



Published in final edited form as:

Ann Surg. 2012 October ; 256(4): 586–594. doi:10.1097/SLA.0b013e31826a9602.

Progress in the Diagnosis of Appendicitis: A Report from Washington State's Surgical Care and Outcomes Assessment Program (SCOAP)

The SCOAP Collaborative, Frederick Thurston Drake¹, Michael G. Florence², Morris G. Johnson³, Gregory J. Jurkovich⁴, Steve Kwon¹, Zeila Schmidt⁵, Richard C. Thirlby⁶, and David R. Flum^{1,2}

¹University of Washington Medical Center, Seattle, Washington

²Swedish Medical Center, Seattle, Washington

³Skagit Valley Hospital, Mt. Vernon, Washington

⁴Harborview Medical Center, Seattle, Washington

⁵University of Washington Surgical Outcomes Research Center, Seattle, Washington

⁶Virginia Mason Medical Center, Seattle, Washington

Abstract

BACKGROUND and OBJECTIVES—Studies suggest that CT and US can effectively diagnose and rule-out appendicitis, safely reducing negative appendectomies (NA); however, some within the surgical community remain reluctant to add imaging to clinical evaluation of patients with suspected appendicitis. The Surgical Care and Outcomes Assessment Program (SCOAP) is a physician-led quality initiative that monitors performance by benchmarking processes of care and outcomes. Since 2006, accurate diagnosis of appendicitis has been a priority for SCOAP. The objective of this study was to evaluate the association between imaging and NA in the general community.

METHODS—Data were collected prospectively for consecutive appendectomy patients (age > 15) at nearly 60 hospitals. SCOAP data are obtained directly from clinical records, including radiology, operative, and pathology reports. Multivariate logistic regression models were used to examine the association between imaging and NA. Tests for trends over time were also conducted.

RESULTS—Among 19,327 patients (47.9% female) who underwent appendectomy, 5.4% had NA. Among patients who were imaged, frequency of NA was 4.5%, whereas among those who were not imaged, NA was 15.4% ($p < 0.001$). This association was consistent for males (3% vs. 10%, $p < 0.001$) and for reproductive-age females (6.9% vs. 24.7%, $p < 0.001$). In a multivariate model adjusted for age, sex, and WBC, odds of NA for patients not imaged were 3.7 times the odds for those who received imaging (95% CI 3.0 – 4.4). Among SCOAP hospitals, use of imaging increased and NA decreased significantly over time; frequency of perforation was unchanged.

CONCLUSIONS—Patients who were not imaged during work-up for suspected appendicitis had over three times the odds of NA as those who were imaged. Routine imaging in the evaluation of patients suspected to have appendicitis can safely reduce unnecessary operations. Programs such as SCOAP improve care through peer-led, benchmarked practice change.

Corresponding Author: Frederick Thurston Drake, MD, Department of Surgery, University of Washington, 1859 NE Pacific Street, Box 356410, Seattle, WA 98195, ftdrake@u.washington.edu, mobile: 801.647.4754.

Presented at the 132nd Annual Meeting at the American Surgical Association April 27, 2012, San Francisco, California

INTRODUCTION

Surgical convention suggests that clinical assessment is usually sufficient to make the diagnosis of acute appendicitis. Under this view, a certain frequency of so-called negative appendectomies (NA) – in which a non-inflamed appendix is removed from patients mistakenly suspected to have appendicitis – is acceptable in order to prevent under-diagnosis, delay in definitive therapy, and an attendant increase in the risk of appendiceal perforation. However, recent studies have found that the addition of advanced diagnostic imaging to the clinical evaluation of suspected appendicitis is associated with a reduction in the frequency of NA without an associated increase in the frequency of perforation.^{1–16} Surgeons and Emergency Medicine physicians now commonly employ imaging in the work-up of appendicitis, and many of the most recent studies are devoted to evaluating diagnostic protocols, such as sequenced ultrasound (US) and computed tomography (CT) pathways designed to limit exposure to ionizing radiation.^{17–23} Despite growing acceptance of imaging and widely replicated performance results in tertiary centers, the accuracy of diagnostic imaging in some community settings has not achieved that reported by clinical studies, the utility of imaging across diverse community settings has not been established as a safe means of reducing unnecessary operations, and many surgeons feel that CT is not necessary and overused.^{24–27}

The Surgical Care Outcomes and Assessment Program (SCOAP) is a physician-led quality surveillance program that began in 2006 and has subsequently enrolled essentially all hospitals in Washington State. Data are collected prospectively by trained abstractors and statewide reports are issued (individual institutions are de-identified). Many aspects of surgical care are reported including specific processes of care and clinical outcomes. Performance benchmarks are established by high-achieving hospitals for both processes and outcomes. Although SCOAP data is collected primarily as a quality improvement endeavor, it is a source of data for observational research studies. Unlike administrative datasets in which ICD-9 codes are used to obtain information about diagnosis and treatment, SCOAP relies on review of clinical records for consecutive patients undergoing specific procedures, including those who have appendectomies.

In 2008, we reported results from 15 SCOAP hospitals on the frequency of NA and use of imaging. That report noted substantial variation in both the use of imaging and in NA between hospitals and that NA correlated most closely with diagnostic accuracy (concordance between radiology reports and pathology reports).²⁴ The current report describes results from 55 hospitals over the last 5 years to: (1) investigate the association between the use of imaging and negative appendectomy in the general community, with a focus on patients at high-risk for misdiagnosis, (2) estimate performance characteristics of imaging modalities within a broad clinical environment, and (3) evaluate whether progress in safely reducing NA has continued as SCOAP expanded.

METHODS

Study population and setting

Although 55 Washington hospitals currently participate in SCOAP, hospital enrollment has been a gradual process as hospitals have joined each quarter over the past 6 years. Data are collected at the two pediatric hospitals in the state, and SCOAP also gathers information on children having operations at general hospitals; however, the current study population has been restricted to patients age 15 years or older who underwent appendectomy in a non-pediatric SCOAP hospital between January 1, 2006 and December 31, 2011. Participating hospitals submit data for all appendectomies performed within the institution for each enrollment year.

Data characteristics and primary outcome

Demographic information, clinical characteristics, radiology interpretations, operative indications, operative findings, and pathology results are abstracted from the clinical record using standardized definitions. Abstracted data are audited for quality control and to verify that charts are being evaluated in a similar way among participating sites. The data for appendectomy represent consecutive non-elective appendectomies performed at participating sites. A comorbidity index score, modeled on the Charlson comorbidity index, is calculated based on documentation in the clinical record of the following co-morbid conditions: coronary artery disease, asthma, diabetes, HIV/AIDS, diabetes, and/or elevated serum creatinine. White blood cell (WBC) count is based on the result obtained most proximal to the appendectomy. Body Mass Index (BMI) is calculated from recorded height and weight. Review of the patient's pathology report determines whether the appendix was diseased at the time of operation. Positive pathology results include confirmed or consistent with appendicitis or appendiceal tumor. Perforation of the appendix is based on pathologic diagnosis, and frequency of perforation was calculated excluding patients with NA (i.e., percent perforation = patients with perforation/all with positive appendiceal pathology). Imaging results are based on the final radiologist interpretation and are reported as consistent with appendicitis, not consistent with appendicitis, or indeterminate. An appendectomy is characterized as a NA in the absence of appendicitis or tumor/mass. The imaging report and pathology report are considered concordant if the imaging results are consistent with appendicitis and the pathology is positive or if imaging results are not consistent with appendicitis and pathology does not show evidence of disease. Indeterminate radiographic findings are considered discordant, regardless of pathologic findings. The primary outcome of interest was NA. Research projects utilizing SCOAP data are approved by the Washington State Department of Health Institutional Review Board.

Analytic Methods

Patients with appendiceal pathology were compared to those without appendiceal pathology to identify distinguishing characteristics between the two groups. Categorical variable comparisons were evaluated for significance using Pearson's chi-square test (significance set at $\alpha = 0.05$). Student's T-test was used to compare continuous variables ($\alpha = 0.05$). Odds ratios (and 95% confidence intervals) for variables predictive of misdiagnosis were calculated based on *a priori* hypotheses. A one-way analysis of variance model (multiple linear regression on a binary variable) was used to evaluate whether the proportion of NA differed significantly among comorbidity categories. Tests of trend over time were calculated using the Cochran-Armitage test for trends in the odds. Following the unadjusted analysis, we evaluated the association between imaging and NA for the presence of confounding by other covariates; variables potentially available to be included in this logistic regression model were those patient characteristics listed in Table 1. Covariates were included in this explanatory logistic regression model if they were known from the surgical literature or from clinical experience to be associated with misdiagnosis *and* if a differential association was detected in univariate analysis between the exposures of interest (i.e., imaging vs. no imaging) and the potential covariate. Using these criteria, a parsimonious, logistic regression model was developed that included age, gender, and WBC as covariates in the relationship between imaging use and NA. Using a generalized estimating equation, the model was also adjusted for clustering of patients by institution. Reproductive-aged women were previously identified as a group of patients at high risk for misdiagnosis; therefore, we separately considered a sub-cohort of women age 15 to 50. STATA version 12 was used for all analyses (STATA Corp., College Station, TX).

We estimated sensitivity and positive predictive value (PPV) for CT and US. Additionally we compared frequency of NA among patients imaged by the two most common modalities,

US and CT. An overall comparison was performed, and, because some institutions have imaging protocols based on age, we also made comparisons within three age groups (15–30, 31–65, and >65).

RESULTS

Cohort Characteristics

19,327 adolescent and adult patients underwent appendectomy (47.9% female, mean age 39.4 years, standard deviation 16.6). Ninety-one percent of patients underwent some form of pre-operative imaging (CT, US, and/or MRI). Among all patients with appendectomy, 1042 patients (5.4%) had NA. Overall frequency of perforation, as a percentage of patients with appendicitis, was 15.8 percent. Patients with NA were more often women, younger, and had a lower WBC count. BMI and comorbidity score were similar between patients with NA and patients with appendicitis. Equal proportions of patients underwent a laparoscopic procedure (Table 1).

Imaging vs. No Imaging

Among those with NA, a significantly smaller proportion received pre-operative imaging compared to those patients with appendicitis (75.3% vs. 92.6%, $p < 0.001$). For patients who had pre-operative imaging, the frequency of NA was 4.5%, significantly lower than the frequency of NA (15.4%) among those who did not have pre-operative imaging (odds ratio = 3.90, 95% CI 3.34 – 4.55, $p < 0.001$). After adjusting for gender, age, WBC count, and clustering by site, the odds of NA among those who did not undergo pre-operative imaging were 3.7 times the odds of NA for those who did undergo pre-operative imaging (95% CI 3.01 – 4.42, see Table 2). Adjusted for imaging, the odds ratio for NA among women compared to men was 2.10 (95% CI 1.76 – 2.51, $p < 0.001$). Although women were twice as likely as men to undergo NA, imaging among male patients was also associated with a significantly lower frequency of NA (3% vs. 10%, $p < 0.001$). Frequency of perforation was the same between patients who were and were not imaged: among those who were (and who had appendicitis), perforation was 15.8 percent; among those who were not, perforation was 15.6% ($p = 0.16$).

Reproductive-age Women

There were 6,632 women age 15–50 who underwent appendectomy, representing 34.4% of all appendectomies. Almost 95% underwent some form of diagnostic imaging. Among reproductive-aged women, frequency of NA was 8.1 percent. Nine percent of these patients were perforated compared to 15.8% in the entire cohort. Among women of reproductive age who received any form of pre-operative imaging, the frequency of NA was 6.9%, whereas it was 24.7% among reproductive-aged women who received no imaging (crude odds ratio 4.48, 95% CI 3.49 – 5.64, $p < 0.001$). In the multivariate model adjusted for age, WBC, and clustering by hospital, the odds of NA were 3.46 times the odds for those who did undergo pre-operative imaging (95% CI 2.43 – 4.94, see Table 3). Frequency of perforation was the same between those who had imaging and those who did not (9.9% vs. 9.7% respectively, $p = 0.48$).

Performance Characteristics of CT and US

Among all ages, 4.1% of patients who had CT underwent NA compared to 10.4% of patients who had US ($p < 0.001$). In the both the adolescent/young adult and middle-age categories, NA was significantly less common when CT was used compared to US (4.6% vs. 12% and 3.8% vs. 8.6%, respectively, $p < 0.001$ for both). Only 29 elderly patients underwent US, so a comparison versus CT was not considered robust in this age group. Among elderly patients

who underwent CT, frequency of NA was 3.6 percent. In patients who were not imaged, percent NA ranged from 14.1% to 16.3% depending on age group (Figure 1). The sensitivity of CT scan for appendicitis was estimated to be 93.2%, and for US, sensitivity was estimated to be 47.8 percent. PPV of CT scan was 97.6%, and for US, PPV was 94.3 percent.

Temporal Trends

We evaluated trends in imaging use and percent NA over the duration of SCOAP. The proportion of patients who received imaging in the workup of suspected appendicitis has been consistently rising (Figure 2). This is seen among SCOAP hospitals overall ($p < 0.001$), and also within hospital groups stratified by the year in which they joined SCOAP (though with more year-to-year variability). Concomitantly, in SCOAP overall, there has been a significant decline in the annual rate of NA ($p < 0.001$), though, again, there is year-to-year variability within subgroups of hospitals (Figure 3). Over this same time period, the percent of appendicitis patients who were perforated has not changed (Figure 4). Frequency of perforation ranged from 14.9% to 16.8%, but there was no temporal trend ($p = 0.63$).

DISCUSSION

In this cohort of older-adolescent and adult patients cared for in SCOAP hospitals over a six-year period, the use of advanced diagnostic imaging increased and the frequency of NA decreased. Among patients who received pre-operative imaging, NA was substantially less frequent than among patients who did not receive pre-operative imaging. When this relationship was adjusted for other predictors of negative appendectomy, failure to obtain imaging was associated with a 3.7-fold increase in odds of NA. Among women of reproductive age, the relationship with imaging was especially pronounced (25% NA vs. 7% NA). However, the age- and gender-adjusted regression suggests that, even among men, there is a strong association between pre-operative imaging and decreased odds of NA. As a group, SCOAP hospitals have prioritized the use of diagnostic imaging in the evaluation of suspected appendicitis as part of a commitment to safely reducing unnecessary operations. Although yearly variation is evident, data over the last six years suggest that these goals are being met by SCOAP hospitals.

The sensitivity of CT scan in this population (93.2%) was lower than some of the studies of CT in the highly-structured environment of studies in academic centers; however, this is within the range reported in the literature. Cumulative sensitivity of ultrasound studies in SCOAP hospitals was disturbingly low at 47.8%. Close inspection of this data revealed that a large number of patients with indeterminate results on ultrasonography were ultimately found to have appendicitis at appendectomy, which substantially reduced the modality's sensitivity. PPV for both studies was high (94% and 98%), however, suggesting that positive results on either CT scan or US are useful findings in the evaluation of a patient with suspected appendicitis.

One of the aims of this study was to evaluate diagnostic performance of CT and US in the general community as it compares to that published in the literature. A rigorous meta-analysis of CT and US published by Doria *et al* in 2006, included 57 studies (both retrospective and prospective) and more than 13,000 patients.²⁸ Studies were included only if absolute numbers of true-positives, true-negatives, false-positives, and false-negatives were available and adults and children were considered separately. In adults, sensitivity and specificity of CT were both 94%, and for US, sensitivity was 83% and specificity 93%. More recently, a large single-center study prospectively evaluated CT performance in 2871 consecutive adults imaged for suspected appendicitis and obtained thorough clinical follow-up of operative and non-operative patients; sensitivity was 98.5%, specificity was 98%,

NPV was 99.5%, and PPV was 93.9 percent.²⁹ Other recent studies have shown similar high performance for CT,^{30,31} including one study that evaluated low-dose radiation CT.³² Regarding ultrasound, Rettenbacher *et al* prospectively followed 350 patients evaluated by US for suspected appendicitis and determined a sensitivity of 98%, a specificity of 98%, a PPV of 96%, and a NPV of 99 percent.¹ The sensitivity we estimated for CT among SCOAP hospitals is within the 95% confidence interval reported by Doria (92%–95%), but as a group, SCOAP hospitals have not achieved the high bar set by studies performed in academic centers. Furthermore, although US had a substantial positive predictive value, the frequency of equivocal results limited its performance in terms of sensitivity. Certainly, for surgeons to include imaging results in their clinical decision making, they have to have confidence in the results, and SCOAP has made imaging accuracy a priority. Performance measures and statewide benchmarks for CT and US accuracy are provided to participating hospitals, and SCOAP is collaborating with radiology colleagues to address mechanisms for improvements.

Accurate imaging can provide three important functions in the evaluation of suspected appendicitis: provide evidence for a diagnosis of appendicitis, provide evidence against a diagnosis of appendicitis, and suggest alternative diagnoses. All are important, but the current study focused primarily on the second function. Reducing unnecessary operations is good for patients and for healthcare systems; previous studies have shown substantial increases in both length of stay and hospital charges for patients with NA compared to patients with appendicitis.³³ The current data from SCOAP hospitals suggest that the use of imaging is associated with a reduction in NA in the general community. Two other recent studies have assessed this association prospectively in patients with suspected appendicitis, and in both, pre-operative imaging changed management decisions, reducing negative appendectomy.^{2,3} In one of these studies, 152 patients were randomized to mandatory CT or selective-CT based on clinical examination. In the mandatory CT group, the frequency of NA was 2.6% versus 13.9% in the selective-CT group with no difference in perforation.² In addition to these prospective studies, numerous observational, retrospective analyses of appendectomy patients have shown an association between increased use of imaging and a decrease in the frequency of NA.^{4–14} This association was found for pediatric patients in some studies⁴ but not all.^{5,34}

The current study has several limitations. How patients were allocated to imaging or no-imaging is not captured by our dataset, and although the logistic regression models control for confounding by age, gender, and WBC (a marker of clinical severity), unmeasured confounding by indication may still be present. It is possible that this would lead to a conservative bias if complex or clinically uncertain patients were more likely to undergo imaging. The potential influence of laparoscopy on the measured frequency of NA is also uncertain. Administrative database analysis from the 1990s in Washington state³⁵ suggested that patients undergoing laparoscopy were more likely to have NA than those patients undergoing open appendectomy; however, a later analysis of SCOAP hospitals showed that there was no trend between a hospital's use of laparoscopy and frequency of NA.¹⁴ This latest analysis of SCOAP data is consistent with the latter finding in that patients undergoing laparoscopy were no more likely to have NA than those undergoing open appendectomy. Since SCOAP does not collect data for patients who undergo laparoscopy but do not have appendectomy if no appendicitis is found, the contribution of exploratory laparoscopy to decreasing NA cannot be judged from this dataset. Finally, in an earlier SCOAP study, in which hospitals were the unit of analysis, the correlation between accuracy of imaging (defined as pathology and radiology concordance) and institutional rates of NA was stronger than the correlation between NA and frequency of imaging-use.²⁴ In the current study, which treats patients as the unit of analysis, the impact of accuracy was not assessed. Although this would not be expected to change the association between the use of imaging

and frequency of NA (because accuracy of imaging does not impact patients who are not imaged), it could confound the inference that increased imaging among SCOAP hospitals has led to less NA among SCOAP hospitals. If accuracy is also improving, institutional rates of NA could decrease both from improved accuracy as well as from increased use of imaging. This is the topic of ongoing analyses. There may also be other variables beyond indication for imaging, use of laparoscopy, and accuracy of imaging that are unmeasured confounders. One solution to such confounding would be a statewide trial that randomized patients to mandatory imaging or selective-use of imaging, but this may not be feasible.

A further potential limitation is the possibility for sampling bias since the SCOAP cohort does not represent a truly random sample of the state's total appendectomy volume; however, for this to substantially alter the study results, those hospitals not participating would have to be outliers in terms of appendicitis care. By the end of 2011, 55 of the 75 hospitals in the state that perform appendectomies were actively contributing data to SCOAP. Contributors include both pediatric hospitals in the state and the state's active-duty military hospital, but it does not include the Veterans Affairs Medical Center in Seattle. The 20 hospitals that do not contribute to SCOAP are diverse in size, geographic location, and ownership; of those not enrolled, median appendectomy volume for 2010 was 22 cases, and only 4 non-enrolled hospitals performed greater than 100 cases. Utilizing the Washington State Comprehensive Hospital Abstract Reporting System (CHARS), which collects information on all discharges from non-federal hospitals in the state, we estimated Washington's total 2010 volume of non-incidental, non-elective appendectomy (for patients 18 years only) to be 6124 cases. SCOAP collected 5005 such cases for the same year, representing 82% of the state's total appendectomy volume. For 2011, CHARS data was not available, but with the addition of 6 new hospitals to SCOAP between the end of 2010 and the end of 2011, we expect that the proportion of the state's appendectomies captured by SCOAP has continued to increase as it has every year since 2006, SCOAP's first year. In 2006, data was captured from 14 hospitals, representing approximately 20% of the state's appendectomy volume.

Finally, our estimates of diagnostic performance (sensitivity and PPV) also involve limitations. Because this patient cohort is generated by patients who undergo appendectomy, data on most "true negatives" are not available. Patients who were correctly determined by CT or US not to have appendicitis were not included in this dataset unless the study was over-ruled, the patient was operated on, and had a NA. This makes a determination of specificity [true negatives/(true negatives + false positives)] and negative predictive value [true negatives/(true negatives + false negatives)] impossible. However, if it is assumed that very few patients with acute appendicitis do not undergo appendectomy, estimations of sensitivity [true positives/(true positives + false negatives)] and positive predictive value (PPV) [true positives/(true positives + false positives)] are possible. There may be some loss of "true positives" if the scan was overruled by the physicians, the patient discharged and ultimately had an appendectomy at another hospital (if the patient returned to the same hospital, the original CT information would be captured by SCOAP). There may be some loss of "false positives" if the CT scan was correctly overruled and the patient did not proceed to surgery. There may be some loss of "false negatives," if the patient was discharged and ultimately had an appendectomy at another hospital. Loss of "true positives" tend to reduce the observed performance of the imaging modality; loss of "false negatives" and "false positives" tend to increase the observed performance.

Many of the overlapping issues that arise in a consideration of imaging in suspected appendicitis are areas of active investigation and collaboration among SCOAP-affiliated surgeons and an increasingly broad coalition of academic and community radiologist partners. Given the previously detected association between accuracy of imaging and

reductions in NA, we are currently developing a standardized CT report for imaging in suspected appendicitis that will soon undergo piloting and validation testing. Attention to CT radiation dose, a variable newly captured by SCOAP, has revealed substantial variation in levels of radiation delivered during CT scan for appendicitis. Standardization of dose levels may be one way of reducing unnecessarily high radiation exposure, and the potential benefit to patient safety is being investigated by the SCOAP community. There is an ongoing effort to compare accuracy of CT scans in which IV and enteral contrast are both used to CT scans in which only IV contrast is used; given the time and cost savings that accrue from not using oral contrast, plus the advantage of avoiding oral intake among patients who typically feel very poorly, abandoning oral contrast has potential for significant improvements in the CT evaluation of appendicitis. For care of patients with suspected appendicitis, these developments represent some of the latest efforts within this physician-led system of continuous quality improvement.

The current investigation evaluated the association between imaging and NA across a large population served by diverse institutions. The data suggest that including pre-operative imaging in the workup of suspected appendicitis can lead to a reduction in unnecessary operations, especially among women of reproductive age; these modalities may also uncover alternative diagnoses (e.g., gynecologic pathology), some of which (e.g., Crohn's disease) are better managed non-operatively. Data from SCOAP further suggest that CT is more effective than ultrasound at accurately detecting acute appendicitis. In populations for which ionizing radiation is a concern, however, sequenced algorithms of US followed by CT scan for inconclusive US results may be appropriate. This latest report from SCOAP demonstrates the value of programs that facilitate collaborative, peer-driven quality improvement based on benchmarks for processes of care and for outcomes.

References

1. Rettenbacher T, Hollerweger A, Gritzmann N, et al. Appendicitis: should diagnostic imaging be performed if the clinical presentation is highly suggestive of the disease? *Gastroenterology*. 2002; 123(4):992–8. [PubMed: 12360459]
2. Lee CC, Golub R, Singer AJ, et al. Routine versus selective abdominal computed tomography scan in the evaluation of right lower quadrant pain: a randomized controlled trial. *Acad Emerg Med*. 2007; 14(2):117–22. [PubMed: 17192450]
3. Santos DA, Manunga J Jr, Hohman D, et al. How often does computed tomography change the management of acute appendicitis? *Am Surg*. 2009; 75(10):918–21. [PubMed: 19886134]
4. Applegate KE, Sivitt CJ, Salvator AE, et al. Effect of cross-sectional imaging on negative appendectomy and perforation rates in children. *Radiology*. 2001; 220(1):103–7. [PubMed: 11425980]
5. Bendeck SE, Nino-Murcia M, Berry GJ, et al. Imaging for suspected appendicitis: negative appendectomy and perforation rates. *Radiology*. 2002; 225(1):131–6. [PubMed: 12354996]
6. Coursey CA, Nelson RC, Patel MB, et al. Making the diagnosis of acute appendicitis: do more preoperative CT scans mean fewer negative appendectomies? A 10-year study. *Radiology*. 2010; 254(2):460–8. [PubMed: 20093517]
7. Guller U, Rosella L, McCall J, et al. Negative appendectomy and perforation rates in patients undergoing laparoscopic surgery for suspected appendicitis. *Br J Surg*. 2011; 98(4):589–95. [PubMed: 21259233]
8. Jones K, Pena AA, Dunn EL, et al. Are negative appendectomies still acceptable? *Am J Surg*. 2004; 188(6):748–54. [PubMed: 15619494] *Piper Am J Surg*. 2008
9. Piper HG, Rusnak C, Orrom W, et al. Current management of appendicitis at a community center--how can we improve? *Am J Surg*. 2008; 195(5):585–8. discussion 588–9. [PubMed: 18367143]
10. Rao PM, Rhea JT, Rattner DW, et al. Introduction of appendiceal CT: impact on negative appendectomy and appendiceal perforation rates. *Ann Surg*. 1999; 229(3):344–9. [PubMed: 10077046]

11. Rao PM, Feltmate CM, Rhea JT, et al. Helical computed tomography in differentiating appendicitis and acute gynecologic conditions. *Obstet Gynecol.* 1999; 93(3):417–21. [PubMed: 10074991]
12. Wagner PL, Eachempati SR, Soe K, et al. Defining the current negative appendectomy rate: for whom is preoperative computed tomography making an impact? *Surgery.* 2008; 144(2):276–82. [PubMed: 18656636]
13. Webb, EM.; Nguyen, A.; Wang, ZJ., et al. *AJR Am J Roentgenol.* Vol. 197. United States: 2011. The negative appendectomy rate: who benefits from preoperative CT?; p. 861-6.
14. Weyant MJ, Eachempati SR, Maluccio MA, et al. Interpretation of computed tomography does not correlate with laboratory or pathologic findings in surgically confirmed acute appendicitis. *Surgery.* 2000; 128(2):145–52. [PubMed: 10922984]
15. Horton MD, Counter SF, Florence MG, et al. A prospective trial of computed tomography and ultrasonography for diagnosing appendicitis in the atypical patient. *Am J Surg.* 2000; 179(5):379–81. [PubMed: 10930484]
16. Walker S, Haun W, Clark J, et al. The value of limited computed tomography with rectal contrast in the diagnosis of acute appendicitis. *Am J Surg.* 2000; 180(6):450–4. discussion 454–5. [PubMed: 11182396]
17. 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. [PubMed: 21238665]
18. Gaitini D, Beck-Razi N, Mor-Yosef D, et al. Diagnosing acute appendicitis in adults: accuracy of color Doppler sonography and MDCT compared with surgery and clinical follow-up. *AJR Am J Roentgenol.* 2008; 190(5):1300–6. [PubMed: 18430847]
19. Garcia Pena BM, Mandl KD, Kraus SJ, et al. Ultrasonography and limited computed tomography in the diagnosis and management of appendicitis in children. *JAMA.* 1999; 282(11):1041–6. [PubMed: 10493202]
20. Krishnamoorthi R, Ramarajan N, Wang NE, et al. Effectiveness of a staged US and CT protocol for the diagnosis of pediatric appendicitis: reducing radiation exposure in the age of ALARA. *Radiology.* 2011; 259(1):231–9. [PubMed: 21324843]
21. Poletti PA, Platon A, De Perrot T, et al. Acute appendicitis: prospective evaluation of a diagnostic algorithm integrating ultrasound and low-dose CT to reduce the need of standard CT. *Eur Radiol.* 2011; 21(12):2558–66. [PubMed: 21805194]
22. Poortman P, Oostvogel HJ, Bosma E, et al. Improving diagnosis of acute appendicitis: results of a diagnostic pathway with standard use of ultrasonography followed by selective use of CT. *J Am Coll Surg.* 2009; 208(3):434–41. [PubMed: 19318006]
23. Ramarajan N, Krishnamoorthi R, Barth R, et al. An interdisciplinary initiative to reduce radiation exposure: evaluation of appendicitis in a pediatric emergency department with clinical assessment supported by a staged ultrasound and computed tomography pathway. *Acad Emerg Med.* 2009; 16(11):1258–65. [PubMed: 20053244]
24. Cuschieri J, Florence M, Flum DR, et al. Negative appendectomy and imaging accuracy in the Washington State Surgical Care and Outcomes Assessment Program. *Ann Surg.* 2008; 248(4): 557–63. [PubMed: 18936568]
25. Huynh V, Lalezarzadeh F, Lawandy S, et al. Abdominal computed tomography in the evaluation of acute and perforated appendicitis in the community setting. *Am Surg.* 2007; 73(10):1002–5. [PubMed: 17983068]
26. Frei SP, Bond WF, Bazuro RK, et al. Appendicitis outcomes with increasing computed tomographic scanning. *Am J Emerg Med.* 2008; 26(1):39–44. [PubMed: 18082779]
27. Sarkaria IS, Eachempati SR, Weyant MJ, et al. Current surgical opinion of computed tomography for acute appendicitis. *Surg Infect (Larchmt).* 2004; 5(3):243–52. [PubMed: 15684795]
28. 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.
29. Pickhardt PJ, Lawrence EM, Pooler BD, et al. Diagnostic performance of multidetector computed tomography for suspected acute appendicitis. *Ann Intern Med.* 2011; 154(12):789–96. W-291. [PubMed: 21690593]

30. Dearing DD, Recabaren JA, Alexander M. Can computed tomography scan be performed effectively in the diagnosis of acute appendicitis without the added morbidity of rectal contrast? *Am Surg.* 2008; 74(10):917–20. [PubMed: 18942613]
31. Kim K, Rhee JE, Lee CC, et al. Impact of helical computed tomography in clinically evident appendicitis. *Emerg Med J.* 2008; 25(8):477–81. [PubMed: 18660392]
32. Platon A, Jlassi H, Rutschmann OT, et al. Evaluation of a low-dose CT protocol with oral contrast for assessment of acute appendicitis. *Eur Radiol.* 2009; 19(2):446–54. [PubMed: 18797875]
33. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. *Arch Surg.* 2002; 137(7):799–804. discussion 804. [PubMed: 12093335]
34. Weyant MJ, Eachempati SR, Maluccio MA, et al. The use of computed tomography for the diagnosis of acute appendicitis in children does not influence the overall rate of negative appendectomy or perforation. *Surg Infect (Larchmt).* 2001; 2(1):19–23. [PubMed: 12594877]
35. Flum DR, Morris A, Koepsell T, et al. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. *JAMA.* 2001; 286(14):1748–53. [PubMed: 11594900]

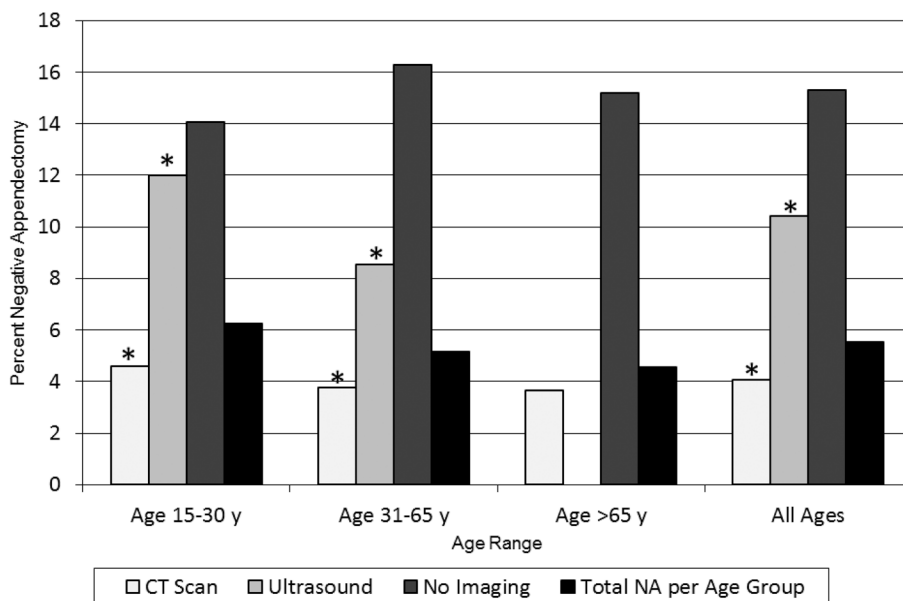


Figure One. Percent Negative Appendectomy by Imaging Modality, Stratified by Age Range
 Asterisks indicate statistically significant comparison within each age category. No US results are reported for the elderly age group; only 29 elderly patients underwent preoperative US, which was felt too small of a number to make a comparison valid.

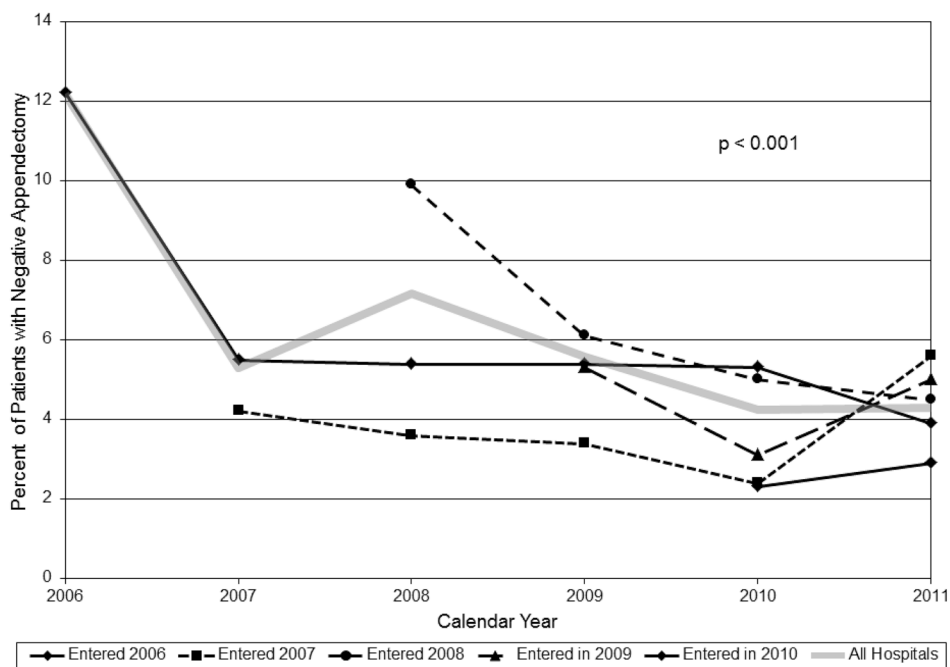


Figure Two. Appendectomy Patients Who Received Pre-operative Imaging
 Hospitals are stratified by the year in which they joined SCOAP (represented by black lines that begin with the year in which that group of hospitals first submitted cases to SCOAP). The solid gray line indicates the number of patients who underwent pre-operative imaging as a percentage of all appendectomy patients from all hospitals participating in SCOAP in any particular year. Test for trend for the overall percentage of pre-operative imaging was significant ($p < 0.001$).

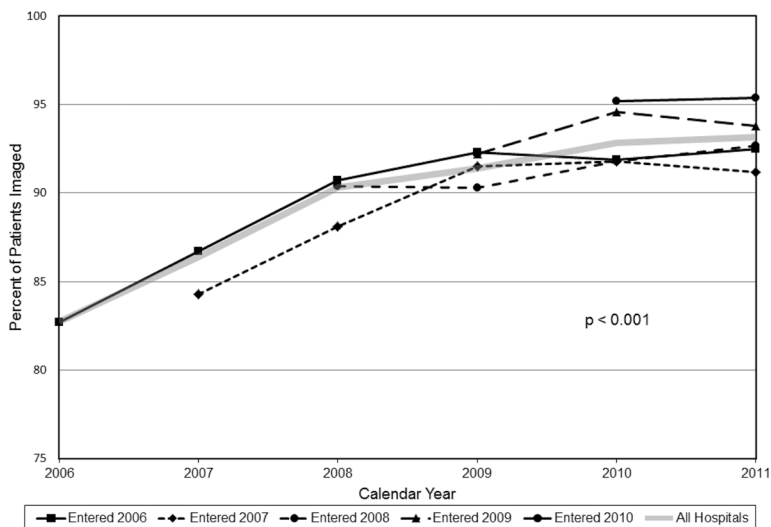


Figure Three. Appendectomy Patients with Negative Appendectomy

Hospitals are stratified by the year in which they joined SCOAP (represented by black lines that begin with the year in which that group of hospitals first submitted cases to SCOAP). The solid gray line indicates the number of negative appendectomies as a percentage of all appendectomy patients from all hospitals participating in SCOAP in any particular year. Test for trend for the overall percentage of negative appendectomies was significant ($p < 0.001$).

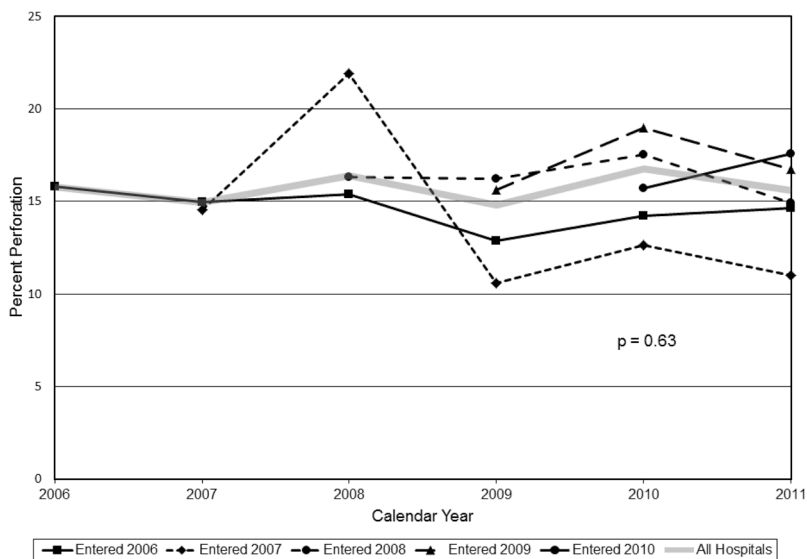


Figure Four. Appendicitis Patients with Perforation

Hospitals are stratified by the year in which they joined SCOAP (represented by black lines that begin with the year in which that group of hospitals first submitted cases to SCOAP). The solid gray line indicates the number of patients with perforation as a percentage of all patients with appendicitis from all hospitals participating in SCOAP in any particular year. There was no temporal trend for the overall frequency of perforation ($p = 0.63$).

TABLE 1

All Appendectomy, Appendicitis, and Negative Appendectomy

	All Appendectomy (n = 19,327)	Appendicitis (n = 18193)	Negative Appendectomy (n = 1042)	P-value
Number female patients	47.86%	46.9%	64.0%	P < 0.001 *
Mean age (SD)	39.4 (16.6)	39.5 (16.6)	37.4 (15.9)	P < 0.001 **
Pre-operative Imaging	91.3%	92.2%	75.3%	p < 0.001 *
Negative Appendectomy	5.4%	–	–	
Laparoscopy	84.8%	84.9%	83.5%	P = 0.24 *
Perforation	15.0%	15.8%	–	
Mean WBC (SD)	13.2 (4.3)	13.3 (4.3)	10.4 (4.0)	P < 0.001 **
Mean BMI (SD)	27.7 (6.0)	27.8 (6.2)	27 (6.3)	P < 0.001 **
Co-morbidity Index:				NS
0	86.5%	86.6%	85.5%	
1	10.9%	10.8%	12.6%	
2	1.7%	1.7%	1%	
3	0.9%	0.9%	0.5%	

SD = standard deviation

* = Pearson Chi-squared test

** = Independent Student's T-test (two-sided)

P-values reflect results of statistical comparison between patients with appendicitis and patients with NA

Table 2

Multiple Logistic Regression on Odds of Negative Appendectomy (full cohort)

	Odds Ratio	95% Confidence Interval
Imaging vs. No Imaging (unadjusted)	3.90	3.34 – 4.55
Imaging vs. No Imaging (adjusted)	3.65	3.01 – 4.42
Gender (female vs. male)	2.10	1.76 – 2.51
Age	0.99	0.98 – 0.99
WBC count	0.86	0.84 – 0.89

Table 3

Multiple Logistic Regression on Odds of Negative Appendectomy (reproductive-age women)

	Odds Ratio	95% Confidence Interval
Imaging vs. No Imaging (unadjusted)	4.48	3.49 – 5.64
Imaging vs. No Imaging (adjusted)	3.46	2.42 – 4.94
Age	0.99	0.99 – 1.0
WBC count	0.86	0.84 – 0.88

Outcome of interest is negative appendectomy; primary exposure is the use of imaging versus no imaging (imaging could be CT scan or US). Potential confounders (gender, age, WBC count) were identified *a priori* and assessed in a univariate fashion (see Table 1) prior to inclusion in the regression model. WBC and age are included as continuous variables. A generalized estimating equation function was utilized to adjust the model for clustering by hospital.