Improving Benchmarking by Using an Explicit Framework for the Development of Composite Indicators of Pediatric Care Quality

Jochen Profit, MD, MPH[a,b,c], , *Katri Typpo, MD, MPH*[a,b,c], *Sylvia J. Hysong, PhD*[b,c], *LeChauncy Woodard, MD, MPH*[b,c], *Michael A. Kallen, PhD, MPH*[d], *Laura A. Petersen, MD, MPH*[b,c]

[a] Department of Pediatrics, Baylor College of Medicine; Texas Children's Hospital; Houston, TX, USA

[b] Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, TX, USA

[c] Houston Veterans Affairs (VA) Health Services Research and Development Center of Excellence, Health Policy and Quality Program, Michael E. DeBakey VA Medical Center; Houston, TX, USA

[d] The University of Texas M D Anderson Cancer Center, Department of General Internal Medicine, Ambulatory Treatment and Emergency Care, Houston, TX

Corresponding Author: Jochen Profit, MD, MPH, Houston VA HSR&D Center of Excellence (152), 2002 Holcombe Boulevard, Houston, TX 77030, phone: 713.794.8609, fax: 713.748.7359, email: profit@bcm.edu

The authors have no financial relationships relevant to this article to disclose.

GRANT SUPPORT: This project is supported by NICHD K23 HD056298-01 (PI Jochen Profit, MD, MPH) and in part by the Houston VA Health Services Research & Development

(HSR&D) Center of Excellence (HFP90-020). Dr. Petersen is a recipient of the American Heart Association Established Investigator Award (Grant number 0540043N). Dr. Hysong is a recipient of a VA HSR&D Career Development Award (CD2-07-0181).

Email addresses:

JP: profit@bcm.edu

KT: ktyppo@bcm.edu

SH: hysong@bcm.edu

LW: lwoodard@bcm.edu

MK: makallen@mdanderson.org

LP: laurap@bcm.edu

KEY WORDS: quality of care, composite indicator, benchmarking

A Framework for the Development of Composite Indicators of Pediatric Care Quality

ABSTRACT

Background: The measurement of health care provider performance is becoming more widespread. Physicians have been guarded about performance measurement, in part because the methodology for comparative measurement of care quality is underdeveloped. Comprehensive quality improvement will require comprehensive measurement, implying the aggregation of multiple quality metrics into composite indicators.

Objective: To present a theoretical framework to develop comprehensive, robust, and transparent composite indicators of pediatric care quality and to highlight aspects specific to quality measurement in children.

Methods: We reviewed the scientific literature on composite indicator development, on health systems, and quality measurement in the pediatric health care setting. Frameworks were selected for explicitness and applicability to a hospital-based measurement system.

Results: We synthesized various frameworks into a comprehensive model for the development of composite indicators of pediatric quality of care. Among its key premises the model proposes identifying structural, process, and outcome metrics for each of the Institute of Medicine's six domains of quality (safety, effectiveness, efficiency, patient-centeredness, timeliness, and equity) and presents a step-by-step framework for embedding the quality of care measurement model into composite indicator development.

Conclusions: The framework presented offers researchers an explicit path to composite indicator development. Without scientifically robust and comprehensive measurement of the quality of health care, performance measurement will ultimately fail to achieve its quality improvement goals.

BACKGROUND

In recent years, composite indicators of care quality have been used more widely to measure and track provider performance in adult medicine.(1-7) In pediatrics, similar interest exists in performance measurement, but, to our knowledge, composite indicators of quality have not yet been developed. This may be because composite indicators in health care have come largely as a by-product of so called "value-based purchasing" initiatives, where payers reimburse providers based on comparative performance (benchmarking).(8-10) To date, such reimbursement policies have been less prevalent in pediatrics (9;11); nevertheless, payers and policy makers are moving ahead with benchmarking activities.(12;13)

Composite indicators can provide global insights and trends on quality not just for external benchmarking against other providers or institutions, but importantly also to facilitate the tracking of internal quality improvement efforts within institutions. However, the process of developing composite indicators is complex, and the editorial choices required of developers may significantly influence performance ratings.(14)

The unique contribution and purpose of this paper is to (a) advocate for using composite indicators as an approach to measure quality in pediatrics, and (b) present a framework for the development of composite indicators based on a combination of previously presented frameworks on both quality measurement and composite indicator development. While we will focus on aspects important to pediatrics, we believe that the application of this framework provides a comprehensive roadmap for the continuous improvement of quality measurement for all populations.

COMPOSITE INDICATORS OF QUALITY

Composite indicators of quality combine multiple metrics of quality into an aggregate score. Table 1 (adapted from Nardo(15)) summarizes the advantages and disadvantages of using composite indicators, regardless of field or purpose. We will facilitate a discussion of the advantages and disadvantages of composite indicators focusing on their two probable uses, benchmarking and quality improvement.

Composites for Benchmarking

Benchmarking of providers based on only one or a few indicators of quality infers a strong correlation of performance across metrics of quality and assumes that measured aspects of quality are informative of overall quality of care. In other words the structure and processes of care are systematically interconnected to yield high performance in many aspects of care. However, this has not been found in the extant literature. Several articles have highlighted weak correlations among metrics of quality.(16;17) Thus isolated metrics of quality are unlikely to be informative of quality of care, whereas composite indicators are better suited to reflect an overall construct of quality.

Composite indicators of quality are communicable to many stakeholders and can be leveraged to induce competition on quality. Payers of health care increasingly employ these measurements to inform and direct patients' choice of providers through selective contracting. Patients may gain from provider competition for quality and through the ability to make informed health care choices. While to date there is little evidence that benchmarking information affects patient choice of provider(18), consumer attitudes may change as the quality

and dissemination formats of quality information improve. However, any benefit to patients is dependent on the accuracy of classifying providers as superior or inferior. Given the variability in methods and quality of existing indicators, there is possibly significant misclassification of providers as outliers.(19)

Composite indicators function by providing a simplified representation of the underlying quality of care construct. In fact, simplification is their main appeal. There is a danger, however, that overly simplistic policy messages derived from composites may be misleading or misused to support narrow agendas. Moreover, if indicators lack scientific soundness, transparency, and content validity to providers being measured, they are unlikely to produce desired improvements in patient health status. Some of these dangers can be countered by using dissemination formats that convey results accurately while avoiding oversimplification (such as the ability to "drill down" into individual components of the composite), and by making the process of indicator development explicit and transparent to all stakeholders. Nevertheless, it is likely that composites used for benchmarking will be subject to methodological and political challenge from providers disagreeing with results.

Composites for Quality Improvement

Used for quality improvement, an accurate and reliable composite index of quality may induce multi-dimensional, systems-based improvement activities that may spread (20) throughout the care service and the institution at large, *i.e.* resulting in systems-based improvements not only in the care area targeted for improvement but also in related care areas.

Composites can help providers translate a bewildering wealth of information into quality action and track action effects throughout the care delivery system. To illustrate, the Vermont

Oxford Network tracks the quality of health care delivery of over 700 neonatal intensive care units worldwide, with clinically rich information available for many processes and care outcomes. However, it may be difficult for neonatal intensive care providers to translate large data volumes into effective quality improvement efforts. In practice, providers often focus on improving a performance outlier. Yet, targeting individual quality metrics may lead to piecemeal rather than system-based efforts in quality improvement. Indeed, without composite indicators, the effect of improvement of one metric on others and on overall quality may be difficult to monitor. It is important to emphasize here that we do not advocate replacing the measurement of individual metrics of quality with a composite indicator. The composite merely summarizes the information contained in the individual metrics and makes that information more digestible.

Recognizing that there are numerous editorial choices in the development of composite indicators and that quality of care can be defined in overly simplistic ways, we propose a composite-based approach to measuring pediatric care quality by combining the Organization for Economic Cooperation and Development (OECD) composite development methodology (15) and Profit et al.'s framework for measurement of pediatric care quality (9).

Development of composite indicators

As do other organizations, the OECD has significant institutional expertise in developing, applying, and evaluating composite indicators; it has, in fact, published guidelines for composite indicator development.(15;21;22) These guidelines have begun to be used in other settings of health care.(23) What differentiates the OECD from other organizations is its highly explicit,

transparent, and evaluative approach to composite indicator development. Their proposed methods promote internal and external statistical and methodological consistency and offer users choices of building blocks at each step in composite indicator construction, tailored to the task at hand.

Table 2 shows the ten steps suggested for developing composite indicators.(15) We present a here a theoretical example of composite score development for pediatric intensive care unit (PICU) quality that follows the OECD system and combines multiple frameworks of quality.

Example: Developing a PICU Quality Composite Indicator

Step 1 - Framework: We base the framework for a PICU indicator on the work of Arah, (24) Roberts(25), the Institute of Medicine (IOM)(26), and Donabedian(27) (see Figure 1). Details of this framework have been described elsewhere.(9) In brief, the figure models a patient's path through the health care system and highlights opportunities and challenges for measurement. The model emphasizes innate and external modifiers of health that determine baseline illness severity and that should be addressed via risk adjustment or risk stratification. Quality of health care measurement combines the frameworks of the IOM and Donabedian, resulting in a quality matrix (see Table 3). Metrics within the matrix can be combined to form a composite indicator of quality. Such an indicator becomes an intermediate outcome metric, which can then be used to assess the effect of new health policies or changes in medical care on long-term health outcomes.

We propose filling the quality matrix with disease- or disease category-specific metrics of quality to create a balanced scorecard of overall quality of care and promote the goal of ensuring that providers are responsive to the quality expectations of all stakeholders, including payers and patients. In many areas of medicine, available metrics may span several domains of quality, may share a cell with other metrics, or may not exist for certain cells of the matrix; the latter measurement state clearly indicates the need for future research. For example, the absence of equity metrics in Table 3's matrix is of note and could be addressed through further research on equity reports.(28)

Step 2 - Metric Selection: Selection should rely on a rigorously developed process. Favourable metric characteristics include: relevance (i.e., relevant domains of care); validity (reflecting the desired measurement construct); reliability (avoiding undue systematic bias around measurement point estimates); malleability (giving providers identifiable control over their performance); interpretability (inducing reasonable action plans); timeliness (improving the effect of feedback); and availability (easy to collect). In our example, the Pediatric Data Quality Systems (Pedi-QS) Collaborative Measures Workgroup (www.pediqs.com) is a joint consensus panel formed by the National Association of Children's Hospitals and Related Institutions, Child Health Corporation of America, and Medical Management Planning tasked with recommending pediatric quality metrics to the Joint Commission. In 2005, the Work Group recommended eight process and outcome quality metrics for use in the PICU, which we have placed into the matrix (see Table 3).

Step 3 - Initial Data Analysis: In this step the data are prepared for analysis. Consideration should be given to the exclusion of outlier data points, such that resulting performance ratings are not unduly influenced by extreme values. In addition, the data need to

be uniform in their directionality. For example, a high ventilator associated pneumonia (VAP) rate indicates poor quality, but a high level of compliance with VAP prevention practices indicates the opposite. In the composite, both metrics have to point in the same direction.

Step 4 - Missing Data: Treatment of missing data may influence hospital performance. For example, if missing data are excluded, or if data are imputed using a hospitals' average performance, then hospitals have an incentive to game the system, by not providing data on patients with poor outcomes. On the other hand, if metrics are binary (yes/no), assuming the adverse outcome to have occurred in case of missingness may encourage complete record keeping. Methods for imputing missing data have been described elsewhere.(29;30) An important consideration here is a judgement regarding whether data are missing at random, or whether missingness signals differences in underlying case mix between institutions (*e.g.,* missing VAP rate data may not be randomly distributed but reflect poor record keeping and/or poor outcomes).

Step 5 - Normalization: From the selected metrics, a base case composite is constructed using a combination of *a priori* agreed on methods. Metrics with different units and scales cannot be aggregated before being transformed to a common scale (normalization). Of the many existing choices for normalization, ranking and assignment to a categorical scale (*e.g.,* star rating) are used most commonly; other choices (e.g., standardization; distance to a reference metric) should also be considered.

Step 6 - Weighting and aggregation: This step is crucial in the development of a composite indicator, as decisions about the attribution of weights to metrics as well as metric aggregation may significantly influence performance assessment results. Weights must reflect the importance, validity, reliability, and malleability of individual metrics; metrics with

contradictory quality signals (*e.g.,* safe and effective, but not efficient) must be weighted appropriately.

Weighting. The two basic methods used to arrive at metric weights are statistical (*e.g.,* principal component analysis, factor analysis, multivariate techniques) and participatory methods (variations on eliciting expert opinion). Note that equal weighting does not imply an absence of weights: Under this approach each metric is given a weight of one. An equal weighting scheme may introduce an element of double counting if two metrics prove to be highly correlated (*e.g.,* VAP rates and VAP prevention practices).

Benefits of the statistical approach to weighting include its relative fairness and its freedom from bias. In contrast to the participatory approach, its primary disadvantage is that resultant weights may lack face validity.

Aggregation. In this phase the metrics are put together to form the composite indicator. The primary decision involved in choosing an aggregation method hinges on whether providers should be allowed to compensate for poor performance in one metric with superior performance in another. There are three principal choices: full compensation (additive), partial compensation (multiplicative), and no compensation (non-compensatory).

Because of its simplicity, the additive aggregation technique is most widely used. However, it implies full compensability and may therefore result in a biased composite indicator, with an error of dimension and direction not easily determined.

Multiplicative aggregation allows for partial compensability, which makes it more difficult to offset a bad indicator with a good one. This is in line with our concept of quality in which a quality performance metric is intended to foster superior quality throughout domains of care and not promote trade-offs between areas of strength and weakness.

Non-compensatory methods demand achieving excellence in all metrics of quality or at least achieving minimum standards of quality, thereby promoting multi-dimensional improvement efforts. One variant of non-compensatory methods, the "All-or-None Measurement" approach, has been recently propagated as a means to foster excellence in quality.(31) However, it has been argued that this particular approach is likely imprecise and may provide perverse incentives, such as promoting treatment irrespective of how small the potential benefit and how great the patient burden or risk.(32) Disadvantages of other non-compensatory methods include statistical complexity and non-discrimination of the magnitude of differences in indicator scores.

Step 7 - Uncertainty and Sensitivity Analysis: The effect of subjective choices and chance variation in the underlying data on provider performance can be modelled in higher order Monte Carlo experiments. The importance of sensitivity analysis cannot be overemphasized. Composite indicators must be sufficiently robust in discriminating outliers on both extremes of performance in order to enhance their usefulness and engender provider trust. Thus, stability of results in sensitivity analysis provides an important quality check of the composite indicator as well as of the underlying framework and data.

Step 8 - Links to other metrics: If composite indicators of quality for related pediatric populations existed, these indicators could be linked to the indicator developed here. Composite indicators can thereby "grow" in a cross-sectional and longitudinal manner. For example, composite indicator of quality of related specialties whose patients frequently require PICU care (e.g., pulmonology) could be combined with a PICU indicator and thus provide a better image of quality for specific patient populations across disease episodes.

Step 9 - De-construction: For presentation purposes, the composite indicator can be deconstructed to reveal contributions from individual metrics to overall performance.

Step 10 - Presentation and Dissemination: Presentation formats can be user-friendly, such as charts that include metrics of uncertainty (*e.g.*, confidence intervals). Electronic publications can link to further detail on individual metrics.

Pediatric Aspects of Composite Indicator Development

Several aspects of composite indicator quality of care measurement feature prominently for pediatric care and are worth discussing. The number of yearly admissions for pediatric patients is smaller than that for adults, making sample size a significant issue.(33) Metric development may therefore require ongoing data collection over several years and across multiple institutions. The aggregation of several metrics into a composite indicator may alleviate this problem, in that information from multiple quality metrics can be combined and thereby increase the power to detect a quality signal (Step 6); however, this is an empirical question that needs to be addressed in future research.

Health related quality of life is an important outcome of care quality (Step 2). However, children under the age of 5 are typically unable to reliably answer quality of life questions, so caregiver proxy opinion has been used as a reasonable substitute.(34;35) However, because parental assessment of their children's quality of life may be positively biased (36), health related quality of life ratings may need to be obtained from health professionals or the general public. Recommendations for cost-effectiveness analysis favour the general public's perspective (37); yet such ratings are strongly influenced by responder personal experience with health status (38) and may also reflect the availability and quality of chronic care management and the degree

of health system integration. Future research should try to address these important methodological gaps that remain in the measurement of health related quality of life of young children.(39) Until such research, the uncertainty in quality of life ratings should be reflected in lower relative weightings, so as to not threaten the external validity of the composite indicator.

Much of the job of pediatric health professionals is to prevent illness or illness exacerbation. Therefore, metrics of primary prevention should be given particular consideration during the metric selection process (Step 2). Childhood illness may potentially lead to long-lasting, even devastating, adverse outcomes, permanently altering children's developmental trajectories.(40) Thankfully, high quality rehabilitation and educational services can support children's unique adaptation to injury, enabling them to reach full potential.(41) This implies that measurement of health care quality should emphasize longitudinal linkages to health outcomes over time (Step 8), which will provide an opportunity for validation of the composite indicator and offer opportunities for further linkage to additional social well being outcomes to help assess the quality of larger societal systems, including social support and educational systems. Currently, few such metrics exist, and much research will be needed to develop them.

CONCLUSION

Composite indicators are being more widely used to measure health care provider performance and may have benchmarking or quality improvement purposes. However, failure to adopt rigorous indicator development methods will undermine their ultimate usefulness in improving quality and instead encourage physician perception that performance measurement is unreliable and inaccurate.(42-44) Although individually published elsewhere, the combination of OECD's performance metric development methodology with Profit et al.'s quality matrix

framework results in a unique approach for quality measurement that is fair, scientifically sound, and promotes the all-important provider buy-in. Future work should evaluate the feasibility and value of this approach in helping make accurate benchmarking and quality improvement decisions.

Author Contributions

JP and LP led the conceptualization, design, writing and revision of the manuscript. KT contributed to adaption of the content to the pediatric intensive care unit setting. MK contributed to the composition of a revised framework for composite indicator measurement and adaptation of the methods to the health care setting. KT, SH, LW, and MK contributed to writing and revision of the manuscript. JP is guarantor of the paper.

Reference List

[1] Premier Hospital Quality Incentive Project. http://www.premierinc.com/quality-safety/tools-services/p4p/hqi/index.jsp

[2] Schoen C, Davis K, How SKH, Schoenbaum SC. U.S. health system performance: a national scorecard. *Health Aff* 2006;w457-w475.

[3] Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, et al. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med* 2007;365:486-96.

[4] Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. *JAMA* 2003;289:305-12.

[5] Epstein AJ. Do cardiac surgery report cards reduce mortality? Assessing the evidence. *Med Care Res Rev* 2006;63:403-26.

[6] Grossbart SR. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Med Care Res Rev* 2006;63:29S-48.

[7] Petra E, Varughese P, Epifania L, Buneo L, Scarfone K. Use of quality index tracking to drive improvement in clinical outcomes. *Nephrology News and Issues* 2006;20:67-83.

[8] Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006;145:265-72.

[9] Profit J, Zupancic JA, Gould JB, Petersen LA. Implementing pay-for-performance in the neonatal intensive care unit. *Pediatrics* 2007;119:975-82.

[10] Profit J, Petersen LA. Pay for performance is growing up. *Arch Pediatr Adolesc Med* 2007;161:713-4.

[11] Freed GL, Uren RL. Pay-for-performance: An overview for pediatrics. *Journal of Pediatrics* 2006;149:120-4.

[12] Kuhmerker K, Hartman T. Pay-for-performance in state Medicaid programs: a survey of state Medicaid directors and programs. New York, NY: The Commonwealth Fund; 2007. Accessed on

[13] The Leapfrog Group. The Leapfrog Group fact sheet. www.leapfroggroup.org/leapfrog-factsheet

[14] Jacobs R, Goddard M, Smith PC. How robust are hospital ranks based on composite performance measures? *Med Care* 2005;43:1177-84.

[15] Nardo M, Saisana M, Saltelli A, Tarantolo S, Hoffman A, Giovanini E. Handbook on constructing composite indicators: methodology and user guide. Paris, France: OECD Publishing; 2005.

[16] Rosenthal GE. Weak associations between hospital mortality rates for individual diagnoses: implications for profiling hospital quality. *Am J Public Health* 1997;87:429-33.

[17] Wilson IB, Landon BE, Marsden PV, Hirschhorn LR, McInnes K, Ding L, et al. Correlations among measures of quality in HIV care in the United States: cross sectional study. *BMJ* 2007;335:1085-91.

[18] Schauffler HH, Mordavsky JK. Consumer reports in health care: do they make a difference? *Annual Review of Public Health* 2001;22:69-89.

[19] Williams SC, Koss RG, Morton DJ, Loeb JM. Performance of top-ranked heart care hospitals on evidence-based process measures. *Circulation* 2006;114:558-64.

[20] Asch SM, McGlynn EA, Hogan MM, Hayward RA, Shekelle P, Rubenstein L, et al. Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. *Ann Intern Med* 2004;141:938-45.

[21] Kelley ET, Hurst J. Health Care Quality Indicator Project - conceptual framework. Paris, France: OECD Publishing; 2006. Accessed on 9-8-2006.

[22] Mattke S, Epstein AM, Leatherman S. The OECD Health Care Quality Indicators Project: history and background. *Int J Qual Health Care* 2006;18:1S-4.

[23] Brand DA, Saisana M, Rynn LA, Pennoni F, Lowenfels AB. Comparative analysis of alcohol control policies in 30 countries. *PLoS Medicine* 2007;4:e151.

[24] Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD Health Care Quality Indicators Project. *Int J Qual Health Care* 2006;18:5S-13.

[25] Roberts MJ, Hsiao W, Berman P, Reich MR. Getting Health Reform Right: A Guide To Improving Performance And Equity. Oxford: Oxford University Press; 2003.

[26] Institute of Medicine. Crossing The Quality Chasm: A New Health System For The 21st Century. Washington, DC: National Academy Press; 2001.

[27] Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;44:166S-206.

[28] Weinick RM, Flaherty K, Bristol SJ. Creating equity reports: a guide for hospitals. The Disparities Solutions Center, Massachusetts General Hospital; 2008. Accessed on

[29] Brick JM, Kalton G. Handling missing data in survey research. *Stat Methods Med Res* 1996;5:215-38.

[30] Duffy ME. Handling missing data: a commonly encountered problem in quantitative research. *Clin Nurse Spec* 2006;20:273-6.

[31] Nolan TP, Berwick DMM. All-or-none measurement raises the bar on performance. *JAMA* 2006;295:1168-70.

[32] Hayward RA. All-or-nothing treatment targets make bad performance measures. *Am J Manag Care* 2007;13:126-8.

[33] Chien AT, Dudley RA. Pay-for-performance in pediatrics: proceed with caution. *Pediatrics* 2007;120:186-8.

[34] Varni JW, Limbers CA, Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL 4.0 Generic Core Scales. *Health Qual Life Outcomes* 2007;5:2.

[35] Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. *Health Qual Life Outcomes* 2007;5:1.

[36] Chesney M, Lindeke L, Johnson L, Jukkala A, Lynch S, Disch J, et al. Comparison of child and parent satisfaction ratings of ambulatory pediatric subspecialty care. *J Pediatr Health Care* 2005;19:221-9.

[37] Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness In Health And Medicine. Oxford, England: Oxford University Press; 1996.

[38] Streiner DL, Saigal S, Burrows E, Stoskopf B, Rosenbaum P. Attitudes of parents and health care professionals toward active treatment of extremely premature infants. *Pediatrics* 2001;108:152-7.

[39] Zwicker JG, Harris SR. Quality of life of formerly preterm and very low birth weight infants from preschool age to adulthood: a systematic review. *Pediatrics* 2008;121:e366-e376.

[40] Halfon N, Hochstein M. Life course health development: an integrated framework for developing health, policy, and research. *Milbank Q* 2002;80:433-79, iii.

[41] McCormick MC, Brooks-Gunn J, Buka SL, Goldman J, Yu J, Salganik M, et al. Early intervention in low birth weight premature infants: results at 18 years of age for the Infant Health and Development Program. *Pediatrics* 2006;117:771-80.

[42] Casalino LP, Alexander GC, Jin L, Konetzka RT. General internists' views on pay-for-performance and public reporting of quality scores: a national survey. *Health Aff* 2007;26:492-9.

[43] Beckman H, Suchman AL, Curtin K, Greene RA. Physician reactions to quantitative individual performance reports. *Am J Med Qual* 2006;21:192-9.

[44] Young GJ, Meterko M, White B, Bokhour BG, Sautter KM, Berlowitz D, et al. Physician attitudes toward pay-for-quality programs: Perspectives from the Front Line. *Med Care Res Rev* 2007;64:331-43.

Tables

Table 1. Advantages and Disadvantages of Composite Indicators	
Advantage	**Disadvantage**
• Facilitate communication with other stakeholders and promote accountability • Summarize complex issues for decision-makers • Facilitate benchmarking • Assess progress over time • Induce innovation in quality improvement	• Provide misleading messages about quality if poorly constructed or misinterpreted • Lead to simplistic policy conclusions • Can be misused, if the construction process is not transparent and lacks sound statistical or conceptual principles • Selection of metrics and weights can be challenged by other stakeholders

Table 2. Developing a Composite Indicator	
Step	**Description**
1	Developing a theoretical framework
2	Metric selection
3	Initial data analysis
4	Imputation of missing data
5	Normalization
6	Weighting and aggregation
7	Uncertainty and sensitivity analysis
8	Links to other metrics
9	De-construction
10	Presentation and dissemination

Table 3. Quality matrix for a pediatric intensive care unit quality index						
	Safe	**Effective**	**Efficient**	**Pt-centred**	**Timely**	**Equitable**
Structure				Nurse-to-patient ratio	Intensivist in house 24 hours a day	
Process	*Medication Safety Practice, Central line infection prevention practice,* VAP prevention practices	*Review of unplanned readmissions*		*Pain assessment on admission, Periodic pain assessment*	Time to receive antibiotics for sepsis	
Outcome	VAP rate, BSI rate, UTI rate, Unplanned extubation rate	*SMR, Unplanned readmission rate*	*Severity adjusted LOS*		Failed extubation rate	
Pt - patient, VAP - ventilator associated pneumonia, BSI - blood stream infection, UTI - urinary tract infection, LOS - length of stay, SMR - standardized mortality ratio. The *italicized items* are the eight core metrics in Pedi-QS report. The other items were initially rejected either because of lack of evidence or difficulty in measurement.						

Figure Legends

Figure 1. Theoretical Framework for Measuring Quality of Care. Solid arrows indicate interactions; dotted arrows indicate potential use of composite indicator to measure health care delivery, predict health status and inform health policy at the health systems and societal level.

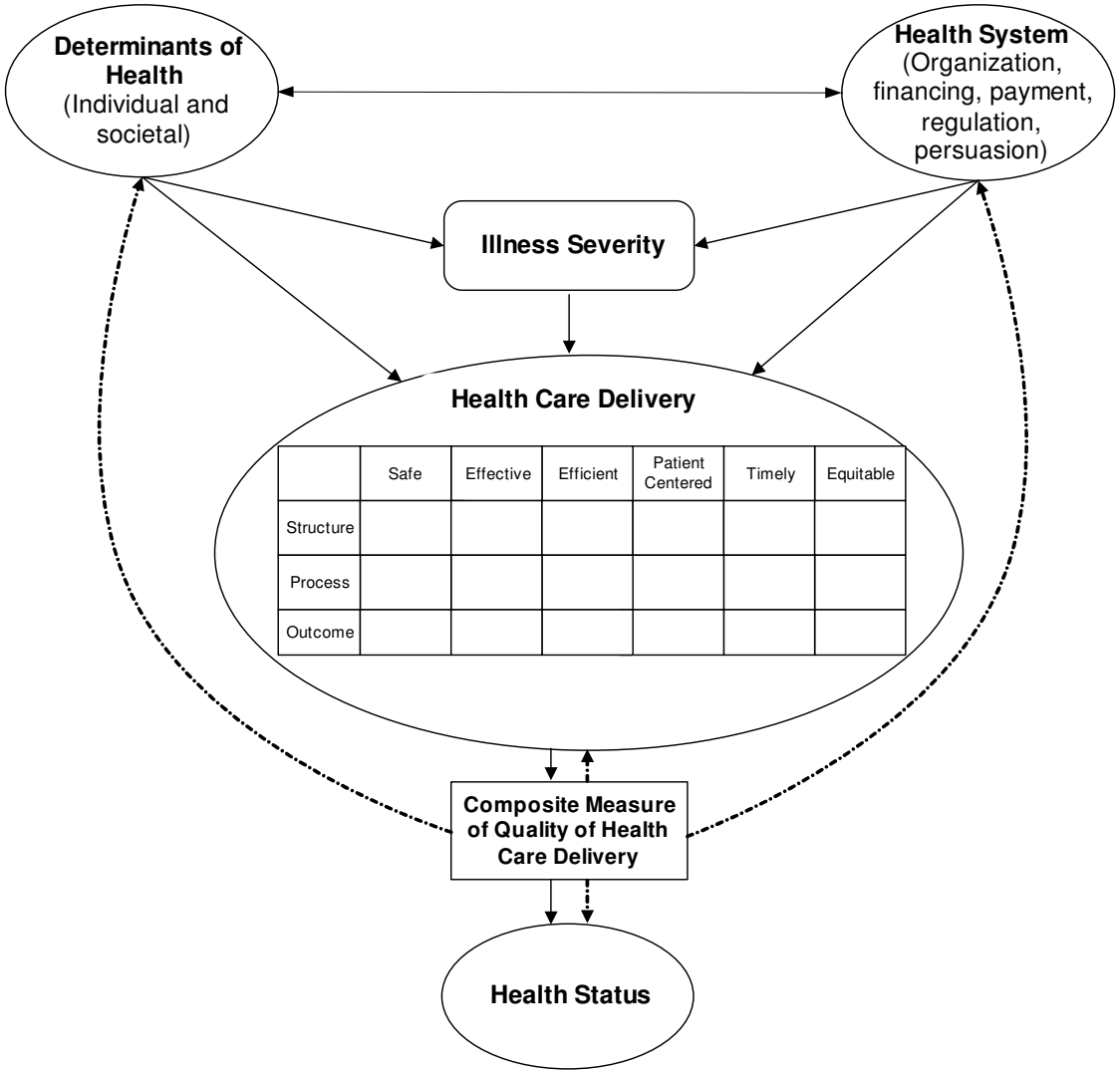
Determinants of Health (Individual and societal)
Health System (Organization, financing, payment, regulation, persuasion)
Illness Severity
Health Care Delivery
Safe
Effective
Efficient
Patient Centered
Timely
Equitable
Structure
Process
Outcome
Composite Measure of Quality of Health Care Delivery
Health Status

Figure 1