Journal of Pediatric Surgery 50 (2015) 144–148

Contents lists available at ScienceDirect

Journal of Pediatric Surgery

journal homepage: www.elsevier.com/locate/jpedsurg

Reducing computed tomography scans for appendicitis by introduction of a standardized and validated ultrasonography report template

Jason W. Nielsen [a], Laura Boomer [a], Kelli Kurtovic [a], Eric Lee [b], Kevin Kupzyk [d], Ryan Mallory [c], Brent Adler [b], D. Gregory Bates [b], Brian Kenney [a,*]

[a] Department of Surgery, Nationwide Children's Hospital and The Ohio State University College of Medicine, Columbus, OH, USA
[b] Department of Radiology, Nationwide Children's Hospital and The Ohio State University College of Medicine, Columbus, OH, USA
[c] The Ohio State University College of Medicine, Columbus, OH, USA
[d] University of Nebraska Medical Center, Omaha, NE, USA

ARTICLE INFO

Article history:
Received 4 October 2014
Accepted 6 October 2014

Key words:
Acute appendicitis
Computed tomography
Ultrasound
Secondary signs

ABSTRACT

Purpose: Computed tomography (CT) for the diagnosis of appendicitis is associated with radiation exposure and increased cost. In an effort to reduce the diagnostic use of CT scans, we implemented a standardized ultrasound report template based on validated secondary signs of appendicitis.
Methods: In September 2012, as part of a quality improvement project, we developed and introduced a four category standardized ultrasound report template for limited right lower quadrant abdominal ultrasounds. Outcomes for patients undergoing ultrasound or CT scan for appendicitis between 9/10/2012 and 12/31/2013 (Period 2, n = 2033) were compared to the three months prior to implementation (Period 1, n = 304).
Results: In Period 1, 78 of 304 (25.7%) patients had appendicitis versus 385 of 2033 (18.9%) in Period 2 ($p = 0.006$). Non-diagnostic exams decreased from 48% to 0.1% ($p < 0.001$). Ultrasound sensitivity improved from 66.67% to 92.2% ($p < 0.001$). Specificity did not significantly change (96.9% to 97.69%, $p = 0.46$). CT utilization for appendicitis decreased from 44.3% in Period 1 to 14.5% at the end of Period 2 ($p < 0.001$).
Conclusions: Implementation of a standardized ultrasound report template based on validated secondary signs of appendicitis nearly eliminated non-diagnostic exams, improved diagnostic accuracy, and resulted in a striking decrease in CT utilization.

Acute appendicitis is the most frequent cause of acute abdominal surgery in children [1]. The total lifetime cumulative incidence rate of appendicitis is 9% and has been increasing annually [2]. The diagnosis of appendicitis is most prevalent during the second decade of life, specifically between the ages of 10 and 14 [2]. Despite the frequency of appendicitis, diagnosis can be challenging [3,4].

Both ultrasound and CT have been reported to improve diagnostic accuracy in appendicitis [5]. CT scan rates for the diagnosis of appendicitis have been increasing nationally [6,7]. Although CT is reported to have a higher sensitivity than ultrasound, ongoing concerns have been raised about the radiation exposure and increased costs associated with CT [8,9]. Projections estimate that a solid cancer will result at a rate of 25.8 to 33.9 cases per 10,000 abdominal CT scans for girls and 13.1 to 14.8 cases per 10,000 abdominal CT scans for boys [10]. Strategies to increase the utility of ultrasound as a diagnostic tool for appendicitis are desirable to reduce radiation exposure and decrease costs, but ultrasound has challenges as well. Appendix visualization rates vary and ultrasound exhibits significant user dependency [11–13].

Others have sought to decrease CT rates by establishing diagnostic protocols using pediatric appendicitis scores and surgeon assessment [3,14]. Increased use of magnetic resonance imaging (MRI) has been proposed [15], but MRI is also associated with challenges including cost, time and potential need for sedation to obtain an accurate study.

We designed and implemented a standardized ultrasound reporting template based on validated secondary signs of appendicitis in order to increase the diagnostic accuracy of the ultrasound exam and to simultaneously decrease CT utilization.

1. Materials and methods

1.1. Template design

A collaborative group from the quality improvement, pediatric surgery, and pediatric radiology departments met in 2012 to discuss a standardized ultrasound reporting template for appendicitis. The current literature was reviewed to design a template with maximal sensitivity and specificity. A maximal outer diameter of <7 mm and a maximal appendiceal wall thickness of <1.7 mm with graded compression were considered normal [16]. Secondary signs were defined as hyperechogenic periappendiceal fat, fluid collection consistent with an appendicular abscess, and local dilation and hypoperistalsis of the

* Corresponding author at: Nationwide Children's Hospital, 700 Children's Drive, Suite ED324, Columbus, Ohio 43205-2696. Tel.: +1 614 722 3928; fax: +1 614 722 3903.
E-mail address: brian.kenney@nationwidechildrens.org (B. Kenney).

http://dx.doi.org/10.1016/j.jpedsurg.2014.10.033
0022-3468/

bowel consistent with focal peritonitis [17]. Other sonographic findings including hyperemia of the appendix, free fluid, lymphadenopathy, and appendicoliths were included as part of the template for assessment but were not considered secondary signs in the analysis. Radiologists were asked to classify patients into four categories: 1. Normal appendix; 2. Appendix not visualized or partially visualized without secondary signs of appendicitis; 3. Appendix not visualized or partially visualized with secondary signs of appendicitis; 4. Acute appendicitis [17]. Category 1 and 2 reports were considered negative for appendicitis while Category 3 and 4 reports were considered positive.

Three criteria were required for a compliant ultrasound report: the template had to be used by the radiologist, one of the four categories had to be selected, and the category selected had to match the information in the report. A sample ultrasound template consistent with Category 1 is shown in Fig. 1. Non-diagnostic exams were defined as ultrasound reports where the description was insufficient to make or exclude the diagnosis of appendicitis. An IRB exemption was granted for this quality improvement project (IRB# 13–00734).

1.2. Patients

From 9/10/2012 to 12/31/2013 (Period 2, n = 2033) records from all patients undergoing abdominal ultrasound evaluation in our emergency department (ED) were prospectively reviewed. Only ultrasounds performed for suspected appendicitis were included in the analysis. Patients enrolled in a concurrently running study on the non-operative management of appendicitis were excluded. Demographic data, ultrasound reports, and diagnostic accuracy were compared to the period prior to implementation of the template from 6/1/2012 to 9/9/2012 (Period 1, n = 304). All CT scans obtained in our ED for evaluation of appendicitis during Periods 1 and 2 were also reviewed. All patients undergoing imaging for suspicion of appendicitis were used to calculate the CT utilization rate. Final diagnosis was determined by histopathology and defined as transmural inflammation of the appendix [18,19].

1.3. Statistical analysis

Descriptive statistics were calculated, including frequencies, percentages, means, and standard deviations. To statistically compare Periods 1 and 2, chi-square tests were performed for categorical variables. Fisher's Exact Test was used when the variable was dichotomous and the number of data points was below 5. T-tests were used to make group comparisons for continuous variables. Measures of test accuracy including sensitivity, specificity, and predictive values were determined by standard methods. Results were tracked using statistical process control (SPC) methodology with control charts (p-charts) per established quality improvement practices. Compliance was defined as ultrasound reports that used the template and selected a category per established guidelines.

1.4. Normal work flow

Initial clinical evaluation of all patients with abdominal pain is performed by our ED physicians. ED physicians make the determination to obtain imaging. Ultrasound is our primary diagnostic imaging modality and is preferentially ordered but is not mandated. The ED physicians had full license to order both initial and secondary imaging for appendicitis prior to surgical consultation. ED Physicians could involve surgery in the decision if they wished but this was not mandated during the time of the study and surgery involvement prior to ordering imaging was inconsistent.

2. Results

2.1. Patient demographics/epidemiology

Demographic data between Periods 1 and 2 were similar. The rate of appendicitis among patients undergoing ultrasound decreased from 25.7% (78/304) in Period 1 to 18.9% (385/2033, p = 0.006) in Period 2. At the same time, the average number of ultrasound exams per month increased from 92.3 in Period 1 to 129.4 in Period 2. Despite differences in appendicitis rates the negative appendectomy rate (NAR) was unchanged (Period 1 = 8.23% vs. Period 2 = 8.76%, p = 0.8).

2.2. Diagnostic accuracy

The sensitivity and negative predictive value (NPV) of our ultrasound exams improved to greater than 90% after the template, and non-diagnostic exams were nearly eliminated (Table 1). The specificity and positive predictive value (PPV) of ultrasound also improved but not significantly (Table 1). Compliance to the template led to a lower NAR, lower CT utilization rate, and higher specificity (Table 2). The rate of non/partial visualization was unchanged between Periods (55.9% vs. 54.1%, p = 0.55).

The predictive values of ultrasound varied for each template category. The NPV of Categories 1 & 2 and the PPV of Category 4 were high. Category 3 was our poorest performing category with a PPV of 76% (Table 3). Falsely positive Category 3 patients had a female predominance (12/17 or 70.6%, p = 0.137).

Figure 1

Appendix: The appendix is identified in the right lower quadrant.
Appendix size: The appendix is less than 7 mm in outer diameter measuring 4.2 mm.
Wall thickness: Normal, less than 1.7 mm in thickness measuring 0.8 mm
****Appendicolith:** No appendicolith was identified.
Perforation: None
***Abscess:** None
****Fluid:** There is no free fluid.
***Periappendiceal fat:** The periappendiceal fat is normal.
****Vascularity:** Normal vascularity was observed without hyperemia.
****Mesenteric lymph nodes:** No pathologically large lymph nodes were observed.
***Adjacent bowel loops:** Peristalsing normal appearing bowel loops were observed.
Additional abnormalities: None

Impression: Normal appendix.

Fig. 1. Sample Category 1 standardized right lower quadrant limited ultrasound report template. Primary criteria include appendiceal size and wall thickness. * Approved secondary signs. ** Additional signs not considered approved secondary signs.

Table 1
Diagnostic utility of ultrasound pre/post template.

	Pre-Template n = 304	Post-Template n = 2033	P Value
Sensitivity	66.7%	92.2%	<0.001
Specificity	96.9%	97.7%	0.466
Positive Predictive Value (PPV)	88.1%	90.3%	0.599
Negative Predictive Value (NPV)	89.4%	98.2%	<0.001
Non-diagnostic exams	48.0%	0.1%	<0.001

Table 3
CT rates and outcomes by category.

	Total	Appendicitis	Non-appendicitis	% Correct	CT #	CT Rate
Category 1	767	4	763	99.5% (NPV)	32	4.2%
Category 2	871	25	846	97.1% (NPV)	301	34.6%
Category 3	71	54	17	76.1% (PPV)	37	52.1%
Category 4	322	301	21	93.5% (PPV)	14	4.4%

Depiction of the total number of patients, positive predictive values (PPV) and negative predictive values (NPV), and CT scan rate by category in Period 2. Category 1 and 2 were considered negative for appendicitis, Category 3 and 4 were considered positive.

2.3. CT utilization and template compliance

After initiation of the template, the abdominal CT utilization rate dropped in stepwise fashion by 67.3% (baseline 44.3% to 14.5%, $p < 0.05$, Fig. 2). CT scan rates were high in Category 2 and 3 when compared to Categories 1 and 4 (Table 3). The average CT scan rate for Category 2 ultrasounds in the first three months after the template was 41%, and this decreased to 26% in the last three months of 2013. Ultrasound was the initial imaging study in 90.5% of patients in Period 1 vs. 97.3% in Period 2 ($p < 0.001$).

Compliance to the template increased over time in stepwise fashion to a high of 95.4%. Each increase in template compliance was statistically significant and directly correlated with a statistically significant decrease in our CT utilization rate (Fig. 2).

3. Discussion

Development of a structured ultrasound report template based on validated size criteria and secondary signs of appendicitis nearly eliminated non-diagnostic reports, improved diagnostic accuracy, and led to a 67.3% decrease in CT rate. The sensitivity, specificity, and positive and negative predictive values reported here are consistent with the reported values in the literature for the secondary signs we selected [17].

In attempting to decrease our CT utilization rate for appendicitis we faced many challenges. First, the appendix is not fully visualized in about half of our patients undergoing ultrasound examination. Reported rates of visualization in the literature range from 24.4% to 69.3% [11]. Lack of visualization prior to the template was generally considered non-diagnostic and led to high CT utilization. The template was not created to increase our rate of visualization, and this remained the same between the two periods (55.9% vs. 54.1%, $p = 0.55$). Instead, the template was designed to increase our ability to diagnose appendicitis even in the absence of visualization by utilizing secondary signs.

Next, before we instituted the new report template, secondary signs of appendicitis that might have aided in diagnosis were inconsistently interpreted. Many signs that were not specific such as lymphadenopathy and free fluid were emphasized in the report, leading to confusion and more CT scans.

Finally, some clinicians expressed a distrust of ultrasound as a definitive diagnostic test for appendicitis. The understandable fear of a missed diagnosis often led, and continues to lead, to CT utilization even when the suspicion for appendicitis is low. For example, Category 2 ultrasounds have an NPV of 97%, but CT scans were still ordered in 34.6% of these patients. Although this number has decreased from an initial average of 41% down to 26%, there continues to be room for improvement.

Category 3 ultrasounds represent both our smallest category (3.49%) and our most difficult category to diagnose. Category 3 ultrasounds are considered positive for appendicitis but only 54 of 71 (76.1%) patients in this category had appendicitis. Several pathologies can cause inflammation in the right lower quadrant especially among females who predominate in this category. When the appendix is not visualized distinguishing between these causes is difficult and these patients will likely receive the most diagnostic benefit from CT scanning. Our observation of higher CT rates (52.1%) in Category 3 is consistent with this diagnostic reality.

Our NAR of 8.76% appears high likely due to our strict definition of transmural inflammation for appendicitis, a definition which is supported by the literature [18,19]. This NAR is consistent with reported ranges in the literature [5,18,19], though the definition of negative appendectomy varies widely. Some studies define negative appendectomies as the absence of inflammation [18]. This likely underestimates the true NAR. Other NARs are defined entirely by billing codes without pathology [6]. Some studies use more than one definition of NAR and report multiple values [18]. Given this variance in definition it is difficult to directly compare NARs between institutions. Template compliance resulted in a significantly lower NAR (6.3%). Also, a large portion of our negative appendectomies (17/37) were classified as normal by ultrasound (e.g. the patients were taken to the operating room based on clinical signs of appendicitis, despite the ultrasound report). With improved understanding of the value and specificity of the template, the NAR should improve.

The decrease in CT scan utilization after the template was largely related to a decrease in CT scans ordered after non-diagnostic ultrasounds. The change is also reflective of an increased voluntary usage of sonogram over CT. This increase in voluntary usage and confidence in ultrasound is demonstrated by the increased rate of ultrasound utilization from Period 1 to Period 2 (92.3 per month to 129.4 per month). The increased use of ultrasound is a good alternative to high CT usage given its lower cost and lack of radiation, but even ultrasound may be over used as many ultrasounds continue to be ordered even when the suspicion for appendicitis is low. Unfortunately, no perfect test for appendicitis exists and even with CT incorrect diagnoses and negative appendectomies occur [5,6].

The most important contributions of the template are likely related to the improved communication between services and the standardization based on evidence-based principles. The template was specifically designed to eliminate subjective observation that might lead to confusion. It creates an actionable interpretation rather than a simple narrative impression. Improved communication through handoffs and patient checklists is currently a popular topic and leads to better patient outcomes [20,21]. Also, recent standardization with the implementation of appendicitis clinical pathways has been shown to improve patient care [22].

The implementation of a standardized ultrasound template based on established secondary signs of appendicitis led to a significant decrease

Table 2
Compliance to the template improves outcomes.

	Compliant	Non-Compliant	P Value
Total Reports	1741	292	
Sensitivity	92.7%	90.7%	0.641
Specificity	97.9%	90.4%	<0.001
Negative Appendectomy Rate (NAR)	6.3%	24.6%	<0.001
CT Rate	15.8%	37.3%	<0.001

The effect of the template is further demonstrated when non-compliant exams are separated from compliant exams. Ultrasound reports had to meet three criteria to be compliant. 1. The template had to be used, 2. One of the four categories had to be selected, 3. The category selected had to match the approved diagnostic criteria.

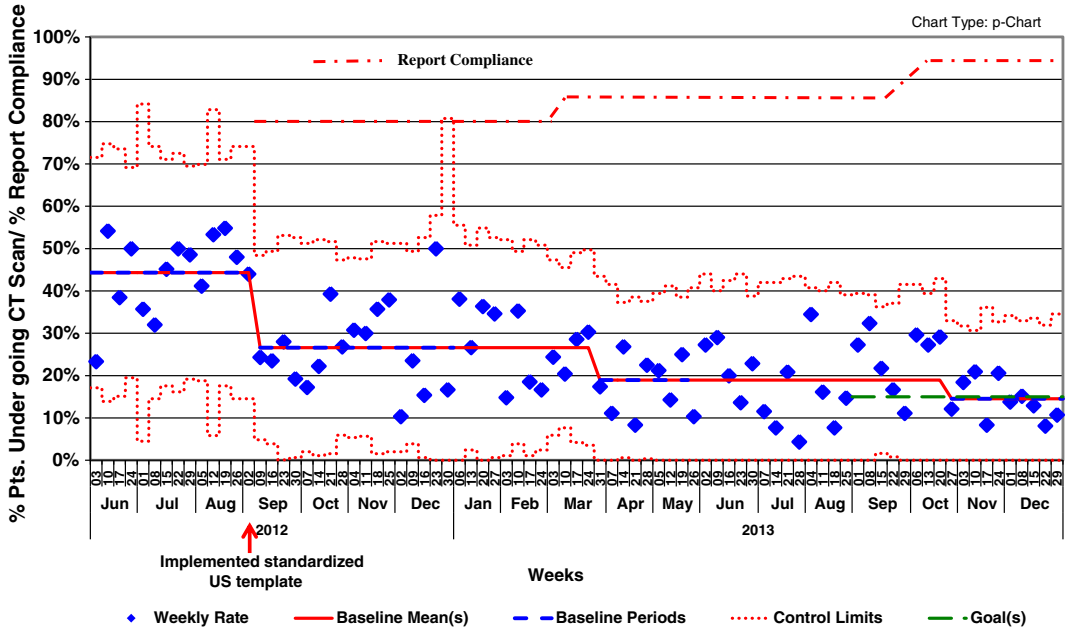

Fig. 2. Control Chart (p-chart) demonstrating weekly CT scan utilization rate with superimposed ultrasound report compliance rate. **Blue Diamonds**: Weekly CT scan rate (number of patients who got a CT/number of patients assessed for appendicitis). **Baseline Means**: Calculated mean CT scan rates. **Baseline Periods**: Time periods used to calculate a new mean CT scan rate after a shift in CT scan rate was detected. **Control Limits**: Control limits are calculated as $\overline{p} \pm 3 * \sqrt{\frac{\overline{p}*(1-\overline{p})}{n}}$ where $\overline{p}$ = the baseline mean and n = the number of samples for the specific time period. Points lying outside of the control limits or patterns within the control limits suggest a new source of variation or a change in an established process. **Report Compliance**: The superimposed baseline mean from our control chart (p-chart) tracking compliance to the ultrasound template. Note how each successive increase in compliance correlates with a decrease in CT scan utilization.

in CT utilization as well as improved accuracy. This change has been sustained for more than a year and continues to improve with improved interdisciplinary compliance. A simple strategy based on established evidence can significantly impact patient care and outcomes.

Appendix A. Discussions

Presented by Dr. Jason Nielsen, Columbus, OH.

Discussant: Dr. Charles Stolar (Santa Barbara, CA) How did you get the radiologists and their ultrasound technicians to buy in to this?

Response: Dr. Jason Nielsen It's been difficult. As you can see from our compliance curve, it took numerous meetings to get to even the 96% compliance. Essentially we kind of bound the radiologists' hands. They had to choose one of our four categories. They were no longer able to have any hedging which I think is what led to our success. We had a couple of great staff members included on the study here that were really paramount in helping us to accomplish that.

Discussant: Dr. David Skarda (Salt Lake City, UT) Congratulations on an excellent process improvement project. This is outstanding work. One question I have for you is in those patients that have nondiagnostic studies or your category 3 variety did you give any consideration to a period of observation rather than proceeding to CT scan? I think there is a potential for allowing time to decide rather than radiation.

Response: Dr. Jason Nielsen I agree completely. We are currently working to develop a protocol for these patients. There is a lot of literature to suggest that the sensitivity of ultrasound increases with multiple exams so we're in a period where we are trying to determine exactly what to do when we have an ultrasound that we are concerned about, whether we should observe them and potentially repeat it, or just observe and do serial exams. We haven't determined that as yet but I think that is the future direction of this project.

Discussant: Dr. David Skarda In Salt Lake City we have moved towards a period of observation rather than proceeding towards CT scan.

References

[1] Scholer SJ, Pituch K, Orr DP, et al. Clinical outcomes of children with acute abdominal pain. Pediatrics 1996;98(4):680–5.
[2] Anderson JE, Bickler SW, Chang DC, et al. Examining a common disease with unknown etiology: trends in epidemiology and surgical management of appendicitis in California, 1995–2009. World J Surg 2012;36(12):2787–94.
[3] Kosloske AM, Love CL, Rohrer JE, et al. The diagnosis of appendicitis in children: outcomes of a strategy based on pediatric surgical evaluation. Pediatrics 2004;113(1):29–34.
[4] Sivit CJ. Imaging the child with right lower quadrant pain and suspected appendicitis: current concepts. Pediatr Radiol 2004;36(6):447–53.
[5] Bachur RG, Hennelly K, Callahan MJ, et al. Diagnostic imaging and negative appendectomy rates in children: effects of age and gender. Pediatrics 2012;129(5):877–84.
[6] Martin AE, Vollman D, Adler B, et al. CT scans may not reduce the negative appendectomy rate in children. J Pediatr Surg 2004;39(6):886–90.
[7] Tsze DS, Asnis LM, Merchant RC, et al. Increasing computed tomography use for patients with appendicitis and discrepancies in pain management between adults and children: an analysis of the NHAMCS. Ann Emerg Med 2012;59(5):395–403.
[8] Doria AS, Moineddin R, Kellenberger CJ, et al. US or CT for diagnosis of appendicitis in children and adults? A meta-analysis. Radiology 2006;241(1):83–94.
[9] Parker L, Nazarian LN, Gingold EL, et al. Cost and radiation savings of partial substitution of ultrasound for CT in appendicitis evaluation: a national projection. Am J Roentgenol 2014;202(1):124–35.
[10] Miglioretti DL, Johnson E, Williams A, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. JAMA Pediatr 2013;167(8):700–7.
[11] Quigley AJ, Stafrace S. Ultrasound assessment of acute appendicitis in paediatric patients: methodology and pictorial overview of findings seen. Insights Imaging 2013;4(6):741–51.
[12] Pohl D, Golub R, Schwartz GE, et al. Appendiceal ultrasonography performed by nonradiologist: does it help in the diagnostic process? J Ultrasound Med 1998;17(4):217–21.
[13] Lee JH, Jeong YK, Park KB, et al. Operator-dependent techniques for graded compression sonography to detect the appendix and diagnose acute appendicitis. Am J Roentgenol 2005;184(1):91–7.
[14] Adibe OO, Amin SR, Hansen EN, et al. An evidence-based clinical protocol for diagnosis of acute appendicitis decreased the use of computed tomography in children. J Pediatr Surg 2011;46(1):192–6.

[15] Aspelund G, Fingeret A, Gross E, et al. Ultrasonography/MRI versus CT for diagnosing appendicitis. Pediatrics 2014;133(4):586–93.
[16] Goldin A, Khanna P, Thapa M, et al. Revised ultrasound criteria for appendicitis in children improve diagnostic accuracy. Pediatr Radiol 2011;41(8):993–9.
[17] Wiersma F, Toorenvliet B, Bloem J, et al. US examination of the appendix in children with suspected appendicitis: the additional value of secondary signs. Eur Radiol 2009;19(2):455–61.
[18] Mariadason JG, Wang WN, Wallack MK, et al. Negative appendicectomy rate as a quality metric in the management of appendicitis: impact of computed tomography, Alvarado score and the definition of negative appendicectomy. Ann R Coll Surg Engl 2012;94(6):395–401.
[19] Marudanayagam R, Williams GT, Rees BI. Review of the pathologic results of 2660 appendectomy specimens. J Gastroenterol 2006;41(8):745–9.
[20] Starmer AJ, Sectich TC, Simon DW, et al. Rates of medical errors and preventable adverse events among hospitalized children following implementation of a resident handoff bundle. JAMA 2013;310(21):2262–70.
[21] Bergs J, Hellings J, Cleemput I, et al. Systematic review and meta-analysis of the effect of the World Health Organization surgical safety checklist on postoperative complications. Br J Surg 2014;101(3):150–8.
[22] Knott EM, Gasior AC, Ostlie DJ, et al. Decreased resource utilization since initiation of institutional clinical pathway for care of children with perforated appendicitis. J Pediatr Surg 2013;48(6):1395–8.