engineering) disciplines in order to assemble a representative sample of studies where outpatient healthcare performance metrics were used to describe primary or secondary outcome of the research.

Results—We identified two primary deficiencies in outpatient performance metric definitions: incompleteness and inconsistency. From our review of performance metrics, we propose the FASStR framework for the Focus, Activity, Statistic, Scale type, and Reference dimensions of a performance metric. The FASStR framework is a method by which performance metrics can be developed and examined from a multidimensional perspective to evaluate its comprehensiveness and clarity. The framework was tested and revised in an iterative process with both practitioners and researchers.

Conclusions—The FASStR framework can guide the design, development, and implementation of operational metrics in outpatient healthcare settings. Further, this framework can assist researchers in the evaluation of the metrics they are using. Overall, the FASStR framework can result in clearer, more consistent use and evaluation of outpatient performance metrics.

PRECIS

Through literature review and collaborative design, we propose the FASStR framework to provide a systematic approach to healthcare operation metric definition and use.

Keywords

Performance Measurement; Outpatient Care; Operational Performance; Health Services

INTRODUCTION

Performance measurement is fundamental to healthcare, from clinical operations to research. The Agency for Healthcare Research and Quality (AHRQ) recognizes this in its mission statement, to “*develop and spread evidence and tools to measure and enhance the efficiency of health systems— the capacity to produce better quality and outcomes.*”[1]

Measurement of operational performance in healthcare is ubiquitous. We measure patient waiting, clinic durations, staff overtime, costs, patient satisfaction, no-shows, and numerous other metrics. Measures of operational performance provides a basis for evaluating and comparing the performance of different healthcare institutions, operating practices, and policies. While seemingly easy to create, developing usable and comprehensible metrics are quite challenging. Inadequately designed and documented operational metrics creates ambiguity and confusion and potentially lead to incorrect managerial decisions potentially at odds with the objectives of high-quality and safe patient care.[2]

To be maximally useful, a metric needs several important characteristics. First, it should be reliable and valid, measuring its true target consistently across time and contexts.[3] Second, it needs to be consistent with the goals, standards, and practices of the organization within

which it is applied. Third, a metric must be clear and unambiguous in its definition and use. Fourth, it should be as generalizable as possible without reducing its utility or specificity. Finally, it must be relevant to practice and aid in managerial decision-making; metrics that do not directly contribute to the management of the organization or are not sufficiently sensitive to detect meaningful operational changes, are, at best, potential distractions from more critical information.

Inadequately defined and documented metrics can degrade the consistency of metric application from one instance (facility, time period, etc.). This is the essence of “test-retest reliability.” A lack of reliability undermines the sharing and acceptance of critical information, slowing the spread of vital data for efforts such as interventions and improvement collaboratives. Without clear and thorough definitions, we are unable to adequately assess the relevance or utility of a particular metric to any specific environment or situation diminishing the metric’s generalizability. For example, if hospital A defines “patient waiting” differently than hospital B does, but both *assume* identical meanings, they can (wrongly) arrive at quite different conclusions about the efficacy of a particular policy. This can manifest even *within* an organization due to definitional “drift” as employees turnover, rendering historical comparisons meaningless. Finally, organizations may not understand or be able to communicate what is being measured and/or why it changed, thereby potentially reducing buy-in from staff and stakeholders.

Some healthcare organizations have begun to standardize measurement of key performance metrics.[4] The National Quality Forum (NQF) has established a system of metric evaluation and stewardship.[5] However, even widely adopted metrics may use a variety of definitions and then collected and reported along a spectrum of interpretations. Inadequacies in one or more of the characteristics listed above, such as consistency or generalizability, can undermine even well-supported national efforts at improving and standardizing operational metrics in healthcare. NQF’s “Measure Evaluation Criteria” uses five criteria to evaluate the suitability of measures and these include: importance, scientific acceptability of measure properties, feasibility, usability, and related/competing measures. We sought to develop a unifying framework focused on the second property (acceptability of measure properties) to improve the development, deployment, and spread of consistent, well-defined metrics and accelerate translation across organizations and research. Adoption and use of such a framework by researchers and organizations could make study results more comparable, repeatable, and, most importantly, more applicable to practice. This creates a “virtuous cycle,” enabling faster and more complete operational improvement.

There is a need for a framework that helps researchers and practitioners define their metrics of interest more clearly. In order to develop such a measurement framework, we used an iterative process consisting of: 1) examination of the literature, and 2) expert assessment and collaborative design. We brought together a team of international experts in healthcare operations and clinical research with MD and/or PhD degrees in Operations Management and decades of experience in interdisciplinary research to examine a metrics database and collaboratively design an organizing framework.

Literature Review

First, we sampled the literature from the medical, health systems research, and healthcare operations (business and engineering) disciplines to assemble a representative sample of papers in which performance metrics were the focus of the research. We narrowed our search to outpatient clinical operations because, with nearly a billion ambulatory care visits in 2015, this setting serves the largest proportion of patients in the U.S.[6]

The number of studies involving outpatient operations has grown substantially over the past three decades (Figure 1). In the interest of parsimony, we further focused on *patient flow* as our illustrative example of an operational context due to both its ubiquity and importance across a wide range of clinical settings.

Our initial sampling of the patient-flow literature identified 268 papers, of which 126 met our inclusion criteria (published during the inspection period; at least one operational metric was mentioned and quantified in some way). For each article, we identified and documented the operational performance metrics and their definitions (if provided by the authors). We found over 200 different outpatient performance metrics discussed in these 126 papers. We then organized the metrics into an online database of quotes and metric definitions, and other relevant data.

We observed many opportunities for improvement in the literature regarding how thoroughly and clearly performance metrics were defined and documented, with nearly all papers exhibiting at least one deficiency. We organized these deficiencies into two qualitative clusters -- inconsistent and incomplete definitions. We reviewed and processed more than 200 operational metrics such as clinic over time, provider idle time, and utilization. In this paper, we use the “patient waiting” metric to provide illustrative examples.

A. Inconsistent definitions—We found at least 10 different definitions for patient waiting time. Definitions included “time patient is waiting for the doctor,”[7] “number of time slots between arrival and service,” and “waits and delays.”[8–10] The most common definition was “mean time between patient’s arrival at the clinic and when first seen by the physician.”[11–19] Some studies considered waiting in the waiting room while others considered waiting in the exam room.[20] However, subtle variations existed, for example, consider the definition of “arrival.” Does it mean the time a patient entered the clinic, when a patient was registered, or when a patient self-registered in a kiosk? Discrepancies in definitions of seemingly identical metrics diminishes the generalizability of these metrics and their utility in practice.

B. Incomplete definitions —Many metric definitions lacked essential details necessary for use. Some patient waiting metric definitions failed to define the period of time being measured. For instance, consider the patient waiting time definition of “waits and delays;”[8] the reader has no way of knowing when a wait or a delay starts or ends in that study. Surprisingly, 27/68 papers (39.7%) we reviewed that claimed to measure patient waiting time did not provide *any* definition of the metric, suggesting that incomplete definitions are common. Table 1 shows sample definitional deficits.

To illustrate how the results of a study may be affected, consider the following. Envision a hypothetical, single-physician clinic in which the manager uses an algorithm that simultaneously identifies the best patient schedule to both minimize patient waiting time and physician idle time (i.e., time the physician waits between patients). There have been several of these algorithms developed.[7,21–23] The algorithm recommends a *variable-interval schedule*, meaning that appointment slots could be any duration—the slots may not match how long it takes the providers to deliver the care. How the components of patient waiting and physician idle time are defined will influence what happens in practice. For instance, if *total* waiting for all patients during a clinic session is used, the metric is weighted more towards the patients and the resulting schedule will have appointment times spread out more, causing less waiting for patients but more idle time for the physician. Alternatively, if *average* waiting time per patient is used, the resulting schedule will set appointment times closer together (i.e., shorter appointment slots), thus favoring the physician with less idle time and resulting in longer waits for the patients. As seen in Figure 2, the first schedule could result if total patient waiting time is used and the second if average patient waiting time is used. Note that neither metric definition is inherently better than the other; the point is, if metrics are not carefully and thoughtfully defined, this ambiguity can cause unintended consequences, reduce the ability to compare the results to other analyses and generalize to other clinics, and reduce optimal management decision-making.

The FASStR Framework

The framework development process involved several iterative rounds, during which experts examined the literature dataset and proposed and discussed additions or changes to the nascent framework. After each round of revision, pairs of members of the investigator team, including both researchers and practitioners, tested the draft framework by applying it to metrics in an attempt to identify cases where the framework lacked organizational value or insight. In each round, the framework was revised based on this feedback. The team communicated online and held face-to-face meetings at national conferences over a period of 26 months. The process consisted of the following:

1. All research articles were reviewed by two team members. An online database with article/author information and the list of performance measures used was created. In front of each performance metric, the definition of the metric (if any) was included.
2. Each article was assigned to a team member to review and confirm the entries of the data base. This step was also necessary so that all team members became familiar with the database.
3. All team members met at an international conference to discuss the metrics and come up with a preliminary list of the framework dimensions.
4. All of the performance metrics were listed and categorized. Two experts were assigned to review each metric in order to identify what data was missing. The primary reviewer for each metric identified different dimensions for that metric, the secondary reviewer would review and critique.

5. Each team member was assigned a set of metrics to categorize by proposed dimension and identify missing components. This step was repeated until theoretical saturation, the point at which the framework could account for all variations of all metrics.
6. The dimensions were finalized through a face-to-face meeting.

The resulting framework draft was presented at three academic conferences to solicit feedback from both researchers and practitioners and was further refined after each presentation. Feedback from conference attendees was iteratively assessed against the framework and reviewed and discussed by the research team until consensus was reached.

We refined the organization of our findings into five thematic domains comprising the framework name: Focus, Activity, Statistic, Scale type, and Reference. Described below, these dimensions involve the subject of what is being measured, the activity being measured, the calculation and units of measurement, and the comparator to which the measurement is related.

Focus—Included in the majority of papers we reviewed, the Focus of a metric is its subject; the person, entity, or object of interest of the metric. Also called the “unit of analysis,” the focus could be a patient, provider, an exam room, a clinic, an operating room, a division, an entire hospital, or a piece of equipment.

Activity—Activity involves what the metric’s Focus is doing. In other words, what is the action, event, or status the metric is measuring? If the Focus is the *noun* of a metric, Activity is the *verb*. As examples, patients (Focus) could be *waiting* (Activity); an operating room (Focus) could be *occupied* with a procedure (Activity) or being *cleaned* (Activity). The Activity’s definition should be specific enough so that there is no confusion or room for misinterpretation as to when the Activity starts and when it concludes.

Statistic—Statistic is how the metric is arithmetically calculated. Common statistics include pure counts (i.e., sums), central tendency (mean, median), percentiles, variation (standard deviation), the minimum or maximum, and proportions. For metrics expressed as ratios, such as *minutes per clinic session* or *interactions per patient*, the denominator (following after the “per”) should be clearly defined. The timeframe data are collected (e.g., over an hour, clinic session, etc.) should be specified as part of its Statistic. For example, a provider’s idle time during a clinic session is not the same Statistic as a provider’s idle time *per patient* during a clinic session.

Scale type—Scale type represents the units or amounts in which the metric is expressed or, in some cases, the format of the measurement instrument. The Scale type is often inextricably tied to the metric’s Focus or Activity. Using “average number of patients in the waiting room” as an example, the Focus of the metric is patients and the Scale type (its units) is also patients. The Scale type can often be straightforward to determine for time-based scales (e.g., clinic duration), where minutes or hours are typically used. Other common Scale types include Likert scales (e.g., satisfaction) and categories (e.g., yes/no). Some metrics may have multiple units combined in their Scale type definition, such as when

the Statistic is a proportion like “patients per hour.” In that example, both “patients” and “hours” are essential elements of the Scale type for the metric.

Reference—The pre-defined reference to which the value is being compared (if any). Metrics are typically used to guide decision-making so they benefit from having a reference (e.g., benchmark, industry standard). Even if no objective standard for comparison exists, a previous period’s value is commonly used to assess change. The Reference should have a directional implication that is easy to understand whether higher or lower values are desirable.

During our literature review, we found multiple examples of metrics where there was a lack of clarity in all five framework dimensions. Thus, it is important to consider all five when defining a metric. However, there may be times when come of the dimensions is not relevant to the metric’s intended use. For example, there could be an instance where something (i.e., the Focus) is being counted, but not engaging in a specific Activity (e.g., number of unread radiology exams; “being unread” could conceivably be the Activity). However, in general, most metrics would benefit greatly from being documented in a way that attempts to engage all five dimensions of FASStR, and every metric’s development effort should consider all five dimensions to determine if potentially relevant.

The examples shown under each of FASStR’s five dimensions in Figure 3 were largely derived from the literature sample we reviewed. These examples offer guidance in the use of each dimension but may not be comprehensive. If a metric’s definition requires expanding these lists, researchers and practitioners should do so and then clearly document its alignment with the FASStR framework.

Table 2 offers illustrative examples of how FASStR can be applied to common metrics in order to clarify their operationalization and intent. While Table 2 is illustrative, an organization might prefer to define a metric that sounds similar to one of these in a markedly different way. In many cases, an organization will need to add significantly more detail, such as when an organization-specific resource, metric, etc. is a necessary element of one or more of the FASStR dimensions.

DISCUSSION

Through an iterative process we propose FASStR, a new framework to help guide those who are designing, developing, and implementing operational metrics in healthcare settings. The framework’s five dimensions attempt to cover all aspects of a metric’s definitional requirements. Developed by an international, multidisciplinary team of subject-matter experts from academia and practice, and based on a thorough literature review, the FASStR framework provides a method by which every metric can be examined to ensure it is thoroughly developed and completely documented. Editors and reviewers could use this framework to evaluate existing and future metrics in submitted papers similar to how research guidelines are used to enhance the quality of reports resulting for medical research. [24]

In addition, practitioners and those who manage health delivery systems should consider using FASStR in their organizations to help ensure metrics are defined and documented in a clear, consistent manner. This could improve employees comprehensive and help to retain organizational memory in the face of staff turnover. We propose that that studies of clinical operations in the medical literature could use the FASStR framework to ensure representation of the relevant dimensions. This would be similar to a checklist for observational cohort clinical research studies.[32] Both research and practice benefit from better metric definitions when large-scale improvement efforts take place. Collaboratives, for example, can be powerful, but if metrics are interpreted and implemented differently across the various institutions participating in the effort, the prospects for undesirable delays and diminished outcomes are increased. If those who design and define metrics consider all five dimensions for each metric, their work will be less ambiguous, more applicable, more generalizable, more reproducible, and, ultimately, more valuable.

Like any framework, FASStR will improve as it gets more used and applied in more and more diverse contexts. Through this improvement, ambiguities can be resolved and best practices will emerge. One clear potential benefit is for the development of a “standard menu” of well-defined metrics appropriate for various healthcare delivery settings (e.g., outpatient clinics, inpatient wards, operating rooms, emergency departments, etc.). As healthcare organizations gravitate towards using a standard set of metrics defined and implemented in the same way, the opportunity for meaningful comparison should increase and improvement accelerate. FASStR fills an important standardization void that has so far vexed the healthcare industry and limited many of its well-intentioned improvement efforts, but only through adoption in both the application and research domains can its potential benefits be fully realized.

CONCLUSION

Healthcare operational metrics are plagued by inconsistency and incompleteness. Through an iterative process of literature review, multidisciplinary expert assessment, and end-user feedback we propose the FASStR framework to address these deficiencies in operational metrics. The FASStR framework fills a gap in standardization necessary to enhance process improvement and research.

FUNDING AND DISCLOSURES

This research was partially funded by the College of Business at James Madison University.

Dr. Ward was funded by an award from the National Heart, Lung, and Blood Institute (K23HL127130)

REFERENCES

1. Agency for Healthcare Research and Quality Notice Number: NOT-HS-14-005. <http://grants.nih.gov/grants/guide/notice-files/NOT-HS-14-005.html>. Date Accessed July 2018.
2. IHI Triple Aim 2012 <http://www.ihl.org/Engage/Initiatives/TripleAim/Pages/default.aspx>. Date Accessed July 2018.
3. Pedhazur EJ, Schmelkin LP Measurement, design, and analysis: An integrated approach. Psychology Press 2013.

4. American Academy of Family Physicians 2019 https://www.aafp.org/news/practice-professional-issues/20190114measurespaper.html?utm_cmpid=aafp&utm_campaign=aafpnews&utm_div=com&utm_mission=ot&utm_prod=news&hootPostID=7f3727545687730cbb03d76966a20cc1 Date Accessed May 2019.
5. National Quality Forum 2012 http://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx Date Accessed July 2018.
6. Ambulatory Care Use and Physician office visits, National Center for Health Statistics, Center for disease control and prevention. <https://www.cdc.gov/nchs/fastats/physician-visits.htm>. Date Accessed July 2018.
7. Robinson LW, Chen RR Scheduling doctors' appointments: optimal and empirically-based heuristic policies. IIE Transactions. 2003;35(3):295–307.
8. Haraden C, Resar R. Patient flow in hospitals: understanding and controlling it better. Frontiers of health services management 2004;20(4):3.
9. LaGanga LR, Lawrence SR. Clinic overbooking to improve patient access and increase provider productivity. Decision Sciences 2007;38(2):251–76.
10. LaGanga LR, Lawrence SR. Appointment scheduling with overbooking to mitigate productivity loss from no-shows. Proceedings of Decision Sciences Institute Annual Conference, Phoenix, Arizona 2007 Nov 17.
11. Wijewickrama A, Takakuwa S. Simulation analysis of appointment scheduling in an outpatient department of internal medicine. Proceedings of the Winter Simulation Conference 2005.
12. Wijewickrama AK. Simulation analysis for reducing queues in mixed-patients' outpatient department. International journal of simulation modelling 2006;5(2):56–68.
13. Dexter F Design of appointment systems for preanesthesia evaluation clinics to minimize patient waiting times: a review of computer simulation and patient survey studies. Anesthesia & Analgesia 1999;89(4):925. [PubMed: 10512266]
14. Benson R, Harp N. Using systems thinking to extend continuous quality improvement. The Quality letter for healthcare leaders 1994;6(6):17–24. [PubMed: 10136742]
15. Babes M, Sarma GV. Out-patient queues at the Ibn-Rochd health centre. Journal of the Operational Research Society 1991;42(10):845–55.
16. Benussi G, Matthews LH, Daris F, Crevatin E, Nedoclan G. Improving patient flow in ambulatory care through computerized evaluation techniques. Revue d'épidémiologie et de sante publique 1990;38(3):221–6.
17. Vissers J, Wijngaard J. The outpatient appointment system: Design of a simulation study. European Journal of Operational Research 1979;3(6):459–63.
18. Vissers J Selecting a suitable appointment system in an outpatient setting. Medical Care 1979;1207–20. [PubMed: 513884]
19. White MB, Pike MC. Appointment systems in out-patients' clinics and the effect of patients' unpunctuality. Medical Care 1964;133–45.
20. White DL, Froehle CM, Klassen KJ. The effect of integrated scheduling and capacity policies on clinical efficiency. Production and Operations Management 2011;20(3):442–55.
21. Kaandorp GC, Koole G. Optimal outpatient appointment scheduling. Health Care Management Science 2007;10(3):217–29. [PubMed: 17695133]
22. Klassen KJ, Yoogalingam R. Improving performance in outpatient appointment services with a simulation optimization approach. Production and Operations Management 2009;18(4):447–58.
23. Denton B, Gupta D. A sequential bounding approach for optimal appointment scheduling. IIE transactions 2003;35(11):1003–16.
24. Johansen M, Thomsen SF. Guidelines for reporting medical research: a critical appraisal. International scholarly research notices 2016;1 10.1155/2016/1346026
25. Chan K, Li W, Medlam G, et al. Investigating patient wait times for daily outpatient radiotherapy appointments (a single-centre study). Journal of Medical Imaging and Radiation Sciences 2010;41(3):145–51. [PubMed: 31051868]
26. Partridge JW. Consultation time, workload, and problems for audit in outpatient clinics. Archives of disease in childhood 1992;67(2):206–10. [PubMed: 1543380]

27. Klassen KJ, Yoogalingam R. Strategies for appointment policy design with patient unpunctuality. *Decision Sciences* 2014;45(5):881–911.
28. Cayirli T, Veral E, Rosen H. Assessment of patient classification in appointment system design. *Production and Operations Management* 2008;17(3):338–53.
29. Belter D, Halsey J, Severtson H, et al. Evaluation of outpatient oncology services using lean methodology. *Oncology nursing forum* 2012;39(2):136–40. [PubMed: 22374487]
30. Lenin RB, Lowery CL, Hitt WC, Manning NA, Lowery P, Eswaran H. Optimizing appointment template and number of staff of an OB/GYN clinic—micro and macro simulation analyses. *BMC health services research* 2015;15(1):387. [PubMed: 26376782]
31. Borgman NJ, Vliegen IM, Boucherie RJ, Hans EW. Appointment scheduling with unscheduled arrivals and reprioritization. *Flexible services and manufacturing journal* 2018;30(1–2):30–53.
32. Stroke Statement. STROBE Initiative, <https://www.strobe-statement.org/index.php?id=available-checklists>. Accessed on December 19, 2019.

TAKE-AWAY POINTS

When measurements are dissimilar or inadequately defined, comparison of operational performance and interpretation of research is impaired. Through an iterative process of literature review, expert assessment, and collaborative design, we propose a unifying framework for healthcare stakeholders and researchers to assist in the development, definition, and evaluation of operational performance measures. We propose the FASStR framework: Focus, Activity, Statistic, Scale type, and Reference. The FASStR framework can be used for operational and research purposes to provide an objective, systematic approach for researchers, practitioners, and managers to ensure metrics are defined and documented in a clear, consistent manner.

- Poorly defined measurement impairs inter-institutional comparison, interpretation of results, and process improvement in healthcare operations.
- Inconsistency and incomplete performance metrics are common deficiencies in healthcare operations metrics
- The FASStR framework provides a systematic approach to metric development, definition, and evaluation

Figure 1.
Peer-reviewed publications on outpatient healthcare operations, by year, 1988–2017

Figure 2. How different definitions can result in different operational outcomes: a) “Best” scheduled appointment start times when *Total* patient waiting is used; b) “Best” scheduled appointment start times when *Average* patient waiting is used

Figure 3.
The FASStR framework for operational performance metrics in healthcare

Table 1.

A sample of different definitions for “patient waiting time” presented in the literature

Provided definition	Deficit
Average (expected) patient waiting time in minutes for each of the patients who show up over 10,000 replications.[9]	Start-end times of the waiting interval are not specified. Also, what is the length of a replication?
Mean patient waiting time in the consultation room in minutes. Also calculated mean waiting per consultant, 5/10/25 th percentiles.[11]	No formal definition, no time frame, start-end times of the waiting interval not specified.
Weighted average of patient waiting time (in the consultation room including service time) in minutes. Also calculated mean waiting per consultant, 5/10/25 th percentiles.[12]	No time frame. Treatment or service time is included in waiting time.
Mean patient wait time calculated from when a patient arrives at the pre-anesthesia evaluation clinic to when his or her consultation begins over 5,000 simulated clinic sessions.[13]	No units
Mean and standard deviation of patient waiting time in minutes (calculated from time of appointment to patient entry to treatment room) over 10 clinic days, the goal is to reduce patient wait time to 20 min. [25]	None.
Mean patient waiting time in minutes (calculated from time of appointment to the start of consultation). Also calculated 10 th & 90 th percentiles, range of the mean and maximum wait, frequency of the number of patients with over 30 and 60 min. wait across 33 clinics.[26]	No time frame. Data obtained from 33 clinics by questionnaire, thus no details on how data were originally collected by each clinic.
Mean patient waiting time in minutes (calculated from (i) time of appointment and (ii) time of arrival to service start time) over 10,000 simulated clinic sessions.[27]	None.
Mean patient waiting time (calculated from time of appointment for early/punctual patients and from time of arrival for late and walk-in patients) in minutes over 10,000 simulated clinic sessions. Also calculated the standard deviation of patient wait times as a measure of fairness, and the percentage of scheduled patients seen within 30 minutes of their appointment times (UK standard 75%).[28]	None.
Average overall patient wait time in minutes from arrival to service start (chemotherapy infusion).[29]	No time frame.
Average wait time of a patient for the doctor in the exam room in minutes over 30 simulated days Average total wait time of a patient for the resources at various locations of the clinic in minutes over 30 simulated days.[30]	Start time of the waiting intervals not specified.
Percentage of unscheduled (urgent) patients that are seen before their due date (set as 2 hours); while minimizing the maximum wait times for scheduled patients. The goal is 90%.Also calculate the expected waiting times by appointment slot throughout the day as a measure of fairness.[31]	No reference to how wait times are calculated for unscheduled vs. scheduled (from time of arrival or appointment)

Table 2.

Selected examples of operational metrics defined using the FASStR framework

Metric	Focus	Activity	Statistic	Scale type	Reference(s)
Physician non-contact time per clinic	Physician	A physician is busy with tasks other than actively caring for or interacting with a patient	Mean	Time (minutes)	Goal of 60 minutes per clinic session
Door-to-physician time	Patient	Time between initial arrival at clinic registration and physician entering patient's exam room	Mean for seen by a physician during a clinic session	Time (minutes)	Previous clinic
Percent on-time patient arrivals	Patient	Arriving at the clinic no later than 5 minutes after appointment time	% of all patients scheduled during the clinic session	Percentage	Organizational goal of 100%
Exam room utilization	Exam room	Occupied for patient care	Mean percentage of clinic time (minutes), averaged over past month	Percentage and minutes	Previous month's utilization
Third-next appointment availability	Clinic schedule	Availability	Mean # days before the third-next available appointment slot, averaged over 8 appointment types	Time (days)	Peer benchmark
Patient Satisfaction	Patient	Satisfaction	Percentage of patients rating "7" on 7-point scale	Likert	Industry benchmark of 90%
Continuity of care	Patient	Clinic visit seeing preferred physician or provider	Percentage of visits in last 12 months	Percentage	Organizational goal of 80%
Patient waiting	Patient	Waiting, from registration to placement in an exam room	Mean	Time (minutes)	Organizational goal of 30 minutes; previous period