

ORIGINAL ARTICLE

Exploring diagnostic stewardship in the emergency department evaluation of pediatric abdominal pain in a statewide quality collaborative

Alexander T. Janke MD, MHS, MSc^{1,2}  | Kenneth A. Michelson MD, MPH³  |
Keith E. Kocher MD, MPH^{1,2,4,5} | Kristian Seiler BS² | Michelle L. Macy MD, MS^{3,6}  |
Michele Nypaver MD^{1,7}  | Prashant V. Mahajan MD, MBA, MPH^{1,2}  | Rajan Arora MD⁸  |
Courtney W. Mangus MD, MHS^{1,2,7,9} 

¹Department of Emergency Medicine, University of Michigan, Ann Arbor, Michigan, USA

²Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, Michigan, USA

³Division of Pediatric Emergency Medicine, Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

⁴Department of Learning Health Sciences, University of Michigan, Ann Arbor, Michigan, USA

⁵Center for Clinical Management Research, Ann Arbor Veterans Affairs Healthcare System, Ann Arbor, Michigan, USA

⁶Mary Ann & J. Milburn Smith Child Health Outcomes, Research, and Evaluation Center, Stanley Manne Children's Research Institute, Ann & Robert Lurie Children's Hospital of Chicago, Chicago, Illinois, USA

⁷Department of Pediatrics, University of Michigan, Ann Arbor, Michigan, USA

⁸Children's Hospital of Michigan, Detroit, Michigan, USA

⁹Hurley Medical Center, Flint, Michigan, USA

Correspondence

Alexander T. Janke, Department of Emergency Medicine, University of Michigan, Ann Arbor, MI 48109, USA.
Email: atjanke@med.umich.edu

Abstract

Background: Diagnostic stewardship is the effort to optimize diagnostic testing to reduce errors while avoiding overtesting and overtreatment. Abdominal pain and appendicitis in children are essential use cases. Delayed diagnosis of appendicitis can be dangerous and even life-threatening, but overtesting is harmful.

Methods: We conducted a retrospective cohort study of children aged 5–17 years presenting with abdominal pain to 26 EDs within the Michigan Emergency Department Improvement Collaborative (MEDIC) from May 1, 2016, to February 29, 2024. We defined two outcome measures summarized by ED. First, we describe the cross-sectional imaging:appendicitis visits ratio, defined as the count of ED visits resulting in any cross-sectional imaging (CT or MRI) divided by the count of ED visits with a diagnosis of appendicitis. Second, we describe the delayed diagnosis rate, defined by an ED visit for abdominal pain resulting in a discharge and subsequent return visit with a diagnosis of appendicitis within 7 days.

Results: The sample included 120,112 pediatric visits for abdominal pain at 26 EDs; 4967 (4.1%) were diagnosed with appendicitis. The cross-sectional imaging:appendicitis visits ratio varied by site, from as low as 0.2 (95% confidence interval [CI] 0.1–0.2) at a pediatric site to as high as 7.9 (95% CI 4.8–16.4) at an urban ED. The proportion of pediatric ED visits for abdominal pain that resulted in an identified delayed diagnosis of appendicitis was 0.1% (141/120,112). All but four sites had fewer than 10 cases of delayed diagnosis across the study period.

Conclusions: In this retrospective cohort study of 120,000+ ED visits for pediatric abdominal pain, we found that the ratio of visits with cross-sectional imaging to diagnosed cases of appendicitis varied widely across EDs. Delayed diagnosis of appendicitis was uncommon. Adherence to best practices and improved imaging

Supervising Editor: Jody A. Vogel

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

quality may hold promise to improve diagnostic stewardship for children with abdominal pain across EDs.

INTRODUCTION

Diagnostic stewardship is the effort to optimize diagnostic testing to reduce errors while avoiding overtesting and overtreatment.¹ Originally developed with application to infectious disease, the concept is broadly applicable in emergency medicine where clinicians must balance the harms of overtesting with the dangers of missed diagnoses. Abdominal pain and appendicitis in children are canonical examples. Abdominal pain is among the most common reasons children present to the emergency department (ED), and appendicitis is the most common surgical cause.^{2–4} Delayed diagnosis of appendicitis in children with abdominal pain can be dangerous and even life-threatening,⁵ but overtesting is wasteful and sometimes harmful.^{6,7}

The diagnosis of appendicitis is generally confirmed by abdominal imaging. Ultrasound is the safest and most cost-effective initial imaging modality in the evaluation of children with suspected appendicitis⁸ and is recommended by most guidelines.⁹ Ultrasound avoids exposure to ionizing radiation, which is particularly harmful to children.^{10,11} However, ultrasound is not always available around the clock across ED settings and is operator dependent, so nondiagnostic scans are common.¹² Cross-sectional imaging (computed tomography [CT] or, less commonly, magnetic resonance imaging [MRI]) is typically obtained if appendicitis is suspected and ultrasound is either unavailable or nondiagnostic.¹³ Clinicians in emergency medicine rely on their expertise and experience as well as diagnostic tools like ultrasound and blood tests. In some cases, ED clinicians may have access to onsite surgical consultation or be able to observe such patients for serial abdominal examinations. These resources help clinicians to reduce cross-sectional imaging overtesting while guarding against potentially dangerous delayed diagnosis of appendicitis. Using a statewide quality collaborative of diverse community, academic, urban, and rural EDs, our objectives were (1) to describe current practice patterns in use of imaging in the ED evaluation of nontraumatic pediatric abdominal pain and (2) to assess the associated rate of delayed diagnosis of appendicitis among those patients.

METHODS

Design, setting, and population

We conducted a retrospective cohort study of children aged 5–17 years presenting to 26 EDs within the Michigan Emergency Department Improvement Collaborative (MEDIC) from May 1, 2016, February 29, 2024. MEDIC is a quality improvement network with a clinical data registry supporting >40 hospital members across Michigan to enhance ED quality improvement through techniques such as audit and feedback measurement and reporting, evidence

translation to practice change, and knowledge generation and dissemination.¹⁴ There were 26 included MEDIC sites caring for pediatric patients with ongoing participation in the collaborative as of July 2024 that were included in the study. The University of Michigan Institutional Review Board approved this study, informed consent was waived, and we followed the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines.¹⁵

The MEDIC data registry includes electronic health record–extracted data elements for all ED visits among participating sites. We identified children with a chief complaint and/or a triage note indicating the presence of abdominal pain using a bidirectional encoder representations from transformers (BERT) language model previously developed specifically for application in emergency care.¹⁶ The BERT language model is a natural language processing tool that allows for the rapid identification of abdominal pain within millions of records with unique free-text chief complaints and triage notes. Table A1 depicts characteristic examples of free-text complaints and their associated labels for abdominal pain and cough.

Children with a diagnosis of trauma, pregnancy, inflammatory bowel disease, systemic lupus erythematosus, sickle cell disease, and malignancy were excluded from the analytic sample because of unique clinical considerations within these populations. In addition, a prior diagnosis of appendicitis within the MEDIC registry associated with the patient record was similarly excluded from the sample. Finally, patients with a first-time diagnosis of appendicitis but without associated diagnostic imaging were also excluded as the majority of these visits were identified as cases of ED transfers from referral hospitals likely after the diagnosis of appendicitis had been established.

Variables

Variables were obtained from the MEDIC data registry and included visit site, visit date, patient age, sex, race and ethnicity, insurance, presence of a complex chronic condition (CCC)¹⁷ among ED diagnosis codes, and abdominal imaging (ultrasound, CT, MRI). Visit diagnosis codes (International Classification of Diseases, 10th Edition, Clinical Modification [ICD-10]) were also collected. The CCCs are a taxonomy of conditions designed to capture medical conditions expected to last at least 12 months and to involve one or more organ systems requiring specialty pediatric care. Race and ethnicity data are variably reported across sites in the MEDIC data registry but are typically structured data fields or free-text drawn from site electronic health record systems. We categorized and reported on race and ethnicity fields according to guidance from the Michigan Social Health Interventions to Eliminate Disparities (MSHIELD) collaborative quality initiative.¹⁸ Sites were classified as either pediatric, urban, or rural. Pediatric sites were identified

as major academic tertiary pediatric specialty centers participating in MEDIC. Rurality was identified by county based on rural-urban commuting area (RUCA) codes consistent with prior research.¹⁹

Outcomes

To measure and compare variation in imaging practices, we created two diagnostic stewardship outcome metrics for children presenting to the ED with nontraumatic abdominal pain. First, we defined a cross-sectional imaging:appendicitis visits ratio as the ratio of the count of ED visits with at least one cross-sectional imaging study performed divided by the total number of ED visits resulting in a diagnosis of appendicitis. We chose this approach to avoid a potential pitfall of a diagnostic yield measure in this case and to gauge the extent to which sites rely on cross-sectional imaging as an evaluation tool relative to the amount of appendicitis diagnosed. Because cross-sectional imaging may optimally be reserved for equivocal cases, or instances where ultrasound imaging is unavailable, a higher yield (positivity rate) of cross-sectional imaging may not necessarily reflect a lower reliance on cross-sectional imaging overall. As a corollary, for descriptive reporting only, we also report a similar ultrasound:appendicitis visits ratio. Second, the prevalence of delayed diagnosis was defined as the percentage of children who presented with abdominal pain at an index ED visit and were subsequently evaluated, discharged, and returned to the ED with a diagnosis of appendicitis within 7 days. Patients seen in one ED, discharged, and later returning to another ED in the same health system can be identified. This definition of delayed diagnosis of appendicitis is relatively specific and has limitations, as it relies on patients having presented with a chief complaint or triage note indicating specifically abdominal pain at the first visit and may not capture all delayed diagnoses including those crossing between health systems within MEDIC or between MEDIC and non-MEDIC sites.

Statistical analysis

To ensure that this approach captured a cohort where appendicitis was a relevant diagnostic consideration, we described rates of imaging and appendicitis diagnoses within the defined abdominal pain cohort compared to all other pediatric visits within ages 5–17. We reported imaging use, the cross-sectional imaging:appendicitis visits ratio, and the delayed diagnosis rate by site as well as pooled across all ED visits and sites in the analysis. For each site cross-sectional imaging:appendicitis visits ratio, we calculated bootstrapped 95% confidence intervals (CIs). For each ED, we performed 1000 bootstrap resamples with replacement from the original visit data, calculating the ratio of visits with cross-sectional imaging to appendicitis diagnoses for each sample. The 95% CIs for these ratios were then determined using the empirical quantiles from the bootstrap distributions. To assess for an association between the cross-sectional

imaging:appendicitis visits ratio and delayed diagnosis rates at the site level, we constructed a scatterplot and reported a Spearman correlation. We compare available sociodemographic characteristics of appendicitis cases with and without delayed diagnoses. We use independent-sample t-tests and chi-square tests of independence to evaluate for differences in means and proportions as appropriate. As our cohort included some patients seen at multiple ED visits for abdominal pain, we also performed a sensitivity analysis where we included only first-time abdominal pain visits as contained within the MEDIC registry, and we report imaging utilization and cross-sectional imaging:appendicitis visits ratio at those visits alone. All analyses were performed using R (R Project for Statistical Computing Version 4.2.1).

RESULTS

After exclusions, there were a total of 120,112 ED visits for abdominal pain for 94,409 unique patients ages 5–17 and 26 EDs in the analytic sample (Figure 1). Appendicitis was diagnosed in 4967 (4.1%) visits presenting with abdominal pain. Three pediatric centers accounted for 59.9% of visits in the sample. Seventeen urban EDs accounted for 33.9% of the visits, and the remaining 6.2% of visits occurred at the six rural EDs included in this study. Overall, the sample was 59.3% female and 53.3% non-Hispanic White and had a mean age of 11 years. Table 1 depicts visit-level sociodemographic characteristics of the study population. Visits for abdominal pain as identified by our natural language processing tool captured the majority of abdominal imaging (95.5%) and appendicitis diagnoses (96.1%) among all pediatric ED visits ages 5–17 (Table A2).

Among all children identified with abdominal pain, 37,396 (31.1%) underwent abdominal imaging (Figure 2). Ultrasound was obtained in 27.2% (32,676) of visits, while 6.4% (7726) had a CT and 1.9% (2243) had a MRI. Sites varied in their propensity to image, from 12.2% to 50.3% of abdominal pain visits. Across the full study sample, the cross-sectional imaging:appendicitis visits ratio was 2.0 (9914 ED visits with cross-sectional imaging divided by 4967 appendicitis cases). The ratio varied by site, from as low as 0.2 (95% CI 0.1–0.2) at a pediatric site to as high as 7.9 (95% CI 4.8–16.4) at an urban ED (Figure 3). Figure A1 similarly depicts this ratio but for ultrasound:appendicitis visits and finds similar marked variation. We report on a sensitivity analysis, where we only include first-time ED visits across patients, in Table A3.

Across all sites and the full study period, there were 141 delayed diagnoses (2.8% of the 4967 total appendicitis cases) following an index ED visit for abdominal pain with appendicitis diagnosed at a second ED visit within 7 days. Figure A2 depicts raw imaging rates against delayed diagnosis rates, by site. Imaging rates were not generally statistically significantly associated with delayed diagnosis rates. Table A4 depicts characteristics of missed diagnosis cases compared to all other appendicitis cases. Females made up a larger proportion of delayed diagnoses compared to other appendicitis cases (48.2 vs. 39.9%, $p=0.047$) as did Hispanic patients (19.1 vs.

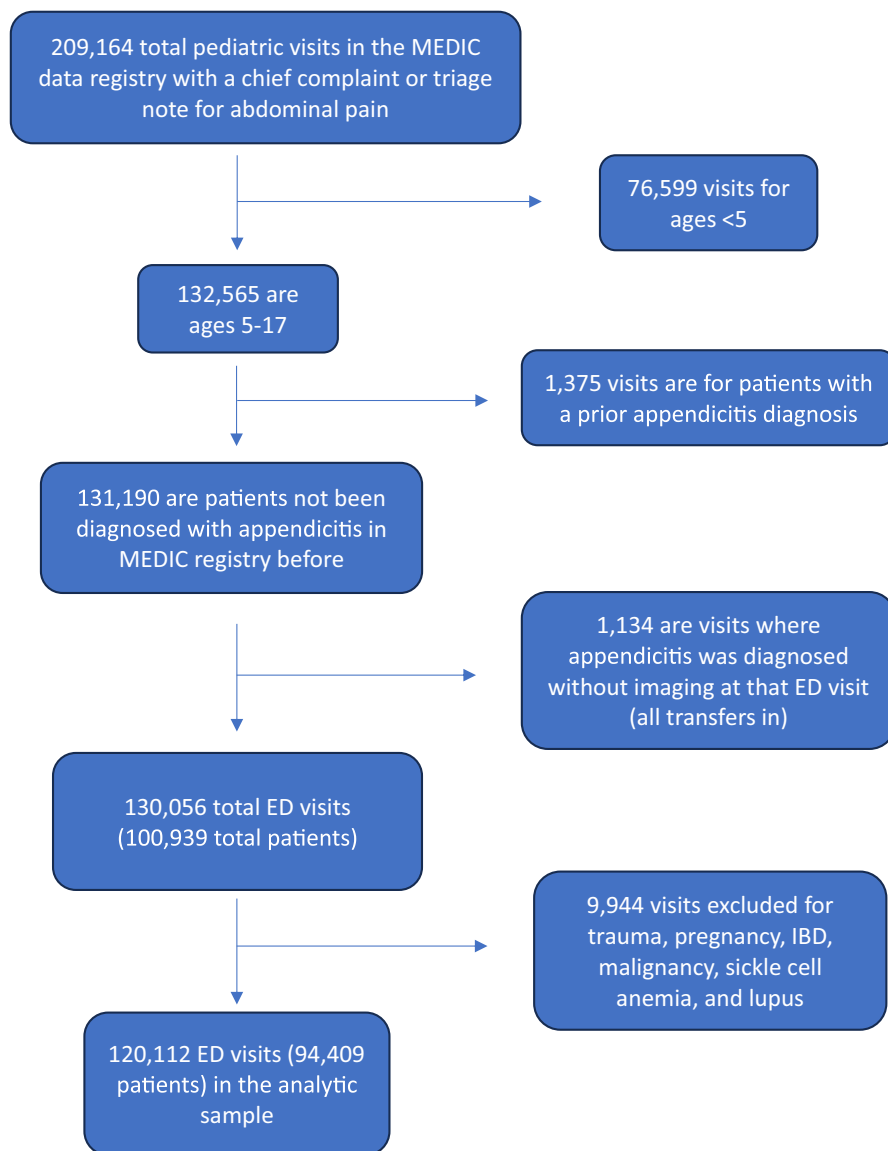


FIGURE 1 Study cohort. Analytic sample constructed for pediatric abdominal pain where appendicitis is likely a relevant diagnostic consideration.

12.4%, $p=0.018$) and self-pay patients (5.7 vs. 0.8%, $p<0.001$). The overall proportion of pediatric ED visits for abdominal pain that resulted in an identified delayed diagnosis of appendicitis was 0.1% (141/120,112). All but four sites had fewer than 10 cases of delayed diagnosis across the study period. Figure 4 depicts cross-sectional imaging:appendicitis visits ratio against delayed diagnosis rates at the site level. The Spearman correlation coefficient for the association of these two measures as in Figure 4 was -0.065 ($p=0.75$), indicating no statistically significant relationship. Figure A2 depicts the association of raw imaging use by type with delayed diagnosis rates. Table A5 shows visit counts, imaging frequency, and appendicitis diagnoses by site.

DISCUSSION

In this cohort study of over 120,000 pediatric visits for abdominal pain at 26 geographically and structurally diverse EDs, we found that

over 30% of children had at least one form of imaging. Over 4% of children had appendicitis, similar to previously described epidemiology of acute pediatric abdominal pain within EDs.^{3,20} Consistent with guidelines,^{5,21} ultrasound was the most frequently employed imaging modality followed by CT and then MRI. Sites varied widely in both their propensity to image and the number of cross-sectional imaging visits compared to number of visits resulting in a diagnosis of appendicitis. For the three pediatric EDs and for one urban ED, the cross-sectional imaging:appendicitis visits ratio was <2 . That is, for every case of appendicitis diagnosed at those EDs, fewer than two cross-sectional studies are performed. For the remainder of sites, this cross-sectional imaging:appendicitis visits ratio varied from 2.5 to nearly 8. Identified delayed diagnoses were uncommon among the cohort who presented with abdominal pain at an index ED visit in the prior 7 days, representing 2.8% of total cases of pediatric appendicitis in the sample.

Pediatric EDs exhibit both relatively low cross-sectional imaging:appendicitis visits ratio and low delayed diagnosis rates, markers

TABLE 1 Visit level characteristics by type of ED.

	Pediatric (n = 3)	Urban EDs (n = 17)	Rural EDs (n = 6)	Overall (n = 26)
Age (years)				
Mean	10.6	11.8	11.7	11.0
SD	3.9	4.0	4.0	4.0
Female	41,384 (57.5)	25,265 (62)	4594 (62)	71,243 (59.3)
Race/Ethnicity				
Asian	928 (1.3)	588 (1.4)	19 (0.3)	1535 (1.3)
Black	20,927 (29.1)	6083 (14.9)	123 (1.7)	27,133 (22.6)
Hispanic	8788 (12.2)	5392 (13.2)	518 (7)	14,698 (12.2)
Indigenous	203 (0.3)	265 (0.7)	108 (1.5)	576 (0.5)
MENA	1587 (2.2)	<10	<10	1587 (1.3)
Multiple	5602 (7.8)	199 (0.5)	84 (1.1)	5885 (4.9)
NHPI	69 (0.1)	91 (0.2)	14 (0.2)	174 (0.1)
White	31,605 (43.9)	25,963 (63.7)	6424 (86.8)	63,992 (53.3)
Payer				
Private	52,411 (72.8)	18,913 (46.4)	5984 (80.8)	77,308 (64.4)
Public	16,531 (23)	17,844 (43.8)	761 (10.3)	35,136 (29.3)
Self-pay	2535 (3.5)	252 (0.6)	76 (1)	2863 (2.4)
CCC	4547 (6.3)	728 (1.8)	184 (2.5)	5459 (4.5)
Total visits	71,965	40,742	7405	120,112

Note: Data are reported as n (%).

Abbreviations: CCC, complex chronic condition (at least one); MENA, Middle Eastern/North African; NHPI, Native Hawaiian/Pacific Islander.¹⁶

of diagnostic stewardship along potentially competing dimensions. The low cross-sectional imaging:appendicitis visits ratios and delayed diagnosis rates are consistent with current literature and likely reflects the relatively greater use and higher diagnostic yield of ultrasonography in diagnosing appendicitis at pediatric sites compared to general EDs. Pediatric EDs typically employ pediatric sonographers and pediatric radiologists who are highly experienced in obtaining and interpreting abdominal ultrasounds obtained on children. Prior studies have demonstrated that nondiagnostic ultrasounds are less common at pediatric EDs compared to community EDs.^{22,23} For one site in our data with a cross-sectional imaging:appendicitis visits ratio of just 0.2, clinical leaders indicated after the present analyses were completed that surgical consultation and observation for serial abdominal examinations are commonly employed in lieu of cross-sectional imaging.

Variability in imaging between sites suggests an opportunity to improve. Overtesting raises several concerns, including exposure to ionizing radiation, increasing cost of care, and prolonging the ED visit.^{7,24-26} Previous work has established that, at least in

pediatric EDs, the increasing use of ultrasound in diagnosing pediatric appendicitis has resulted in a decline in CT imaging without a subsequent increase in appendiceal perforation or ED revisits.²⁷ Other studies have also described the significant variability in ED imaging strategies for suspected appendicitis, though the majority of such studies have focused exclusively on practice variability within pediatric EDs, limiting their generalizability.²⁸⁻³⁰ Benchmarking practice patterns in choice of imaging for pediatric abdominal pain presentations (including the cross-sectional imaging:appendicitis visits ratio) has not been previously described. Given the inherent differences in available resources among EDs, establishing a universal and fair benchmark target may prove challenging. However, as more than 80% of children requiring emergency care are evaluated in general EDs,³¹ standardizing the diagnostic process, including enhancing the capacity to obtain high-quality ultrasound images and accurately interpret results in children and identifying ways to improve pediatric outcomes in these settings, is especially critical.

Despite the challenges, there are opportunities to improve on the problems of overtesting and delayed diagnosis particularly in general EDs with limited resources. Incorporating bloodwork, especially the serum white blood cell (WBC) count, into the diagnostic evaluation has been shown to reduce CT imaging and delayed diagnoses.^{32,33} Further, validated clinical decision tools can help identify children at very low risk of appendicitis who do not require imaging.³⁴⁻³⁶ The pediatric Appendicitis Risk Calculator (pARC) is a newer tool that uses clinical data and the WBC count to stratify children with suspected appendicitis into risk categories, and it has been validated for use in the community ED setting.^{36,37} Use of pARC was projected to reduce delayed diagnoses by about 50% in one study.³⁸ One of the tool's limitations is in its application to children with a nondiagnostic ultrasound, but a recent study by Bravo et al.³⁹ quantified the risk of appendicitis in children with a nondiagnostic ultrasound and a low-risk pARC score and demonstrated that cross-sectional imaging can safely be deferred in these cases. While the study by Bravo et al. requires further testing and validation in general EDs, its premise is promising in enhancing risk stratification tools to guide pediatric imaging and reduce overtesting.

In addition to clinical decision support, a recent joint policy statement from the American Academy of Pediatrics, the American College of Emergency Physicians, and the American College of Radiology identified two additional strategies to reduce overimaging in suspected pediatric appendicitis.²¹ First, for patients with high likelihood of appendicitis who will require transfer regardless of the CT results (for appendectomy or further management of non-surgical abdominal pain), it was recommended that imaging should be deferred to the pediatric specialty center.²¹ Second, when ultrasound is equivocal or unavailable, ED clinicians are encouraged to use shared decision making to explain the available diagnostic options and allow the family to make an informed choice, from watchful waiting at home with close follow-up, ED/inpatient setting observation, and immediate CT imaging to transfer to a pediatric center for further management. Ultimately, for

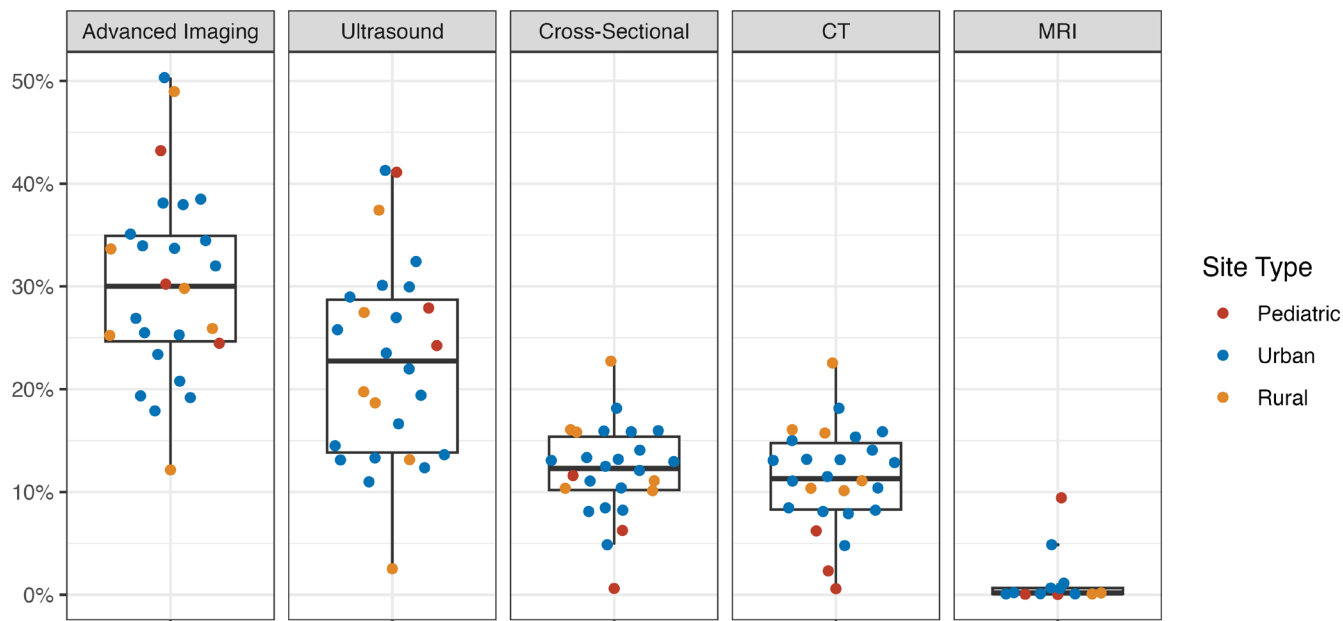


FIGURE 2 Abdominal imaging practices by ED. The proportion of pediatric abdominal pain visits for children ages 5-17 years undergoing abdominal imaging by type in the MEDIC registry. Advanced imaging is inclusive of any one of ultrasound, CT, or MRI. Cross-sectional imaging is inclusive of any one of CT or MRI. The sample includes 26 unique ED sites. MEDIC, Michigan Emergency Department Improvement Collaborative.

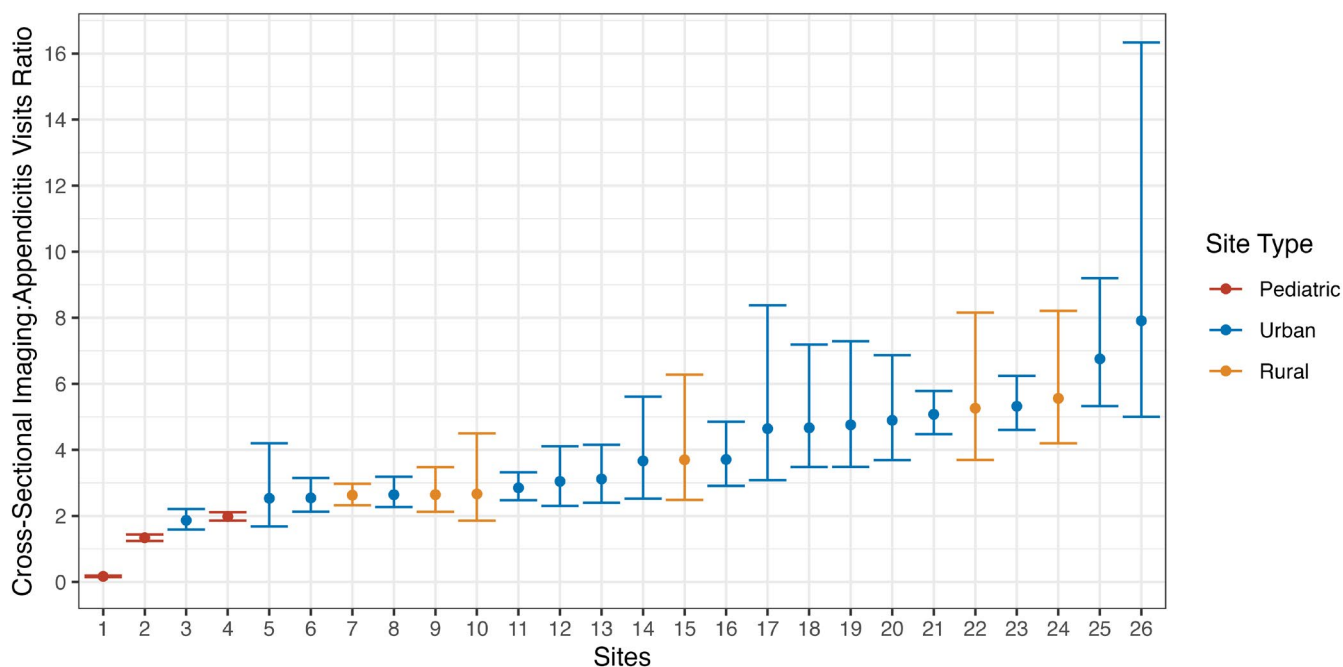


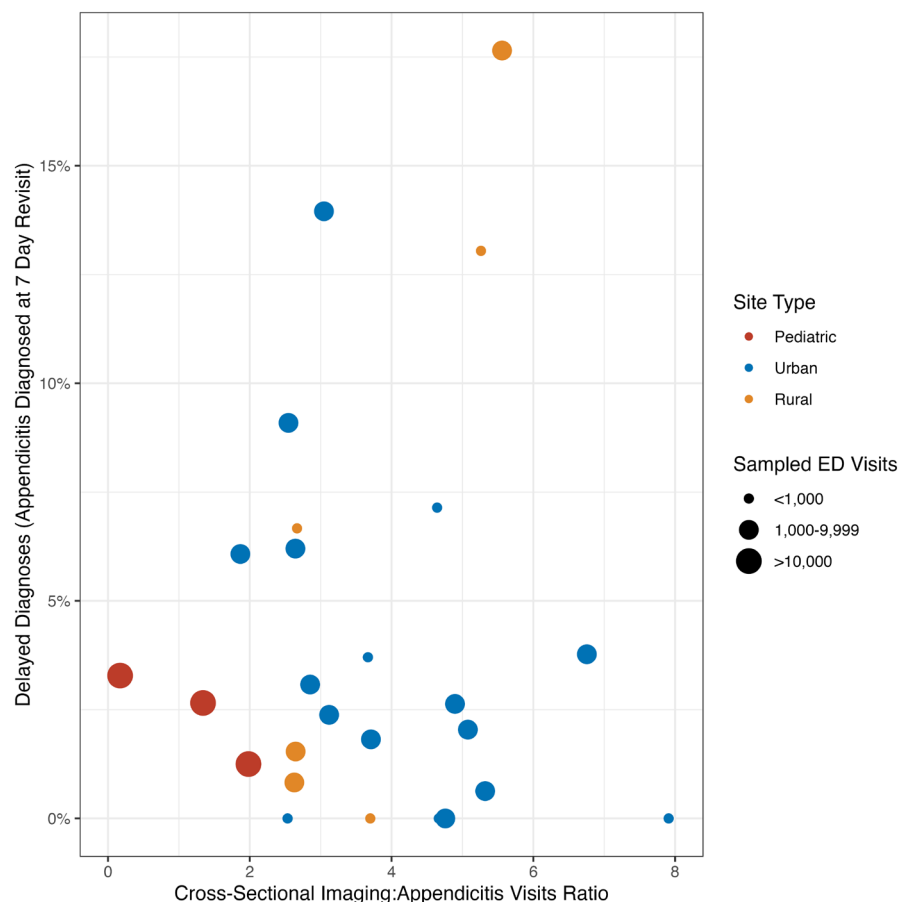
FIGURE 3 Cross-sectional Imaging:appendicitis visits ratio by ED. Cross-sectional imaging:appendicitis visits ratio by ED site in the sample. The cross-sectional imaging:appendicitis visits ratio is defined by the ratio of ED visits with at least one cross-sectional imaging study performed divided by the number of appendicitis cases diagnosed. Bootstrapped 95% CIs are provided.

every child with abdominal pain, an ED clinician must use clinical judgment to balance the risks of diagnostic imaging with the potential harms of delayed diagnosis. Future studies should include a cost-effectiveness analysis to compare these options based on an EDs available resources.

LIMITATIONS

Our study has several limitations. We used the application of a natural language processing tool to identify abdominal pain in the chief complaint or triage note and define the cohort, which may have

FIGURE 4 Cross-sectional imaging use and delayed diagnosis by ED. Estimates of diagnostic stewardship in pediatric abdominal pain across sites in the Michigan Emergency Department Improvement Collaborative. The cross-sectional imaging:appendicitis visits ratio is defined by the number of ED visits with at least one cross-sectional imaging study (CT and/or MRI) performed divided by the number of appendicitis cases diagnosed. Delayed diagnosis of appendicitis is reported as a percentage of overall appendicitis cases and is defined as an ED visit for abdominal pain resulting in discharge and return within 7 days with a new diagnosis of appendicitis. The Spearman correlation coefficient is -0.065 ($p=0.75$).



missed some children presenting without classic symptoms of appendicitis. However, we took additional steps to ensure the validity of our cohort and ensure the majority of appendicitis diagnoses in the sample were captured (Table A2). Additionally, defining potential delays in appendicitis diagnosis as a 7-day return visit misses patients who presented to EDs outside of the MEDIC data registry for their second visit, and therefore our estimate of delayed diagnosis can be expected to underestimate the true rate (and is indeed lower than other recent U.S.-based estimates).⁴⁰ This bias would tend to underestimate delay rates at general EDs compared with pediatric EDs, as patients at these EDs may be more likely to go to a different ED at the second visit.⁴¹ Further, the majority of appendicitis cases were evaluated in pediatric EDs, which is another source of bias when calculating the rate of delayed diagnosis. Finally, while we quantified delayed diagnosis of pediatric appendicitis, we did not assess whether this delay resulted in patient harm or morbidity, including appendiceal perforation, sepsis, bowel resection, or abdominal abscess drainage.⁴² It is likely that some cases of diagnostic delay reflect scenarios where shared decision making was utilized and watchful waiting was initially selected over immediate imaging; however, our data set does not contain information about the decision-making process, the locally available resources, or laboratory testing. Finally, there are other indications for cross-sectional imaging in the evaluation of pediatric patients with abdominal pain—including renal stones or bowel obstruction—that may account for

greater reliance on these studies in certain populations and for which we did not fully assess in this analysis.

CONCLUSIONS

In this retrospective cohort study of 120,000+ ED visits for pediatric abdominal pain, we found that the cross-sectional imaging:appendicitis visits ratio varied widely across EDs. In our sample, delayed diagnosis of appendicitis was uncommon. Pediatric EDs had both the lowest imaging and the lowest delayed diagnosis rates, likely partially reflective of the quality of pediatric ultrasound at those sites. Adherence to best practices, improved imaging availability and quality, and strengthened coordination of care across centers may hold promise to improve diagnostic stewardship in the evaluation of children with abdominal pain across EDs.

FUNDING INFORMATION

Support for MEDIC is provided by Blue Cross Blue Shield of Michigan (BCBSM) and Blue Care Network (BCN) as a part of the BCBSM Value Partnerships program. Although Blue Cross Blue Shield of Michigan and MEDIC work collaboratively, the opinions, beliefs, and viewpoints expressed by MEDIC do not necessarily reflect the opinions, beliefs, and viewpoints of BCBSM or any of its employees.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Alexander T. Janke  <https://orcid.org/0000-0003-1181-714X>

Kenneth A. Michelson  <https://orcid.org/0000-0003-1763-7262>

Michelle L. Macy  <https://orcid.org/0000-0003-4401-7384>

Michele Nypaver  <https://orcid.org/0000-0001-7422-8700>

Prashant V. Mahajan  <https://orcid.org/0000-0003-4278-3560>

Rajan Arora  <https://orcid.org/0000-0001-9816-8140>

Courtney W. Mangus  <https://orcid.org/0000-0001-7957-1049>

REFERENCES

- Morgan DJ, Malani PN, Diekema DJ. Diagnostic stewardship to prevent diagnostic error. *JAMA*. 2023;329(15):1255-1256.
- Alpern ER, Stanley RM, Gorelick MH, et al. Epidemiology of a pediatric emergency medicine research network: the PECARN Core data project. *Pediatr Emerg Care*. 2006;22(10):689-699.
- Scholer SJ, Pituch K, Orr DP, Dittus RS. Clinical outcomes of children with acute abdominal pain. *Pediatrics*. 1996;98(4 Pt 1):680-685.
- Sømme S, Bronsert M, Morrato E, Ziegler M. Frequency and variety of inpatient pediatric surgical procedures in the United States. *Pediatrics*. 2013;132(6):e1466-e1472.
- Michelson KA, McGarghan FLE, Patterson EE, Samuels-Kalow ME, Waltzman MJ, Greco KF. Delayed diagnosis of serious paediatric conditions in 13 regional emergency departments. *BMJ Qual Saf*. 2024;33(5):293-300.
- Tung M, Sharma R, Hinson JS, Nothelle S, Pannikottu J, Segal JB. Factors associated with imaging overuse in the emergency department: a systematic review. *Am J Emerg Med*. 2018;36(2):301-309.
- Marin JR, Mills AM. Developing a research agenda to optimize diagnostic imaging in the emergency department: an executive summary of the 2015 academic emergency medicine consensus conference. *Academic Emergency Medicine*. 2015;22(12):1363-1371.
- Jennings R, Guo H, Goldin A, Wright DR. Cost-effectiveness of imaging protocols for suspected appendicitis. *Pediatrics*. 2020;145(2):e20191352.
- Rosen MP, Ding A, Blake MA, et al. ACR Appropriateness Criteria® right lower quadrant pain—suspected appendicitis. *J Am Coll Radiol*. 2011;8(11):S373-S387.
- Brenner D, Elliston C, Hall E, Berdon W. Estimated risks of radiation-induced fatal cancer from pediatric CT. *AJR Am J Roentgenol*. 2001;176(2):289-296.
- 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):700-707.
- Kotagal M, Richards MK, Flum DR, Acierno SP, Weinsheimer RL, Goldin AB. Use and accuracy of diagnostic imaging in the evaluation of pediatric appendicitis. *J Pediatr Surg*. 2015;50(4):642-646.
- Ramarajan N, Krishnamoorthi R, Gharahbaghian L, Pirrotta E, Barth RA, Wang NE. Clinical correlation needed: what do emergency physicians do after an equivocal ultrasound for pediatric acute appendicitis? *J Clin Ultrasound*. 2014;42(7):385-394.
- Kocher KE, Arora R, Bassin BS, et al. Baseline performance of real-world clinical practice within a statewide emergency medicine quality network: the Michigan emergency department improvement collaborative (MEDIC). *Ann Emerg Med*. 2020;75(2):192-205.
- von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-349.
- Chang D, Hong WS, Taylor RA. Generating contextual embeddings for emergency department chief complaints. *JAMIA Open*. 2020;3(2):160-166.
- Feudtner C, Feinstein JA, Zhong W, Hall M, Dai D. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr*. 2014;14:199.
- Michigan Social Health Interventions to Eliminate Disparities (MSHIELD). Collecting demographic data: best practices for collecting race, ethnicity, language, sexual orientation, gender identity, and disability data. <https://www.michiganshield.org/resources/data-strategy-quality-resources#h.i55f4hixtkfp>. Updated March 2024. Accessed 8/23/2024
- Venkatesh AK, Janke A, Rothenberg C, Chan E, Becher RD. National trends in emergency department closures, mergers, and utilization, 2005-2015. *PLoS One*. 2021;16(5):e0251729.
- Reynolds SL, Jaffe DM. Diagnosing abdominal pain in a pediatric emergency department. *Pediatr Emerg Care*. 1992;8(3):126-128.
- American Academy of Pediatrics Committee on Pediatric Emergency Medicine; Section on Radiology; American College of Emergency Physicians Pediatric Emergency Medicine Committee; American College of Radiology, Marin JR, Lyons TW, et al. Optimizing advanced imaging of the pediatric patient in the emergency department: policy statement. *Ann Emerg Med*. 2024;84(2):e13-e23.
- Alter SM, Walsh B, Lenehan PJ, Shih RD. Ultrasound for diagnosis of appendicitis in a community hospital emergency department has a high rate of nondiagnostic studies. *J Emerg Med*. 2017;52(6):833-838.
- Saito JM, Yan Y, Evashwick TW, Warner BW, Tarr PI. Use and accuracy of diagnostic imaging by hospital type in pediatric appendicitis. *Pediatrics*. 2013;131(1):e37-e44.
- Mathews JD, Forsythe AV, Brady Z, et al. Cancer risk in 680,000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. *BMJ*. 2013;346:f2360.
- Kanzaria HK, Probst MA, Ponce NA, Hsai R. The association between advanced diagnostic imaging and ED length of stay. *Am J Emerg Med*. 2014;32(10):1253-1258.
- Kocher KE, Meurer WJ, Desmond JS, Nallamotheu BK. Effect of testing and treatment on emergency department length of stay using a national database. *Acad Emerg Med*. 2012;19(5):525-534.
- Bachur RG, Levy JA, Callahan MJ, Rangel SJ, Monuteaux MC. Effect of reduction in the use of computed tomography on clinical outcomes of appendicitis. *JAMA Pediatr*. 2015;169(8):755-760.
- Thompson GC, Schuh S, Gravel J, et al. Pediatric emergency research Canada. Variation in the Diagnosis and Management of Appendicitis at Canadian Pediatric Hospitals. *Acad Emerg Med*. 2015;22(7):811-822.
- Bachur RG, Hennelly K, Callahan MJ, Monuteaux MC. Advanced radiologic imaging for pediatric appendicitis, 2005-2009: trends and outcomes. *J Pediatr*. 2012;160(6):1034-1038.
- Rice-Townsend S, Barnes JN, Hall M, Baxter JL, Rangel SJ. Variation in practice and resource utilization associated with the diagnosis and management of appendicitis at freestanding children's hospitals: implications for value-based comparative analysis. *Ann Surg*. 2014;259(6):1228-1234.
- Gausche-Hill M, Ely M, Schmuhl P, et al. A national assessment of pediatric readiness of emergency departments. *JAMA Pediatr*. 2015;169(6):527-534.
- Anandalwar SP, Graham DA, Kashtan MA, et al. Leveraging the combined predictive value of ultrasound and laboratory data to

- reduce radiation exposure and resource utilization in children with suspected appendicitis. *J Surg Res.* 2021;257:529-536.
33. Michelson KA, McGarghan FLE, Patterson EE, Waltzman ML, Samuels-Kalow ME, Greco KF. Clinician factors associated with delayed diagnosis of appendicitis. *Diagnosis (Berl).* 2022;10(2):183-186.
 34. Samuel M. Pediatric appendicitis score. *J Pediatr Surg.* 2002;37(6):877-881.
 35. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986;15(5):557-564.
 36. Kharbanda AB, Vazquez-Benitez G, Ballard DW, et al. Development and validation of a novel pediatric appendicitis risk calculator (pARC). *Pediatrics.* 2018;141(4):e20172699.
 37. Cotton DM, Vinson DR, Vazquez-Benitez G, et al. Validation of the Pediatric Appendicitis Risk Calculator (pARC) in a Community Emergency Department Setting. *Ann Emerg Med.* 2019;74(4):471-480.
 38. Michelson KA, Reeves SD, Grubenhoff JA, et al. Clinical features and preventability of delayed diagnosis of pediatric appendicitis. *JAMA Netw Open.* 2021;4(8):e2122248.
 39. Bravo M, Palnizky-Soffer G, Man C, et al. Identification of children with a nondiagnostic ultrasound at a low appendicitis risk

using a pediatric appendicitis risk calculator. *Acad Emerg Med.* 2024;31(12):1256-1263.

40. Michelson KA, Bachur RG, Rangel SJ, Monuteaux MC, Mahajan P, Finkelstein JA. Emergency department volume and delayed diagnosis of pediatric appendicitis: a retrospective cohort study. *Ann Surg.* 2023;278(6):833-838.
41. Michelson KA, Lyons TW, Bachur RG, Monuteaux MC, Finkelstein JA. Timing and location of emergency department revisits. *Pediatrics.* 2018;141(5):e20174087.
42. Michelson KA, Dart AH, Bachur RG, Mahajan P, Finkelstein JA. Measuring complications of serious pediatric emergencies using ICD-10. *Health Serv Res.* 2021;56(2):225-234.

How to cite this article: Janke AT, Michelson KA, Kocher KE, et al. Exploring diagnostic stewardship in the emergency department evaluation of pediatric abdominal pain in a statewide quality collaborative. *Acad Emerg Med.* 2025;32:309-319. doi:[10.1111/acem.15075](https://doi.org/10.1111/acem.15075)

APPENDIX A

TABLE A1 Examples of free text chief complaint labels.

Abdominal pain
"belly pain, nausea, vomiting"
"x2 days abd pain, -v"
"ptn presents with mom with n/v and abdominal pain after eating"
Cough
"cough, nasal congestion"
"x3 days cough, not eating, siblings sick"
"cough, nasal congestion, COVID+ at home"

Note: Examples of free text chief complaints according to their labels. We used a previously developed bidirectional encoder representations from transformers (BERT) language model,¹⁶ which accepts as input free text for chief complaints and triage notes, and outputs a standardized chief complaint 'label.'

TABLE A2 Advanced abdominal imaging and appendicitis diagnoses among pediatric ED visits with and without abdominal pain.

	No abdominal pain	Abdominal pain
Total visits	701,837	132,565
Total patients	403,462	102,071
Imaging, US	1.1%	26.2%
Imaging, CT	0.6%	6.6%
Imaging, MRI	0.004%	1.8%
Imaging, any advanced	1.6%	30.3%
Appendicitis primary	0.026%	4.3%
Appendicitis any	0.053%	4.8%

Note: Counts of children with and without a chief complaint or triage note for abdominal pain as identified by the applied language model, and the rate of advanced abdominal imaging and appendicitis diagnoses in those groups.

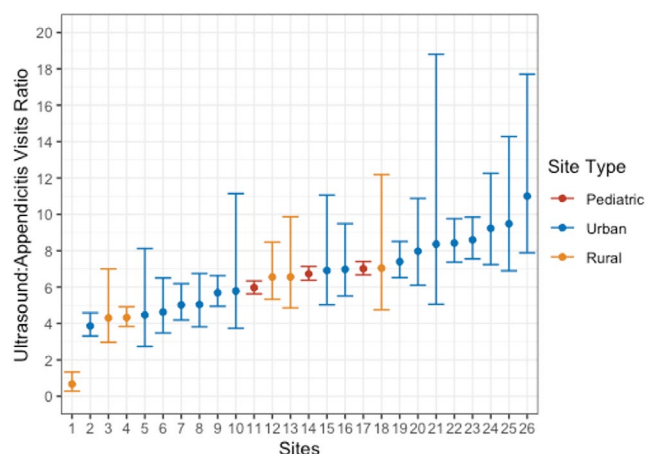


FIGURE A1 Ultrasound: appendicitis visits ratio by ED. Ultrasound: Appendicitis visits ratio by ED site in the sample. The ultrasound:Appendicitis visits ratio is defined by the ratio of ED visits with at least one ultrasound imaging study performed divided by the number of appendicitis cases diagnosed. Bootstrapped 95% confidence intervals are provided.

TABLE A3 Sensitivity analysis for imaging use and cross-sectional imaging:appendicitis visits ratio, including only patients' first time abdominal pain ED visits in data.

	Pooled	Site median [IQR]
Ultrasound	26,643 (28.2%)	23.7 [14.3%–30.0%]
CT	5924 (6.3%)	11.2 [8.2%–14.5%]
MRI	1877 (2.0%)	0 [0%–0%]
CSI:appendicitis ratio	1.72	3.10 [2.24–4.44]

Note: In a sensitivity analysis where first-time abdominal pain visits (excluding subsequent ED visits for the same patient in the data), the sample size is N=94,409.

Abbreviation: CSI:appendicitis ratio, cross-sectional imaging:appendicitis visits ratio.

TABLE A4 Sociodemographics of appendicitis cases by delayed versus not delayed diagnosis.

	Delayed (N = 141)	Not delayed (N = 4826)	p-value
Age (mean)	11.1	11.6	0.172
Female (%)	48.2	39.9	0.047
Race/Ethnicity (%)			
Asian	<10	1.4%	N/A
Black	12.1%	9.6%	0.322
Hispanic	19.1%	12.4%	0.018
Indigenous	<10	0.5%	N/A
MENA	<10	1.8%	N/A
Multiple	<10	2.8%	N/A
NHPI	<10	0.2%	N/A
White	57.4%	64.8%	0.070
Payer (%)			
Private	74.5%	75.1%	0.871
Public	14.9%	21.1%	0.080
Self-Pay	5.7%	0.8%	<0.001
CCC (%)	<10	2.9%	N/A

Note: Sociodemographics and complex chronic conditions among appendicitis cases that were identified as delayed (N=141) and those not identified as delayed diagnoses (N=4826). p-values are reported for pairwise difference in mean (independent sample T-tests) and difference in proportion (chi-squared tests).

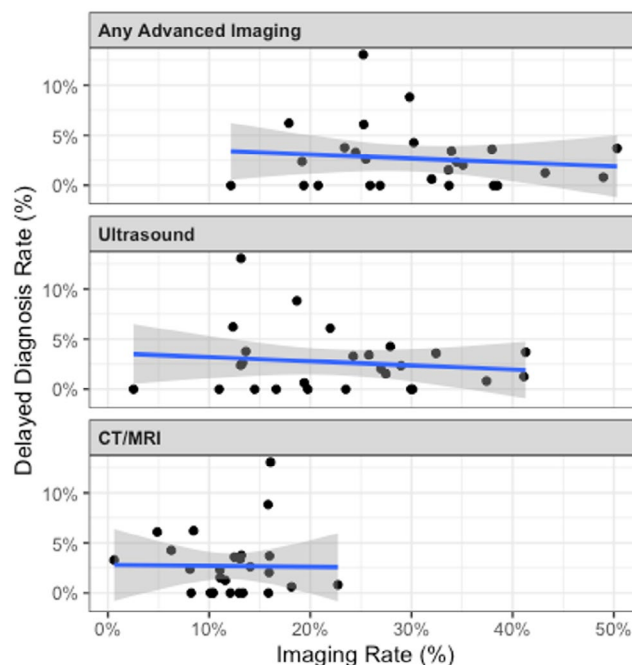


FIGURE A2 Association of imaging use to delayed diagnosis rate by site. Site-level association between the imaging rate (proportion of pediatric abdominal pain visits undergoing imaging, by type) against the delayed diagnosis rate (proportion of ED visits for appendicitis identified as delayed diagnoses). Advanced imaging refers to any of ultrasound, CT, or MRI. Line of best fit is depicted with 95% confidence intervals. Raw associations are not statistically significant throughout.

TABLE A5 Site-level characteristics and event rates.

Site	Type	Abdominal pain visits	Any imaging		Cross-sectional imaging		Appendicitis	
1	Major Pediatric	31,282	7655	24.5%	192	0.6%	1127	3.6%
2	Major Pediatric	20,184	6101	30.2%	1263	6.3%	942	4.7%
3	Major Pediatric	20,499	8858	43.2%	2378	11.6%	1200	5.9%
4	Urban	6247	2192	35.1%	995	15.9%	196	3.1%
5	Urban	6937	1754	25.3%	338	4.9%	181	2.6%
6	Urban	2715	635	23.4%	358	13.2%	53	2.0%
7	Urban	4663	1492	32.0%	846	18.1%	159	3.4%
8	Urban	837	162	19.4%	87	10.4%	11	1.3%
9	Urban	4450	1689	38.0%	556	12.5%	195	4.4%
10	Rural	1551	522	33.7%	172	11.1%	65	4.2%
11	Urban	4030	721	17.9%	341	8.5%	129	3.2%
12	Urban	1616	310	19.2%	131	8.1%	42	2.6%
13	Urban	1687	643	38.1%	204	12.1%	55	3.3%
14	Rural	2798	1370	49.0%	636	22.7%	242	8.6%
15	Urban	1322	337	25.5%	186	14.1%	38	2.9%
16	Rural	1194	356	29.8%	189	15.8%	34	2.8%
17	Urban	1714	582	34.0%	224	13.1%	88	5.1%
18	Urban	1184	408	34.5%	131	11.1%	43	3.6%
19	Rural	714	185	25.9%	74	10.4%	20	2.8%
20	Rural	753	190	25.2%	121	16.1%	23	3.1%
21	Urban	462	96	20.8%	38	8.2%	15	3.2%
22	Urban	1065	410	38.5%	138	13.0%	29	2.7%
23	Urban	706	238	33.7%	112	15.9%	24	3.4%
24	Rural	395	48	12.2%	40	10.1%	15	3.8%
25	Urban	487	131	26.9%	65	13.3%	14	2.9%
26	Urban	620	312	50.3%	99	16.0%	27	4.4%

Note: Site-level visit counts for the primary analysis. Percentages are given where the denominator is total abdominal pain visits, except for delayed diagnoses, which are reported as a percentage of appendicitis cases. Cumulative delayed diagnoses for appendicitis by site to protect privacy of sites and in accordance with existing data use agreements. Cumulative delayed diagnosis counts for appendicitis observed were 77 (2.4%), 51 (3.9%), and 13 (3.3%) for major pediatric, urban, and rural sites, respectively.