Distribution Agreement

Signature:

______________________________ _______________

Shabnam Jain, MD Date

IMPACT OF PHYSICIAN FEEDBACK ON EMERGENCY DEPARTMENT RESOURCE USE, QUALITY AND EFFICIENCY

By

Shabnam Jain
MPH

CMPH

____ [Chair's signature]
[David Howard, PhD]
Committee Chair

____ [Member's signature]
[Brent Johnson, PhD]
Committee Member

IMPACT OF PHYSICIAN FEEDBACK ON EMERGENCY DEPARTMENT RESOURCE USE, QUALITY AND EFFICIENCY

By

Shabnam Jain

MD
Lady Hardinge Medical College
1984

Thesis Committee Chair: David Howard, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Health Outcomes
2012

Impact of Physician Feedback on Emergency Department Resource use, Quality and Efficiency

Background: Emergency Departments (ED) carry a perception of excessive diagnostic testing and therapies. Variation in physician practice is widely prevalent and highlights an opportunity for quality improvement and cost containment. Monitoring resources used in the management of common pediatric ED conditions has been suggested as an ED quality metric.

Objectives: (1) To develop a tool for comprehensive feedback to ED physicians on their practice patterns relative to peers, and (2) To evaluate the impact of physician feedback on ED resource use, quality and efficiency

Methods: Data on resource use by physicians were extracted from electronic medical records at 2 tertiary pediatric EDs for 4 common conditions in mid-acuity: 1) Fever 2) Head Injury 3) Respiratory illness, 4) Gastroenteritis. Condition-relevant resource use was tracked for Lab tests (blood count, chemistry, CRP), Imaging (chest X-ray, abdominal X-ray, head CT scan, abdominal/pelvic CT scan), intravenous fluids, parenteral antibiotics, and intravenous ondansetron; hospital admission was tracked for all 4 conditions. The outcome measure was ED Length of stay (LOS) and the balancing measure was 72-hr return to ED (RR).
Scorecards were constructed using box plots showing physicians their practice patterns relative to peers. Blinded scorecards were distributed quarterly for 5 quarters. Pre- Post-intervention analysis was performed with Sep 1, 2010 as the intervention date. Fisher-exact and Wilcoxon Rank sum tests were used for overall impact. Trend analysis was conducted to account for physician-specific random effects in clustering, patient-level covariates, and time trends.

Results: We analyzed 48,538 patient visits (21,612 Pre and 26,926 Post) seen by 121 physicians (mean 401 patients/physician). Overall, reduction was noted in use of abdominal/pelvic CT scans, head CT scans, chest X-rays, IV antibiotics and IV Ondansetron ($p<0.001$ for all). Hospital admission rate decreased from 7.4% to 6.7% ($p<0.001$). ED LOS decreased from 129 to 125 min ($p<0.001$). 72-hr RR changed from 2.2% to 2.0%. Significant change in trends was noted for admission to hospital ($p<0.05$), use of lab tests ($p<0.001$), and use of IV antibiotics ($p<0.05$).

Conclusion: Our study shows reduction in resource use for commonly seen conditions in the pediatric ED after providing ED physicians with feedback on practice patterns relative to peers. Reduced resource use did not adversely affect quality of care (LOS or Return rate).

IMPACT OF PHYSICIAN FEEDBACK ON EMERGENCY DEPARTMENT RESOURCE USE, QUALITY AND EFFICIENCY

By

Shabnam Jain

MD
Lady Hardinge Medical College
1984

Thesis Committee Chair: David Howard, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Health Outcomes
2012

ACKNOWLEDGEMENTS

- Gary Frank, MD: Medical Director, Quality and Medical Management, Children's Healthcare of Atlanta
- Seton McRae: Quality Analysis Data Engineer, Children's Healthcare of Atlanta

Impact of Physician Feedback on Emergency Department Resource use, Quality and Efficiency

BACKGROUND:

Emergency Department (ED) care is considered to be expensive and carries a perception of excessive diagnostic testing and therapies especially for the non-emergent conditions that often comprise a large share of the case mix in the typical ED. Furthermore, wide variation in resource use that cannot be explained by patient severity or complaint has been demonstrated in both adults and in pediatrics, including emergency medicine.[1-10] Excessive use of resources in healthcare has not been found to improve quality or outcomes, but does impact costs.[11]

Traditional ED quality measures have included measures of timeliness such as length of stay, time to antibiotics, boarder time, safety measures such as errors, hand-washing, and measures of patient-centeredness such as left without being seen rates and patient satisfaction. Unexpected return to ED within 48 or 72-hours is a measure of effectiveness of care that is almost universally monitored in most EDs. Other measure of effectiveness of care such as adherence to evidence based guidelines are not routinely monitored. Recent studies have suggested the importance of measuring efficient use of resources for common pediatric ED conditions.[12,13] However, efforts to streamline resource use by standardizing practice using evidence based guidelines have been ongoing, but significant degrees of variation in practice remain.[5,13,14]

Audit and feedback of physician practice has been used as a tool to improve resource utilization and care efficiency.[15] The hope is that when physicians are provided with data on how their practice compares with their peer group, they will be amenable to discussing practice changes.[16]

While insurers have been profiling providers' resource use patterns, such data is often not available to individual clinicians, is usually not acuity-adjusted, and does not include a combination of metrics that balance resource use and outcome measures. The impact of combining various ED quality measures to provide comprehensive feedback to physicians on their practice patterns compared to peers is not known. In this quality improvement initiative, our objectives were: (1) to develop a tool for comprehensive feedback to ED physicians on their practice patterns relative to peers, and (2) to evaluate the impact of physician feedback on ED resource use, quality and efficiency

METHODS:

Setting and Scorecard development:

The study was conducted at a large pediatric healthcare system's two tertiary level EDs with a combined annual census of over 130,000 pediatric visits. Both sites are staffed by pediatric emergency medicine physicians and urgent care pediatricians. We used existing variation in practice to benchmark physicians against peers in order to highlight high performers; we did this by creating a scorecard showing physicians their resource use and quality metrics for four common ED conditions noted below.

Scorecard Inclusion Criteria: Four common ED conditions were included in the scorecard. To avoid bias based on final diagnosis, inclusion criteria were based on patient presenting complaints. Professional coders classified patient chief complaint into admitting diagnosis using *International Classification of Diseases, Ninth Revision* (ICD-9) codes. The four conditions and the corresponding chief complaint ICD-9 codes included: (a) Fever unspecified (780.60) (age > 2 months only as infants under 2 months receive routine screening tests and are often admitted; (b)Head Injury unspecified (959.01) (age > 3 months only as institutional guidelines apply to infants >3 months); (c) Gastroenteritis-like symptoms: Vomiting alone (787.03), Diarrhea (787.91), Dehydration (276.51); (d) Respiratory illness: Other Dyspnea and Respiratory Abnormality (786.09), Cough (786.2), Wheezing (786.07), Unspecified Asthma (493.90), with exacerbation (493.92), with status asthmaticus (493.91).

Acuity adjustment and peer comparisons: Studies on ED provider feedback have underscored the importance of adjusting for acuity, diagnosis, and patient outcomes, and recommend making comparisons of utilization rates and outcomes for individual physicians against peer-based norms.[16] In this initiative, acuity adjustment was achieved by including only patients in Emergency Severity Index (ESI) category 3 (mid-acuity).[17,18] The potential for practice variation is highest for these mid-acuity patients, and the four conditions included in the scorecard represent nearly 40% of all mid-acuity patients seen in our ED. Furthermore, institutional guidelines exist for all 4 conditions (children with fever; vomiting + diarrhea; asthma, bronchiolitis; and minor head injury). Peer group performance was shown on the scorecard as a standard box plot with 25th, 50th and 75th percentiles, and dashes for 1.5 interquartile ranges. A bold diamond indicated an individual physician's performance against the peer group's performance shown on the box plot (Figure 1).

Quality measures included in Scorecard: Three measures of ED quality of care were included in the scorecard: ED length of stay (LOS - measured as time from MD picking patient until exit from ED) was included as timeliness measure, and rate of 72-hour return for same condition was included as a balancing quality measure. Resource use, a process measure, was monitored specific to the condition as follows: *Fever:* Lab tests (complete blood count, basic or comprehensive chemistry, blood culture, C-reactive protein level counted cumulatively), Chest X-ray, Abdominal X-ray, Antibiotics: intravenous (IV) or intramuscular (IM); *Respiratory Illness*: Chest X-ray; *Head Injury*: CT scan of head without contrast; *Gastroenteritis-like* *symptoms*: Lab tests (basic or comprehensive chemistry, complete blood count), abdominal X-ray, intravenous fluids, CT scan of abdomen, and intravenous Ondansetron.

EDs account for over 50% of hospital admissions, and arguably, this is one of the most expensive resource use decisions made by an ED physician.[19] Therefore, admission rate to the hospital was measured for each of the 4 conditions. The scorecard also included the number of patients seen during the reporting period.

Exclusion Criteria: Patients who left without being seen were excluded. Performance data of physicians who saw less than 10 patients during a reporting period in any condition was not included in the peer group (box plot) calculation for that condition – this was to prevent proportions with low denominator from unduly influencing percentiles. Furthermore, physicians who saw less than 25 patients during either the pre- or post-intervention phase were excluded from analysis as these physicians were not regular ED physicians, or left the practice at some point, and therefore did not meaningfully contribute to change in practice patterns.

Physician Attribution: This was based on the attending physician assigned to the patient during the ED visit. In situations where there was transfer of care during the ED visit, resource use decisions were assigned to the first physician, while the disposition decision and 72-hour return was assigned to the second (dispositioning) physician.

Intervention:

Scorecards were distributed quarterly starting Sep 1, 2010 with July 2009-June 2010 as the first reporting period. Each subsequent reporting quarter represented a prior 12-month rolling average. The 14 month period from July 2009 - Aug 2010 (just before the first scorecard was distributed) was the pre-intervention phase (PRE) and the 16 month period from Sep 2010 – Dec 2011 was the post-intervention phase (POST)

Data Source:

Data were obtained from electronic medical records and administrative data that are stored in an institutional data warehouse. The electronic medical records have electronic signature capture that allows for accurate physician attribution; computerized physician order entry ensured that resource use was accurately assigned to the ordering physician. Data was aggregated at the level of individual providers, who were identified by blinded codes not accessible to investigators.

Statistical Analysis:

The analysis was performed as a Pre- Post intervention analysis. To evaluate overall change in resource use, Fisher exact test and Wilcoxon rank sum tests were performed using SAS software.[20]

To account for physician-specific random effects in clustering, any patient-level covariates, and time trends in resource use before and after the intervention, we compared trends in endpoints prior to and during the intervention. The effect of the intervention was measured by the change in the slope of the trend lines before and after the intervention. Results were noted in terms of marginal effects (e.g. the change in the probability that a patient receives a test pre- and post-intervention.

Statistical power was calculated based on our ability to detect differences in resource use rates between the pre- and post-intervention period. To simplify the analysis, we assumed that there are no underlying, secular trends in endpoints. Based on preliminary data, we assumed that there are 80 physicians, each of whom treats 150 patients in each period (pre- and post-intervention) for each of the four conditions. We used Monte Carlo simulation to estimate power while accounting for physician-level clustering and estimate that for most measures, we will have good power (> 80) to detect reasonable effect sizes.

The institutional review board exempted this study from review because only aggregate data were used with no identifiers and no patient-level intervention.

RESULTS:

During the study period from July 2009 to Dec 2011, there were a total of 336,294 patients seen in the EDs, of which 128,691 were in ESI 3 acuity. Of these, 48,538 met inclusion criteria (PRE – 21,612; POST – 26,926). 121 physicians saw these patients. Figure 2 shows the scorecard of one physician during the study period for all four conditions.

Table 1 shows the overall PRE POST results for the various measures reported on scorecards. Categories reported reflect only those that were included in the QI initiative and reported to providers via the scorecard. Overall, statistically significant reduction was noted in use of abdominal/pelvic CT scans, head CT scans, chest X-rays, IV antibiotics and IV Ondansetron ($p<0.001$). Hospital admission rate decreased from 7.4% to 6.7% ($p<0.001$). ED length of stay for the 4 conditions tracked in the scorecard decreased from 129 min to 125 min ($p<0.001$). The 72-hour return rate changed from 2.2% to 2.0%; while not statistically significant, the rate did *not* increase as use of tests and therapies in the first visit decreased.

Regarding trends in resource use, significant change was noted for admission to hospital ($p<0.05$), use of lab tests ($p<0.001$), and use of IV antibiotics ($p<0.05$) (Figures 3-5).

DISCUSSION:

This study shows reduction in resource use for several commonly seen conditions in the pediatric ED after physicians were provided with feedback on practice patterns, including resource use and quality metrics, relative to their peers. Reduced resource use did not adversely affect quality of care (LOS or Return rate).

Traditional ED quality measures include turnaround time and length of stay, time to antibiotics in newborns or sickle cell patients with fever, as well as patient satisfaction and left without being seen rates).[21-24] Many of these measures do not reflect clinical elements in the care provided to the patient, and are often not directly controlled by ED providers.

In this QI initiative, we took advantage of existing variation in practice to highlight outlier practice patterns in both directions of high as well as low performers. Such benchmarking of providers against peers can highlight opportunities for improvement. A unique feature of our scorecard is that it is balanced and comprehensive, providing data not only on resource use, but also providing return rates as a balancing measure as well as LOS as an efficiency measure. We also adjusted for both acuity and patient complaint, and used common ED conditions for which there were institutional guidelines available to inform practice.

We saw overall reduction in resource use in several resource categories, but trend data did not show similar effects for some of these categories. This may have been due to several factors: (1) trends pre-dating the intervention: use of abdominal/pelvic CT scans, head CT scans, chest X-rays, and IV ondansetron were significantly different in the PRE and POST periods; however, there were trends towards decreasing use noted prior to the intervention; while this decreasing trend was maintained (and in some cases showed a steeper slope), it was not statistically significant; (2) Trends reversed in the PRE and POST period e.g. for use of labs – in Figure 4, the trend line shows increasing use prior to the intervention, and decreasing use after. While the overall resource use did not change during the study, this reflects a real change in practice before and after the intervention. For use of abdominal X-rays, the reverse was noted, with trend somewhat decreasing in the PRE period, but showing increasing use in the POST period (Figure

6). This was the only resource category that showed an absolute increase in use after the intervention; we conjecture that this may reflect the use of abdominal X-rays an alternate imaging modality with decreasing use of abdominal CT scans.

Use of IV antibiotics showed significant decrease after the intervention both in overall rates as well as in significant change in trends (Table 1, Figure 5). In these common, non-severe pediatric conditions, use of antibiotics is often not indicated, and use of IV antibiotics may be a cautionary approach of some providers, and therefore amenable to change in practice. Decreased use of broad spectrum IV antibiotics would have additional downstream benefits including reduced costs, side effects, as well as reducing resistance.

Our study had some limitations. This is a single center study and may not be generalizable to other settings. However, given the increasing use of electronic medical records nationally, as well as standard systems of triage acuity assignment which can be used for severity adjustment, along with nearly universal tracking of return rates and length of stay in most EDs, we feel that scorecards similar to this can be potentially easily adapted for use in many EDs.

Of note, we did not make any other changes in the ED during the study period that would affect resource use. The EDs did get a new EMR in mid-July 2011 – the only foreseeable impact of that pertaining to this study would be an increase in LOS as providers learnt a new system. We did see an increase in overall LOS for a few months after implementation of the new EMR (data available); however, over the entire study period, the LOS specific to the patients included in the scorecards showed a decrease.

CONCLUSIONS:

This study shows reduction in resource use for several commonly seen conditions in the pediatric ED after providing ED physicians with feedback on practice patterns, including resource use and quality metrics, relative to peers. Reduced resource use did not adversely affect quality of care (LOS or Return rate).

Feedback on practice patterns relative to peers can potentially influence provider practice for patients outside the middle acuity and the four conditions we studied, thus having a broader impact. With further refinement, our efforts to develop comprehensive, severity-adjusted performance measures that encompass both quality and resource use can potentially be used to measure “value” at the individual provider level and to identify high-value providers.

REFERENCES:

1. Goldin AB, Garrison M, Christakis D. Variations between hospitals in antireflux procedures in children. Arch Pediatr Adolesc Med 2009;163:658-63.

2. Goldman RD, Scolnik D, Chauvin-Kimoff L, et al. Practice Variations in the Treatment of Febrile Infants Among Pediatric Emergency Physicians. Pediatrics 2009.

3. Hampers LC, Faries SG. Practice variation in the emergency management of croup. Pediatrics 2002;109:505-8.

4. Jain S, Elon LK, Johnson BA, Frank G, Deguzman M. Physician practice variation in the pediatric emergency department and its impact on resource use and quality of care. Pediatr Emerg Care 2010;26:902-8.

5. Landrigan CP, Conway PH, Stucky ER, Chiang VW, Ottolini MC. Variation in pediatric hospitalists' use of proven and unproven therapies. J Hosp Med 2008;3:292-8.

6. Lantos JD, Meadow W. Variation in the treatment of infants born at the borderline of viability. Pediatrics 2009;123:1588-90.

7. Powell EC, Hampers LC. Physician variation in test ordering in the management of gastroenteritis in children. Archives of Pediatrics and Adolescent Medicine 2003;157:978-83.

8. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. N Engl J Med 2010;363:45-53.

9. Tieder JS, Cowan CA, Garrison MM, Christakis DA. Variation in inpatient resource utilization and management of apparent life threatening events. J Pediatr 2008;152:629-35, 35.e1-2.

10. Wennberg JE. Variation in use of Medicare Services in Regions and Selected Academic Medical Centers: Is More Better?; 2005.

11. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Ann Intern Med 2003;138:288-98.

12. Guttmann A, Razzaq A, Lindsay P, Zagorski B, Anderson GM. Development of measures of the quality of emergency department care for. Pediatrics 2006;118:114-23.

13. Knapp JF, Simon SD, Sharma V. Quality of care for common pediatric respiratory illnesses in United. Pediatrics 2008;122:1165-70.

14. Jain S, Sullivan K. Ceftriaxone use in the emergency department: are we doing it right? Pediatr Emerg Care 2002;18:259-64.

15. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Does telling people what they have been doing change what they do? A systematic review of the effects of audit and feedback. Qual Saf Health Care 2006;15:433-6.

16. Ahwah I, Karpiel M. Using profiling for cost and quality management in the emergency department. Healthcare Financial Management 1997;51:48, 50-3.

17. Durani Y, Brecher D, Walmsley D, Attia MW, Loiselle JM. The Emergency Severity Index Version 4: reliability in pediatric patients. Pediatr Emerg Care 2009;25:751-3.

18. Gilboy N, Tanabe T, Travers D, AM R. Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4. In: Implementation Handbook 2012 Edition. Rockville, MD. : Agency for Healthcare Research and Quality. ; 2011.

19. Pitts S, Niska R, Xu J, Burt C. National Hospital Ambulatory Medical Care Survey: 2006 emergency department summary. National health statistics reports 2008;No. 7.

20. SAS 9.1.3. Cary NSII-Ha, Documentation.

21. Hung GR, Chalut D. A consensus-established set of important indicators of pediatric emergency department performance. Pediatr Emerg Care 2008;24:9-15.

22. Welch SJ, Asplin BR, Stone-Griffith S, Davidson SJ, Augustine J, Schuur J. Emergency department operational metrics, measures and definitions: results of the Second Performance Measures and Benchmarking Summit. Ann Emerg Med 2011;58:33-40.

23. Welch SJ, Stone-Griffith S, Asplin B, Davidson SJ, Augustine J, Schuur JD. Emergency department operations dictionary: results of the second performance measures and benchmarking summit. Acad Emerg Med 2011;18:539-44.

24. Pham JC, Kirsch TD, Hill PM, DeRuggerio K, Hoffmann B. Seventy-two-hour returns may not be a good indicator of safety in the emergency department: a national study. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2011;18:390-7.

Figure 1: A box plot showing an individual physician's performance (shown as a blue diamond) against the peer group's performance

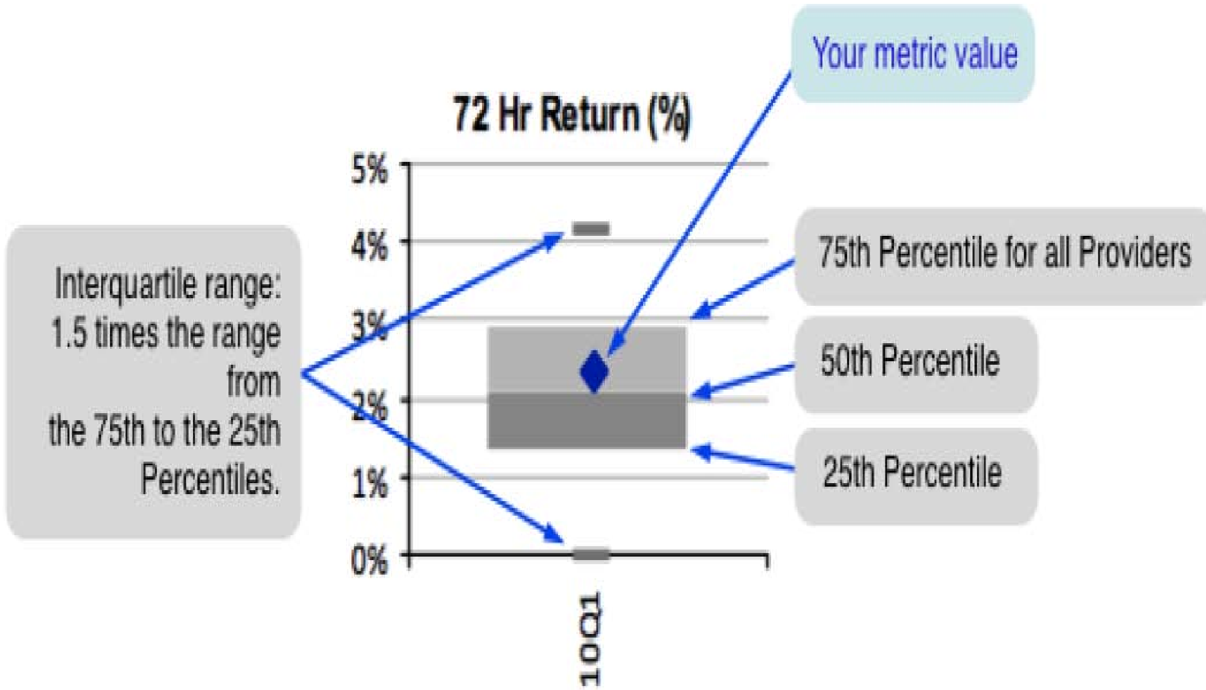

Figure 2: Scorecard for one physician for all four conditions (fever, gastroenteritis-type illness, respiratory illness, and head injury) during the entire study period.

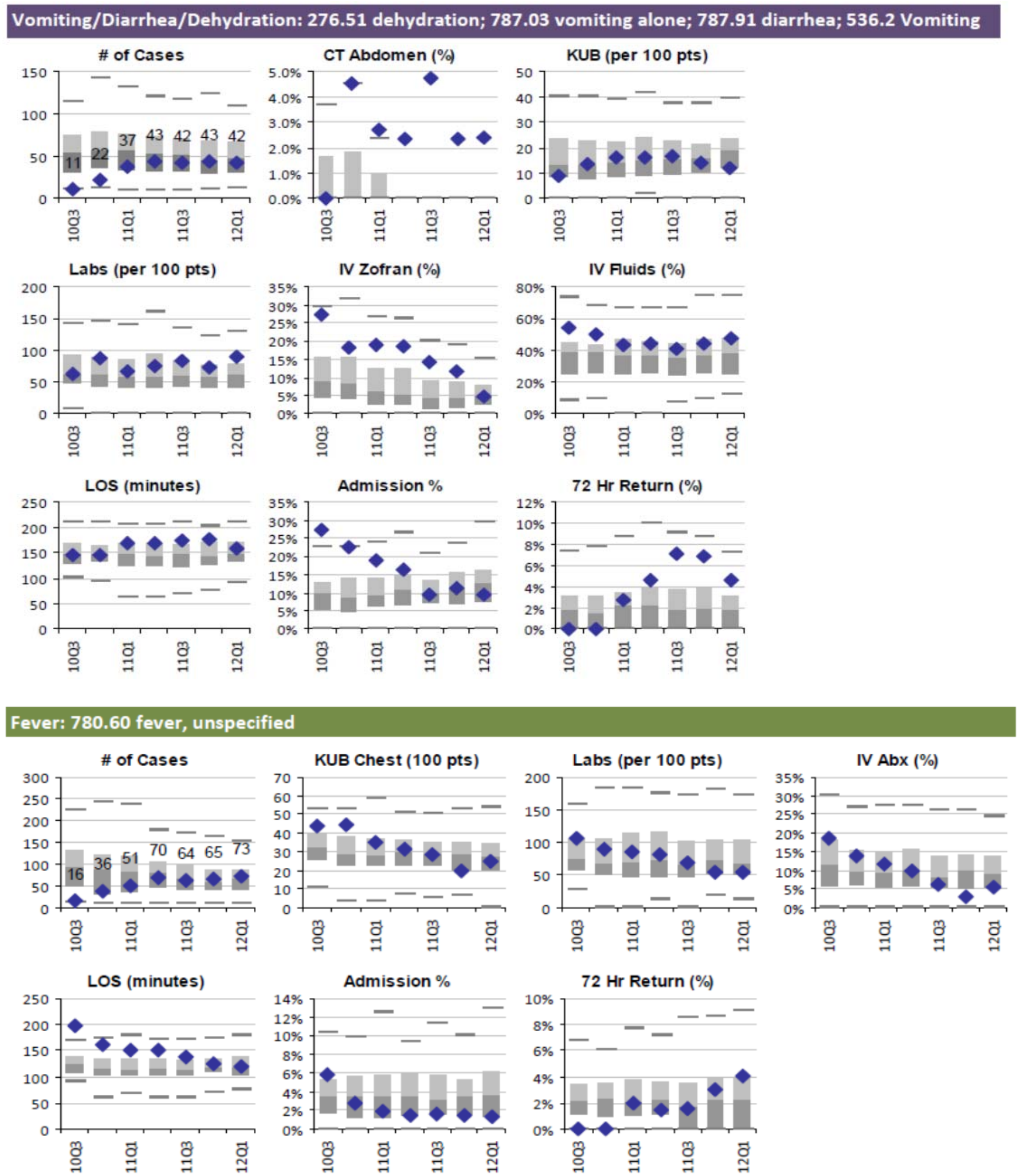

Figure 2 (continued): Scorecard for one physician for all four conditions (fever, gastroenteritis-type illness, respiratory illness, and head injury) during the entire study period.

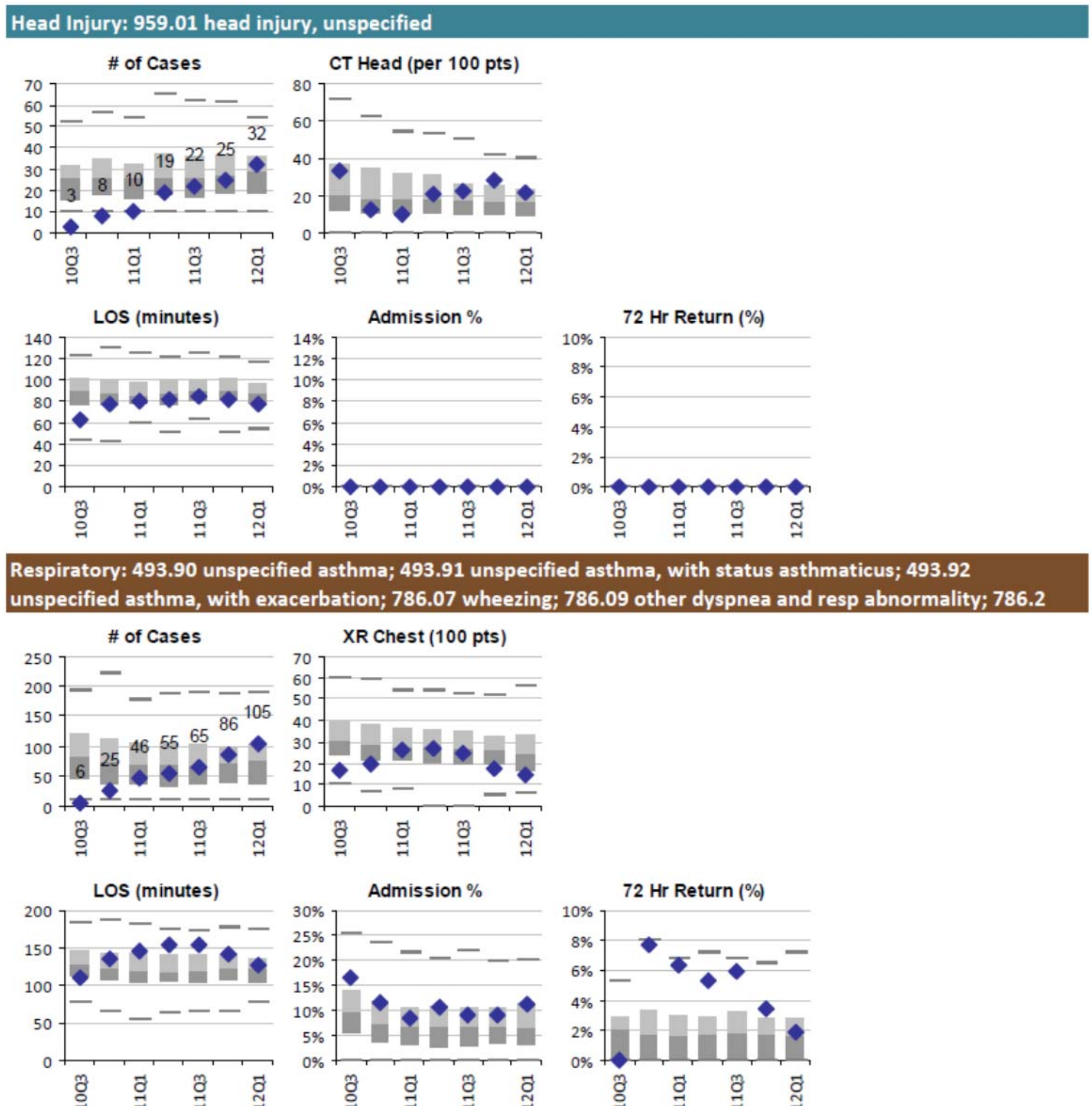

Table 1: Overall Pre- and Post-Intervention results for the quality, efficiency and resource use measures reported on scorecards.

Resource/Outcome	PRE	POST	P-Value
Abdomen/Pelvis CT Scan (%)	1.2	0.6	<0.0001
Head CT Scan (%)	26.0	19.1	<0.0001
Chest X-ray (per patient)	31.7	28.1	<0.0001
IV Antibiotics (%)	12.0	10.8	<0.0001
IV Ondansetron (%)	11.6	8.1	<0.0001
Abdominal X-ray (per patient)	15.7	16.6	ns
Lab Tests (per patient)	71.1	68.7	ns
IV Fluids (%)	37.8	38.6	ns
Hospital Admission (%)	7.4	6.7	<0.0001
Length of Stay (min)	129	125	<0.0001
72-hr Return Rate (%)	2.2	2.0	ns

Figure 3: Trends in rates of hospital admission for the four study conditions showing significant change (p=0.02) after intervention.

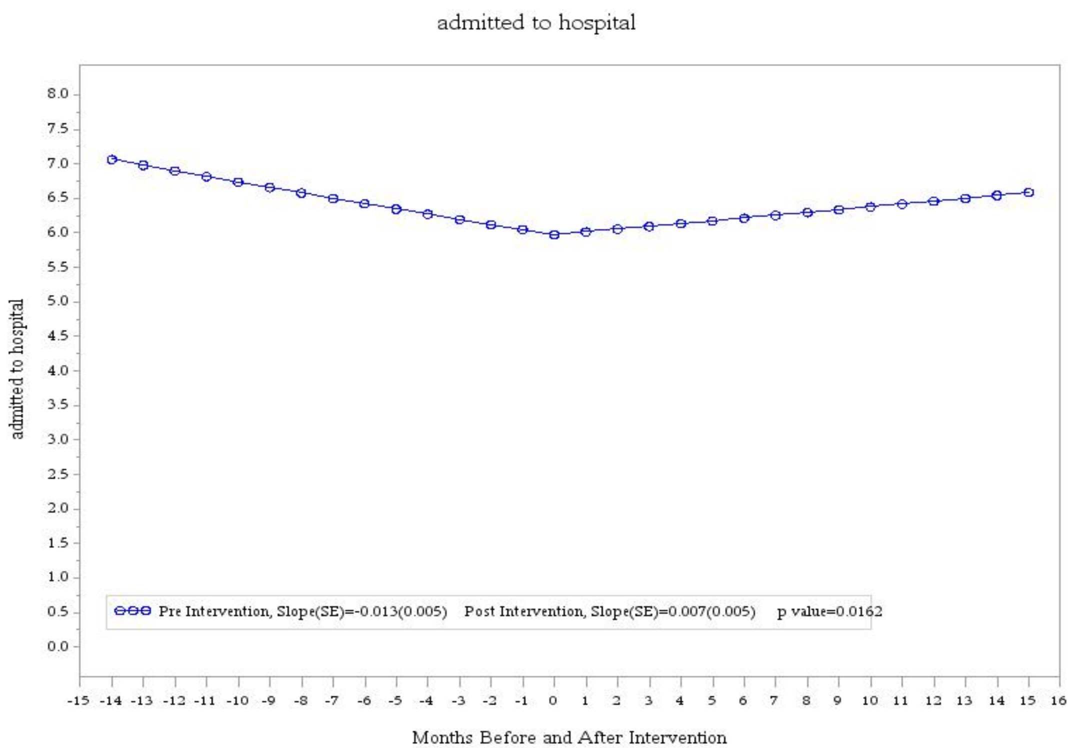

Figure 4: Trends in use of lab tests for the four study conditions showing significant reduction (p<0.001) before and after intervention.

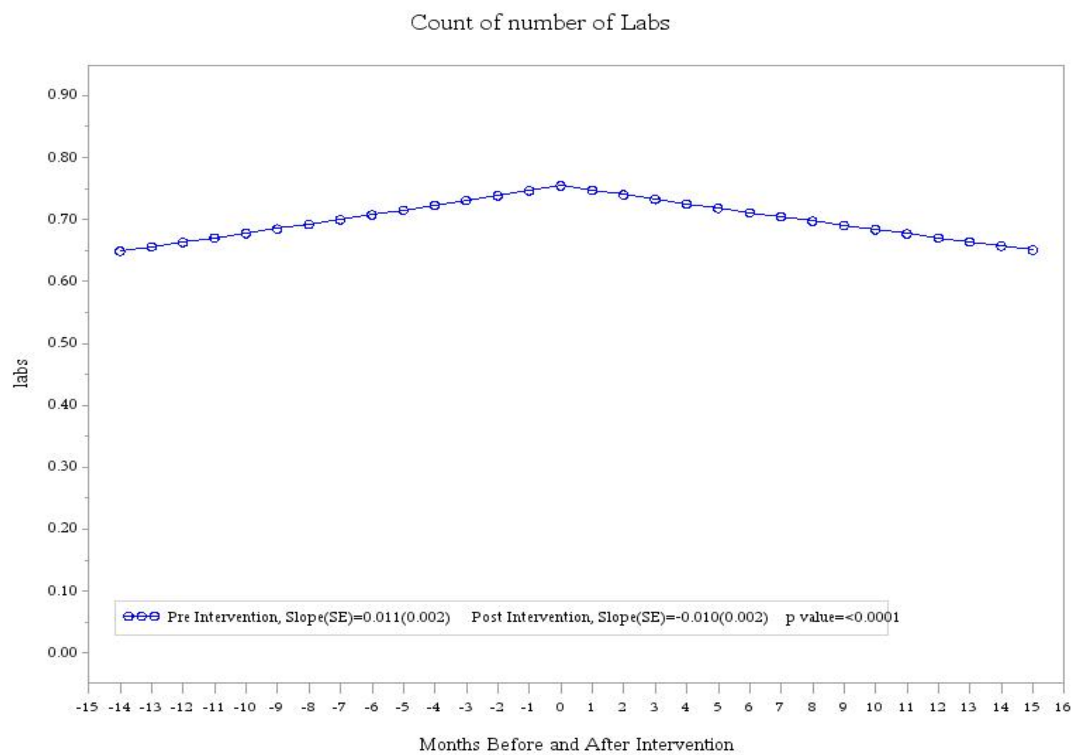

Figure 5: Trends in use of intravenous antibiotics for the four study conditions showing significant reduction (p=0.02) before and after intervention.

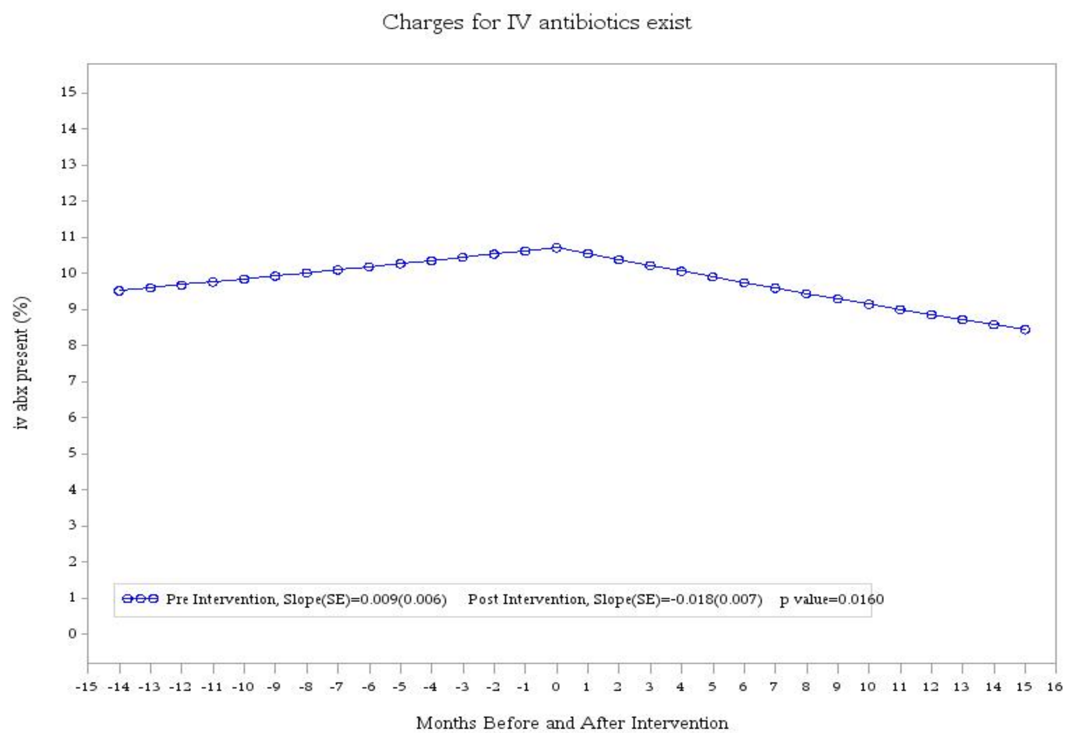

Figure 6: Trends in rates of admission to hospital for the four study conditions showing significant reduction (p=0.02) before and after intervention.