OPEN

Perception of Treatment Success and Impact on Function with Antibiotics or Appendectomy for Appendicitis

A Randomized Clinical Trial with an Observational Cohort

Writing Group for the CODA Collaborative, Callie M. Thompson, MD,*† Emily C. Voldal, BA,‡
Giana H. Davidson, MD, MPH,‡ Sabrina E. Sanchez, MD, MPH,§ Patricia Ayoung-Chee, MD, MPH,‖¶
Jesse Victory, DO,** Mary Guiden,‡ Bonnie Bizzell, MEd, MBA,‡ Jacob Glaser, MD, FACS,††
Christopher Hults, MD,†† Thea P. Price, MD,‡‡ Nicole Siparsky, MD,‡‡ Kristin Ohe, MD,§§
Katherine A. Mandell, MD, MPH,§§ Daniel A. DeUgarte, MD,¶¶ Amy H. Kaji, MD, PhD,¶¶
Lisandra Uribe, BS,‖‖ Lillian S. Kao, MD, MS,*** Krislynn M. Mueck, MD,***
Farhood Farjah, MD, MPH,‡ Wesley H. Self, MD, MPH,* Sunday Clark, ScD,§††
F. Thurston Drake, MD, MPH,§ Katherine Fischkoff, MD,‡‡ Elizaveta Minko, MD,‡‡
Joseph Cuschieri, MD,§§§¶¶¶ Brett Faine, PharmD, MS,‖‖‖‖ Dionne A. Skeete, MD,‖‖‖‖
Naila Dhanani, MD,**** Mike K. Liang, MD,****†††† Anusha Krishnadasan, PhD,‖‖
David A. Talan, MD,‖‖‡‡‡‡ Erin Fannon, BA,‡ Larry G. Kessler, ScD,‡ Bryan A. Comstock, MS,‡
Patrick J. Heagerty, PhD,‡ Sarah E. Monsell, MS,‡ Sarah O. Lawrence, MA,‡
David R. Flum, MD, MPH,‡✉ and Danielle C. Lavallee, PhD, PharmD‡§§§§

Objective: To compare secondary patient reported outcomes of perceptions of treatment success and function for patients treated for appendicitis with appendectomy vs. antibiotics at 30 days.

Summary Background Data: The Comparison of Outcomes of antibiotic Drugs and Appendectomy trial found antibiotics noninferior to appendectomy based on 30-day health status. To address questions about outcomes among participants with lower socioeconomic status, we explored the relationship of sociodemographic and clinical factors and outcomes.

Methods: We focused on 4 patient reported outcomes at 30 days: high decisional regret, dissatisfaction with treatment, problems performing usual activities, and missing > 10 days of work. The randomized (RCT) and observational cohorts were pooled for exploration of baseline factors. The RCT cohort alone was used for

From the *Vanderbilt University Medical Center, Nashville, TN; †University of Utah, Salt Lake City, UT; ‡University of Washington, Seattle, WA; §Boston University Medical Center, Boston, MA; ¶Tisch Hospital NYU Langone Medical Center, New York, NY; ‖Grady Health, Morehouse School of Medicine, Atlanta, GA; **Bellevue Hospital Center NYU School of Medicine, New York, NY; ††Providence Regional Medical Center Everett, Everett, WA; ‡‡Rush University Medical Center, Chicago, IL; §§ Swedish Medical Center, Seattle, WA; ¶¶Harbor-UCLA Medical Center, Torrance, CA; ‖‖Olive View-UCLA Medical Center, Sylmar, CA; ***McGovern Medical School at The University of Texas Health Science Center at Houston, Houston, TX; †††Weill Cornell Medicine, New York, NY; ‡‡‡Columbia University Medical Center, New York, NY; §§§Harborview Medical Center, Seattle, WA; ¶¶¶University of California, San Francisco, San Francisco, CA; ‖‖‖‖University of Iowa Hospitals and Clinics, Iowa City, IA; ****University of Texas Lyndon B. Johnson General Hospital, Houston, TX; ††††University of Houston, HCA Healthcare Kingwood, Kingwood, TX; ‡‡‡‡Ronald Reagan UCLA Medical Center, Westwood, CA; and §§§§BC Academic Science Health Network, Vancouver, BC, Canada.

comparison of treatments. Logistic regression was used to assess associations.

Results: The pooled cohort contained 2062 participants; 1552 from the RCT. Overall, regret and dissatisfaction were low whereas problems with usual activities and prolonged missed work occurred more frequently. In the RCT, those assigned to antibiotics had more regret (Odd ratios (OR) 2.97, 95% Confidence intervals (CI) 2.05–4.31) and dissatisfaction (OR 1.98, 95%CI 1.25–3.12), and reported less missed work (OR 0.39, 95%CI 0.27–0.56). Factors associated with function outcomes included sociodemographic and clinical variables for both treatment arms. Fewer factors were associated with dissatisfaction and regret.

Conclusions: Overall, participants reported high satisfaction, low regret, and were frequently able to resume usual activities and return to work. When comparing treatments for appendicitis, no single measure defines success or failure for all people. The reported data may inform discussions regarding the most appropriate treatment for individuals.

Trial Registration: Clinicaltrials.gov Identifier: NCT02800785.

With growing evidence that antibiotic management can be effective for appendicitis,[1–10] patients have a nonsurgical option for treatment. Researchers and clinicians have largely focused on the rate of appendectomy after initiating antibiotics as the main clinical outcome of antibiotic treatment, but patient reported outcomes (PROs) related to treatment success and impact on function are relevant and less studied outcomes. Prior trials have been limited in study size, lack of heterogeneity of participant characteristics, in the methods, and delayed timing[11] of PRO asssessments.[1–9] In distinction, the Comparison of Outcomes of antibiotic Drugs and Appendectomy (CODA)[10] randomized Clinical Trial (RCT) included 1552 adults who were randomized to appendectomy or antibiotics. CODA used a PRO (EQ-5D)[12] at 30 days as its primary outcome and also included a battery of standardized secondary PROs. In addition, CODA had a parallel observational group of 510 participants who declined randomization, chose their treatment, and were surveyed in the same manner as the randomized cohort.

The CODA trial found antibiotics were noninferior to appendectomy when considering the EQ-5D, with other outcomes variously favoring antibiotics or appendectomy (eg, time until return to work and subsequent hospitalization, respectively).[10] The aim of this study is to compare secondary PROs related to perceptions of treatment success (dissatisfaction and decisional regret) and function (problems with usual activities and at least 2 weeks of missed work) between the two treatment arms at 30 days. To address questions[13] about potentially worse outcomes among people with lower socioeconomic status, we conducted an exploratory analysis to compare secondary PROs in specific subgroups.

METHODS

CODA

The CODA trial is a pragmatic, nonblinded, noninferiority, randomized clinical trial designed to assess whether antibiotic treatment for appendicitis is as effective as surgery. The study was conducted at 25 US medical centers participating in the Comparative Effectiveness Research Translation Network, based at the University of Washington.[14] The trial was funded by the Patient-Centered Outcomes Research Institute. Complete details of the trial design and protocol have been published elsewhere.[10,15]

Data Description

These analyses focus on 4 PROs reported at 30 days after presentation, a time window commonly used to describe outcomes for surgical care.

Decisional Regret

Participants were asked, "Think about the decision to follow through with your initial treatment of appendectomy or antibiotics for your appendicitis. Please tell us how much you agree with the following statements," and responded "Strongly agree," "Agree," "Neither agree nor disagree," "Disagree," or "Strongly disagree" to the following 5 statements: It was the right decision; I regret the choice that was made; I would go for the same choice if I had to do it over again; The choice did me a lot of harm; and The decision was a wise one.[16] Responses were summed to form a regret score, with 0 corresponding to no regret and 100 corresponding to very high regret. We used a cutoff of 50 to differentiate participants with high regret (>50) and patients with low regret (≤ 50).

Dissatisfaction

Participants were asked, "How satisfied are you with your treatment for appendicitis?"; they were considered to be dissatisfied if they answered "Somewhat dissatisfied" or "Very dissatisfied," vs."Somewhat satisfied" or "Very satisfied."

Problems With Activities

We used a question from the EQ-5D questionnaire.[12] Although the primary CODA analysis used the full EQ-5D scale,[12] we chose to focus on a single question because we were interested in a measure of function. Participants were asked to indicate which statement best describes their health status on that day regarding "Usual Activities (eg work, study, housework, family or leisure activities)." Participants were categorized as having no problems with activities ("I have no problems with performing my usual activities") or problems with activities ("I have some problems with performing my usual activities" or "I am unable to perform my usual activities").

Prolonged Missed Work

Those who indicated they were currently employed (working for pay) were asked "Did you miss any days of work since you initially went to the emergency room for signs and symptoms of your appendicitis (include day of presentation)?" and if yes, how many days. Participants who missed more than 10 days of work (colloquially, 2 weeks of work) were considered to have prolonged missed work. Participants missing employment status on all 3 surveys were excluded. This measure has been reported for the CODA trial,[10] but the analysis presented here differs in several ways. In the analysis presented here, we used a shorter recall window (30 days vs. 90 days); we focused on the subgroup of people who were traditionally employed for pay (vs. including responses from all participants); we dichotomized the amount of missed work (vs. examining the number of days of missed work, which has a skewed distribution); and, last, we used multiple imputation to address missing data (vs. a complete case analysis).

Baseline Characteristics

Baseline factors included demographic, sociodemographic, clinical, and radiologic variables, details of which have been reported.[10,15] Preferred language for study materials (Spanish or English) was reported by the participant at the time of screening, and this determined in which language surveys were given. Concern about bills was determined by the question "If you were admitted or readmitted to the hospital for your appendicitis, would you be worried about the bills you would have to pay out of pocket?" Help with health literacy was split into two groups (always/often/sometimes or rarely/never) based on the answer to "How often do you need to have someone help you when you read instructions, pamphlets, or other written material from your doctor or pharmacy?" Household income below the poverty line and/or Medicaid beneficiary (yes or no) was determined based on the reported household income accounting for number of dependents and whether participants reported using Medicaid or another state program for insurance. Level of education was split into two groups: high school or General Educational Development test credential (GED) or less, or some education beyond high school GED. Employment status at baseline was based on the survey answers and supplemented with electronic medical record data; participants were asked "If employed, how often does your job require a lot of physical activity during your work shift?" Answers about employment and physical activity were combined into the variable 'Employment and activity,' with 3 levels: employed and active (job requires physical activity most or all of the time), employed and inactive (job requires physical activity some of the time or less), and unemployed/student/other. The number of dependents (any or none) was determined by the reported "number of family members and friends who depend on your help on a regular basis," and whether participants were sharing responsibilities for those dependents was determined by the question "Is there at least one other adult sharing responsibilities of caring for family members and/or friends who depend on you for help on a regular basis?" (yes, no, or no dependents).

The CODA dataset for this analysis includes all information as of May 21, 2021.

Statistical Methods

Baseline characteristics of participants were described using means (standard deviations) for continuous measures and n (%) for categorical variables.

When comparing outcomes between treatment arms we used an intention-to-treat framework and focused on antibiotics-assigned versus appendectomy-assigned participants (RCT cohort). We then described outcomes among the antibiotics-assigned participants who did and did not undergo appendectomy within 30 days. To address concerns about whether treatment comparisons hold in subgroups of interest, we performed a posthoc subgroup analysis among those whose incomes were below the poverty line and/or Medicaid beneficiaries. Throughout, formal comparisons were made using odds ratios (ORs) from univariate logistic regression; confidence intervals (CIs) around rates were calculated using a log-binomial framework.

To identify baseline factors potentially associated with worse outcomes, we pooled data from the RCT and observational cohorts. We used univariate logistic regression to assess the association between the odds of negative outcomes and various baseline characteristics. Because we believe the mechanism driving potential associations may differ by treatment

TABLE 1. Demographic and Clinical Characteristics* of Participants in the Randomized and Observational Cohorts of CODA (n = 2062)

Baseline Factor		Mean (SD) or n (%)
Cohort (%)	RCT	1552 (75.3%)
	Observational	510 (24.7%)
Treatment arm (%)	Appendectomy	1029 (49.9%)
	Antibiotics	1033 (50.1%)
Age over 50 years (%)	No	1666 (80.8%)
	Yes	396 (19.2%)
Sex (%)	Male	1268 (61.5%)
	Female	794 (38.5%)
BMI (%)	Less than 25	545 (31.3%)
	At least 25 but less than 30	609 (35.0%)
	At least 30 but less than 35	341 (19.6%)
	35 and above	246 (14.1%)
Preferred language (%)	English	1508 (73.1%)
	Spanish	554 (26.9%)
Health literacy help (%)	Never or rarely	1609 (81.4%)
	Sometimes or more	367 (18.6%)
Education (%)	Some beyond HS/GED	1252 (61.8%)
	HS/GED or less	773 (38.2%)
Employment and activity (%)	Employed (inactive)	768 (38.3%)
	Employed (active)	661 (33.0%)
	Unemployed/student/other	574 (28.7%)
Below poverty or Medicaid/state program (%)	No	883 (55.3%)
	Yes	714 (44.7%)
Hispanic (%)	No	1181 (57.3%)
	Yes	880 (42.7%)
Appendiceal diameter (mean [SD])		11.36 (2.87)
Perforation/abscess/fat** (%)	No	1701 (87.1%)
	Yes	253 (12.9%)
Appendicolith (%)	No	1494 (72.5%)
	Yes	568 (27.5%)

*The following variables are addressed in supplemental tables, http://links.lww.com/SLA/D819: worried about bills (yes vs. no), any dependents (yes vs. no), sharing responsibilities for dependents (no, yes, or none), instrumental support score, race (White, Black, American Indian, Asian, Islander, Other, or Multiple), smoking history (yes vs. no), duration of symptoms (1 or more days vs. less than 1 day), average pain in the previous 7 days, discharge pain controlled (no, yes, or no pain), white blood cell count, fever (yes vs. no), nausea/vomiting/anorexia (yes vs. no), Charlson score (above 0 vs. 0), and Alvarado score.

**Perforation/abscess/fat: perforation, abscess, or moderate or severe periappendiceal fat stranding (phlegmon).

Note: Mean and standard deviation (SD) shown for continuous variables, count (n) and percent for categorical variables; see Supplemental Table 1, http://links.lww.com/SLA/D819 for missing data; BMI, body mass index; GED, General Educational Development test credential; HS, high school; RCT, randomized clinical trial.

group, we stratified by treatment and estimated separate associations within the surgery group and the antibiotics group.

Information on regret, satisfaction, and usual activities were missing for 17%, 15%, and 13% of participants, respectively. Among participants whose employment status could be positively confirmed (893 participants in the RCT and 230 in the observational cohort), data on missing work was missing in 12% of participants. Missing data for both outcomes and baseline factors were imputed using multivariate imputation by chained equations algorithms in R software; see Appendix A for details.[17]

Analyses were performed in R statistical software version 40.2.

Percent with negative outcome (95%CI)

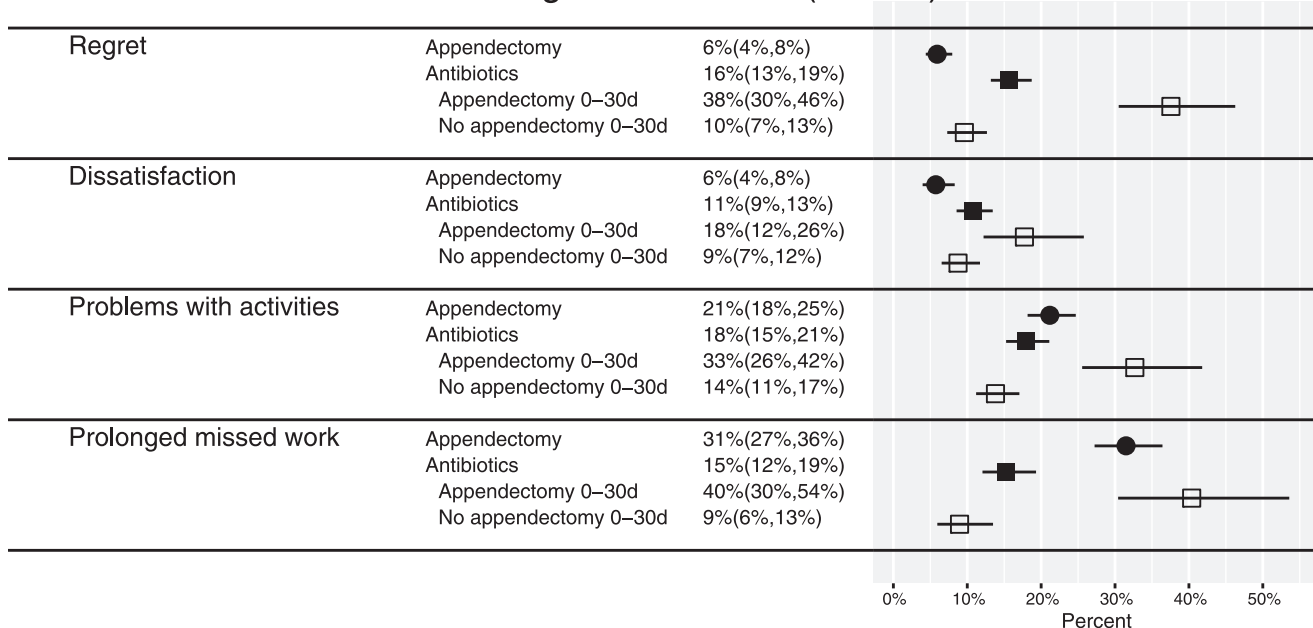|		
Regret	Appendectomy	6%(4%,8%)
	Antibiotics	16%(13%,19%)
	Appendectomy 0–30d	38%(30%,46%)
	No appendectomy 0–30d	10%(7%,13%)
Dissatisfaction	Appendectomy	6%(4%,8%)
	Antibiotics	11%(9%,13%)
	Appendectomy 0–30d	18%(12%,26%)
	No appendectomy 0–30d	9%(7%,12%)
Problems with activities	Appendectomy	21%(18%,25%)
	Antibiotics	18%(15%,21%)
	Appendectomy 0–30d	33%(26%,42%)
	No appendectomy 0–30d	14%(11%,17%)
Prolonged missed work	Appendectomy	31%(27%,36%)
	Antibiotics	15%(12%,19%)
	Appendectomy 0–30d	40%(30%,54%)
	No appendectomy 0–30d	9%(6%,13%)

FIGURE 1. Percent of participants in the randomized cohort with negative perceptions of treatment and impaired function, by treatment arm and 30-day appendectomy status for the antibiotics arm; percentages and 95% confidence intervals are pooled estimates from multiply imputed data sets.

RESULTS

Among 2062 participants (average age 37.5 years, 38.5% female), 1552 (75%) were from the RCT cohort (Table 1). There were differences in sociodemographic characteristics between the RCT and observational cohorts (Supplemental Table 1, http://links.lww.com/SLA/D819). The characteristics of the randomized cohort have been previously reported.[10]

Overall, decisional regret and dissatisfaction were low (10%, 95% CI 9%–12% and 8%, 95% CI 6%–10%, respectively). Problems with usual activities (20%, 95% CI 18%–22%) and prolonged missed work (22%, 95% CI 19%–24%) occurred more commonly (Supplemental Table 2, http://links.lww.com/SLA/D819 contains details of each outcome in the RCT and observational cohorts).

Comparison of PROs Between Treatment Arms (RCT)

Although negative perceptions of treatment (decisional regret and dissatisfaction) were uncommon in both treatement arms (Fig. 1), participants assigned to antibiotics more often had negative perceptions of treatment than those assigned to appendectomy; the OR for dissatisfaction was 1.98 (95% CI 1.25–3.12) and the OR for treatment regret was 2.97 (95% CI 2.05–4.31), comparing antibiotics with appendectomy.

Conversely, those assigned to appendectomy more often reported limitations in function compared with those in the antibiotics group; missing more than 10 days of work in the antibiotics-assigned group was less common (OR 0.39, 95% CI 0.27–0.56). Differences in the problems with usual activities were somewhat smaller (OR 0.81, 95% CI 0.62–1.07 for antibiotics versus appendectomy). In the group that was assigned to antibiotics and then underwent an appendectomy, negative perceptions of treatment and impact on function were much more common compared with all other groups (Fig. 1). Similar findings were noted in the observational cohort (Supplemental Table 2, http://links.lww.com/SLA/D819).

When compared with those assigned to appendectomy, those assigned to antibiotics had higher odds of dissatisfaction and decisional regret. These associations persisted even when focusing only on the subgroup who received antibiotics and did not end up undergoing appendectomy (Figs. 1 and 2). Among those in the antibiotic group, there were trends of lower odds for both impairment of activity and prolonged missed work compared with the appendectomy-assigned group, although this did not hold in the subgroup that was assigned to antibiotics and underwent appendectomy.

Baseline Factors Associated With Outcome in the Antibiotics Group-pooled Cohort

Among those in the antibiotic group, women, those with appendicolith (for regret), and wider diameter appendices (for regret) appeared to more often have decisional regret or dissatisfaction than those without those characteristics (Table 2 and Supplemental Table 3A, http://links.lww.com/SLA/D819). There was no positive association between some sociodemographic characteristics (ie, education, instrumental support) and regret or dissatisfaction. Other sociodemographic characteristics had notable ORs with consistent trends across the two measures of treatment perception, but wide CI. The analysis of sociodemographic characteristics and satisfaction and regret was limited by the relatively uncommon outcomes and small subgroup sizes.

Among those undergoing antibiotic treatment, many sociodemographic and clinical variables were associated with the 2 measures of function, including age and appendicolith (Table 2 and Supplemental Table 3A, http://links.lww.com/SLA/D819). Some of these associations have CIs that excluded 1 for only one of the measures of function but followed the same trend in the other with wider CIs (eg, duration of symptoms); others were directionally dissimilar in the two measures of function (eg, preferred language).

OR for negative outcome (95% CI)

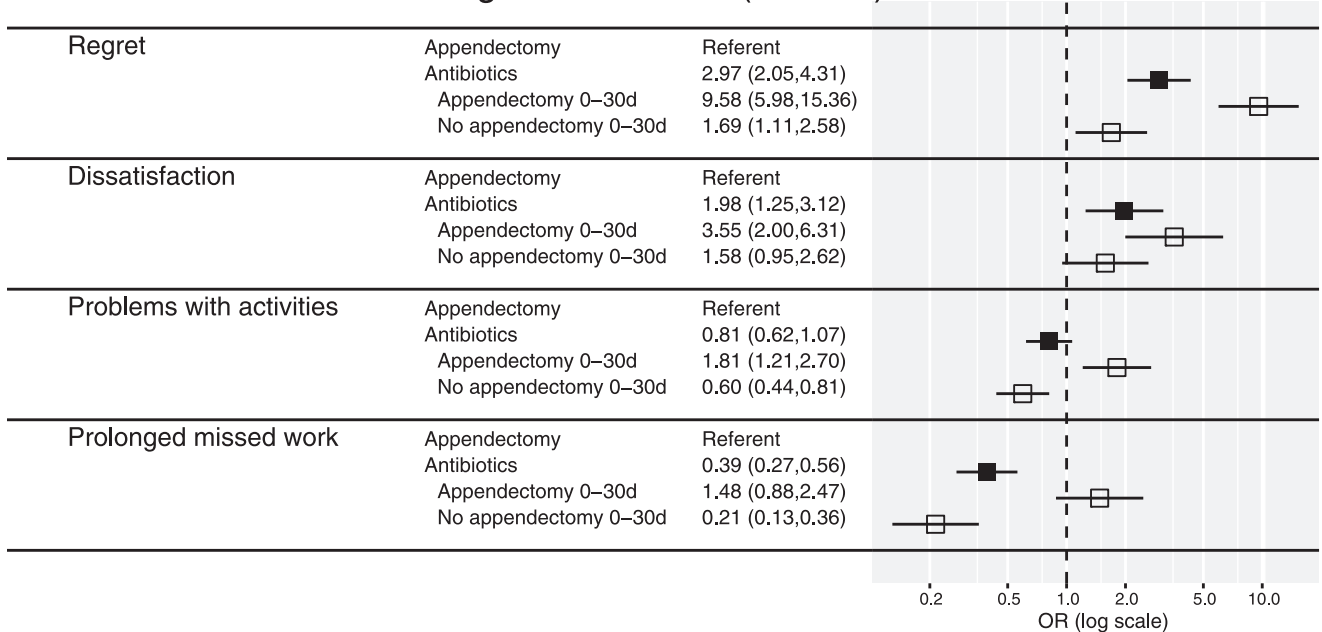|		
Regret	Appendectomy	Referent
	Antibiotics	2.97 (2.05,4.31)
	Appendectomy 0–30d	9.58 (5.98,15.36)
	No appendectomy 0–30d	1.69 (1.11,2.58)
Dissatisfaction	Appendectomy	Referent
	Antibiotics	1.98 (1.25,3.12)
	Appendectomy 0–30d	3.55 (2.00,6.31)
	No appendectomy 0–30d	1.58 (0.95,2.62)
Problems with activities	Appendectomy	Referent
	Antibiotics	0.81 (0.62,1.07)
	Appendectomy 0–30d	1.81 (1.21,2.70)
	No appendectomy 0–30d	0.60 (0.44,0.81)
Prolonged missed work	Appendectomy	Referent
	Antibiotics	0.39 (0.27,0.56)
	Appendectomy 0–30d	1.48 (0.88,2.47)
	No appendectomy 0–30d	0.21 (0.13,0.36)

OR (log scale): 0.2 0.5 1.0 2.0 5.0 10.0

FIGURE 2. Odds ratios for negative perceptions of treatment and impaired function, comparing antibiotics-assigned (overall and by 30-day appendectomy status) to appendectomy-assigned in the randomized cohort; ORs and 95% confidence intervals are pooled estimates from multiply imputed data sets.

Baseline Factors Associated With Outcome in the Appendectomy Group-pooled Cohort

Among those undergoing appendectomy, no factors were associated with dissatisfaction. For regret after appendectomy, there were greater odds of regret for Spanish speakers, those who reported being physically active at work, and Black participants (Table 3 and Supplemental Table 3B, http://links.lww.com/SLA/D819). A number of other factors had notable ORs and consistent associations with regret and dissatisfaction but wide CIs. These included older age and sharing responsibility for dependents.

Many sociodemographic variables were associated with problems with activities or longer missed work, including lower

TABLE 2. Univariate Odds Ratios and 95% Confidence Intervals for Association Between Baseline Factors and Worse Outcomes in the Antibiotics-assigned/Selected Arm in the Combined Cohort

Baseline Factor	Comparison	Regret	Dissatisfaction	Problems with Activities	Prolonged Missed Work
Age	Over 50 years vs. not	1.35 (0.88,2.06)	1.20 (0.73,1.99)	**1.78 (1.20,2.66)**	**2.00 (1.09,3.65)**
Sex	Female vs. male	**1.50 (1.05,2.14)**	**1.93 (1.20,3.11)**	1.28 (0.90,1.83)	1.34 (0.78,2.28)
BMI	(25, 30) vs. <25	1.08 (0.67,1.76)	0.99 (0.55,1.78)	1.24 (0.80,1.93)	1.59 (0.82,3.09)
	(30, 35) vs. <25	1.16 (0.69,1.97)	0.62 (0.31,1.25)	1.27 (0.76,2.15)	*2.22 (1.11,4.44)*
	≥35 vs. <25	1.47 (0.83,2.60)	0.95 (0.46,1.93)	**1.88 (1.10,3.21)**	0.73 (0.24,2.24)
Preferred language	Spanish vs. English	0.98 (0.63,1.54)	1.06 (0.64,1.74)	0.92 (0.62,1.37)	*3.15 (1.83,5.43)*
Health literacy help	Sometimes or more vs. never or rarely	1.29 (0.82,2.04)	0.78 (0.41,1.48)	1.54 (1.00,2.37)	1.34 (0.67,2.68)
Education	HS/GED or less vs. some beyond HS	0.91 (0.61,1.36)	0.68 (0.41,1.12)	0.85 (0.60,1.20)	*2.40 (1.42,4.06)*
Employment and activity	Employed (active) vs. employed (inactive)	0.83 (0.54,1.29)	1.15 (0.65,2.02)	1.23 (0.81,1.86)	1.50 (0.86,2.62)
	Unemployed/student/other vs. employed (inactive)	1.05 (0.68,1.61)	1.23 (0.74,2.06)	**1.60 (1.07,2.40)**	1.57 (0.62,3.99)
Below poverty or Medicaid/state program	Yes vs. no	0.99 (0.63,1.55)	0.90 (0.51,1.57)	1.33 (0.94,1.89)	*2.77 (1.62,4.72)*
Hispanic	Yes vs. no	0.87 (0.59,1.29)	0.85 (0.54,1.34)	0.83 (0.58,1.17)	*2.27 (1.35,3.81)*
Appendiceal diameter	For increase of 1mm	**1.07 (1.01,1.14)**	1.02 (0.95,1.10)	1.06 (1.00,1.12)	**1.16 (1.07,1.26)**
Perforation/abscess/fat	Yes vs. no	1.00 (0.58,1.72)	1.16 (0.59,2.28)	1.01 (0.61,1.65)	1.78 (0.89,3.57)
Appendicolith	Yes vs. no	**1.91 (1.31,2.77)**	1.45 (0.89,2.38)	**1.52 (1.01,2.29)**	**1.92 (1.13,3.27)**

Odds ratios summarize association between baseline factors and odds of regret (decisional regret score over 50), dissatisfaction (somewhat or very dissatisfied with treatment for appendicitis), problems with activities (having problems performing usual activities, eg, work, housework, leisure), and prolonged missed work (missed more than 10 days of work, among those who were employed) at 30 days; bold text indicates that the confidence interval excludes 1, italic text indicates that the odds ratio is large (>2) or small (<0.5), considering a 1-standard deviation increase for continuous baseline factors; missing data in both the baseline factors and outcomes was addressed via multiple imputation; BMI, body mass index; GED, General Educational Development test credential; HS, high school.

TABLE 3. Univariate Odds Ratios and 95% Confidence Intervals for Association Between Baseline Factors and Worse Outcomes in the Appendectomy-assigned/Selected Arm in the Combined Cohort

Baseline Factor	Comparison	Regret	Dissatisfaction	Problems with Activities	Prolonged Missed Work
Age	Over 50 years vs. not	0.59 (0.23,1.46)	0.78 (0.30,2.00)	1.26 (0.86,1.85)	0.76 (0.43,1.37)
Sex	Female vs. male	0.75 (0.40,1.40)	1.31 (0.64,2.67)	0.99 (0.70,1.39)	0.78 (0.52,1.19)
BMI	[25, 30) vs. <25	1.01 (0.50,2.06)	1.04 (0.53,2.03)	1.01 (0.69,1.49)	1.37 (0.81,2.34)
	[30, 35) vs. <25	0.97 (0.41,2.27)	0.73 (0.27,2.00)	1.00 (0.62,1.60)	*2.17 (1.20,3.92)*
	≥35 vs. <25	1.00 (0.41,2.43)	0.51 (0.13,1.95)	1.21 (0.73,1.98)	1.11 (0.57,2.18)
Preferred language	Spanish vs. English	*0.43 (0.20,0.95)*	0.57 (0.24,1.35)	1.04 (0.73,1.49)	*6.49 (4.12,10.24)*
Health literacy help	Sometimes or more vs. never or rarely	1.07 (0.48,2.39)	1.03 (0.45,2.39)	**1.56 (1.03,2.37)**	*3.68 (2.23,6.07)*
Education	HS/GED or less vs. some beyond HS	1.39 (0.76,2.53)	0.86 (0.41,1.80)	**1.41 (1.02,1.96)**	*4.47 (2.86,6.99)*
Employment and activity	Employed (active) vs. employed (inactive)	***2.85 (1.25,6.49)***	1.92 (0.79,4.65)	1.28 (0.88,1.86)	***2.03 (1.34,3.07)***
	Unemployed/student/other vs. employed (inactive)	*2.06 (0.78,5.43)*	1.59 (0.68,3.73)	1.24 (0.85,1.81)	1.42 (0.66,3.07)
Below poverty or Medicaid/state program	Yes vs. no	1.25 (0.63,2.45)	1.12 (0.54,2.32)	1.43 (0.95,2.15)	***3.79 (2.49,5.76)***
Hispanic	Yes vs. no	0.69 (0.37,1.27)	0.69 (0.36,1.32)	0.88 (0.63,1.22)	*4.34 (2.68,7.03)*
Appendiceal diameter	For increase of 1mm	1.01 (0.90,1.13)	0.90 (0.79,1.03)	1.03 (0.97,1.09)	0.99 (0.92,1.07)
Perforation/abscess/fat	Yes vs. no	1.00 (0.42,2.41)	0.95 (0.34,2.62)	1.10 (0.67,1.80)	1.02 (0.57,1.83)
Appendicolith	Yes vs. no	1.10 (0.58,2.06)	1.04 (0.49,2.22)	1.28 (0.86,1.91)	**1.73 (1.14,2.65)**

Odds ratios summarize association between baseline factors and odds of: regret (decisional regret score over 50), dissatisfaction (somewhat or very dissatisfied with treatment for appendicitis), problems with activities (having problems performing usual activities, eg, work, housework, leisure), and prolonged missed work (missed more than 10 days of work, among those who were employed) at 30 days; bold text indicates that the confidence interval excludes 1, italic text indicates that the odds ratio is large (> 2) or small (< 0.5), considering a 1-standard deviation increase for continuous baseline factors; missing data in both the baseline factors and outcomes was addressed via multiple imputation; BMI, body mass index; GED, General Educational Development test credential; HS, high school.

health literacy and less education. Some clinical variables were also associated with these outcomes, such as higher average pain before hospitalization (Table 3 and Supplemental Table 3B, http://links.lww.com/SLA/D819).

Posthoc Analysis

Given that household income below the poverty line and/or Medicaid beneficiary status was associated with an increased odds of limitations in function and prolonged missed work *with both treatments*, and because this characteristic was identified in almost half of the participants in the RCT, we performed a post hoc intention-to-treat analysis comparing the two treatments in this subgroup. We found directionally similar outcomes among this subgroup compared with the entire randomized cohort. Among those with household income below the poverty line and/or Medicaid beneficiaries, the OR for problems with activities was 0.76 (95% CI 0.53–1.1) for antibiotics vs. appendectomy, and the OR for prolonged missed work was 0.32 (95% CI 0.20–0.51).

DISCUSSION

For both appendectomy and antibiotics, there were high levels of satisfaction and low levels of regret. Most participants had a return of ability to perform all usual activities by 30 days and missed less than 2 weeks of work. Though uncommon, dissatisfaction and regret were more often reported by those undergoing antibiotic treatment. This finding was not completely accounted for by the subgroup of antibiotic-assigned participants who ended up undergoing an appendectomy. Conversely, those receiving antibiotics were less likely to have > 2 weeks of missed work. This was not the case for the subgroup that underwent an appendectomy. Exploratory analysis revealed many clinical and demographic characteristics associated with outcomes.

In comparing the treatments for appendicitis there may not be a single measure that can be used to define success or failure for all people. For example, appendectomy after initial treatment with antibiotics is an outcome unique to that treatment strategy, although surgical site infection is an outcome only relevant to those who undergo appendectomy. To address this, CODA used a primary outcome—general health status at 30 days measured by the EQ-5D—as it was expected to be relevant to both treatment arms. Patient advisors and clinicians encouraged the measurement of other PROs such as those related to perception of treatment success (regret and satisfaction) and functional impact (problems with usual activities and > 2 weeks missed work) because these were expected to measure other aspects of care. These measures are increasingly reported in surgical research,[18,19] even though they may relate to both clinical outcomes and nonclinical issues (eg, parking, cafeteria services,[20] and surgeon relatability[21]). A caution for the use of satisfaction as an outcome in trials of treatments is that satisfaction may also be related to expectation. If there is asymmetry in expectation about success between treatment arms (ie, patients in 1 treatment group expect to have an operation whereas patients in the other expect to avoid an operation) then this may be a biased comparison. Alternatively, limitations on usual activities and prolonged time away from work offer decision-makers another perspective on treatment. The first outcome reflects that not all patients have equal physical abilities to perform usual activities at baseline and an inability to assesses for within-patient change. The second outcome—prolonged time away from work—may be an important measure for some and relevant to employers, but may not be relevant for the self-employed, caregivers, or those who work from home or in nontraditional worksites. A key aspect of patient-centered outcomes research is that it informs decision-makers based on the unique characteristics of both patients and providers and how they prioritize different outcomes. Because patients are unlikely to prioritize each of these outcomes equally, varied results across these PROs may be helpful in individualizing treatment decisions.

Findings from this study related to factors associated with these PROs should be considered hypothesis generating. For example, among those receiving antibiotics, female sex was

associated with dissatisfaction and regret. Although a biologic explanation for this finding is possible, there is no known association between female sex and severity of appendicitis,[22,23] and a post hoc analysis found that accounting for appendectomies did not impact the strength of the association between female sex and these outcomes. In general, female sex is not a risk factor for higher regret in healthcare decisions.[24] In RCTs for other conditions, adverse outcomes were associated with higher regret.[25,26] Although complications were more common in subgroups of patients who received antibiotics,[10] a post hoc sensitivity analysis (complete case RCT cohort only) found that complications did not explain the differences in these PROs between treatment groups.

In patients who underwent appendectomy, the exploratory analysis found several factors that appeared to be assocated with functional outcomes and time away from work. Findings may relate to residual pain or symptoms, but many surgeons still recommend lifting restrictions and delaying return to work/ school for 7 to 14 days after appendectomy. Apparent associations with specific sociodemographic characteristics such as Spanish as a preferred language, Hispanic ethnicity, requiring help with health literacy, and having a high school or GED education or less raise many potential hypotheses, including the possibility of communication issues affecting recovery and employment that requires more physical labor. Despite this, an exploratory subgroup re-analysis of the RCT cohort found that in people with household income below the poverty line and/or Medicaid beneficiaries, antibiotics were still favored over appendectomy for function and return to work outcomes.

This study should be considered in light of several limitations. The CODA trial[10,27] found noninferiority of the main analytic outcome (EQ-5D at 30 days) and evaluated a group of clinical and healthcare utilization outcomes, focused only on the randomized arms of the study. The exploratory analysis presented here also includes the patients who refused randomization and selected their treatment and focuses on outcomes not included in the primary reporting, including a focus on negative reactions related to function and perceptions of treatment success. Those who selected their treatment were included in the analysis of factors associated with these outcomes, but we recognize their inclusion may have introduced bias and this should be considered hypothesis generating. Missing data were addressed using multiple imputation, but this may have introduced bias. We replicated the comparison of treatment arms in the RCT cohort including only participants with complete information (a complete case analysis) and had similar findings. The PRO instruments included in this study have all been previously used by researchers but have not necessarily been validated for the treatment of appendictis or in the Spanish language. Key properties, such as minimal clinically important differences and severity cutpoints have not been established for the satisfaction and decisional regret measures, and we employed cutpoints that were not determined by psychometric testing. We did not assess whether descriptions of 'satisfaction' and 'regret' relate to expectations of the healthcare system or prior experiences, and this may differ by demographic subgroups. The question on decisional regret asked, "think about the decision to follow through with your initial treatment of appendectomy or antibiotics," but it is possible that some patients in the randomized cohort interpreted that to be related to the decision to be randomized. For assessments of associations of factors with outcomes, relatively infrequent events required us to focus on unadjusted associations, and for similar reasons we didn't examine the impact of potentially important rare factors like

surgical complications. The analysis of prolonged missed work was only applied to a subset of the participants who reported that they were employed/working for pay. We did not ask whether time away from work was related to paid sick leave nor did we ask about missed work for those who do unpaid labor (eg, stay-at-home parents, caregivers) or those who are self-employed. We focused on the baseline factors associated with these outcomes to help patients and their caregivers make decisions about treatment at the point of care, but excluded potentially important factors that were measured after the index visit. Furthermore, this analysis focused only on one time point; comparisons may not hold up at later time points. Although CODA included most patients with the type of appendicitis typically treated with appendectomy, some patients were excluded (eg, those with an abscess or severe phlegmon), and these results may not apply to patients with those characteristics. The implications of pooling data from the RCT and observational cohorts are unclear. Although we saw similar trends in these PROs between treatment groups in the observational and randomized cohorts, differences in the characteristics of these two cohorts did exist and may have influenced the findings. In an exploratory comparison of outcomes between the RCT and observational cohorts, CIs were wide and covered 1 (supplemental Table 4, http://links.lww.com/SLA/D819). Additionally, RCT participants chose to be randomized to one of the two treatments instead of choosing their treatment, which could impact perception of treatment success. Lastly, these results are exploratory and we did not correct for multiple testing.

In conclusions, for either appendectomy or antibiotics, secondary PROs related to feeling and function reveal high levels of satisfaction and low levels of regret; most participants had a complete return to usual activities by 30 days, and most missed less than 10 days of work in that time period. A small proportion in both treatment groups reported dissatisfaction and regret at 30 days; this was more common in the antibiotics group and not completely explained by appendectomies in that arm. Patients more likely to have dissatisfaction or regret after antibiotics were women, those with an appendicolith, and those with wider diameter appendices. A substantial proportion (though still a minority) in the RCT reported having problems with usual activities and missing more than 2 weeks of work after treatment for appendicitis. These outcomes appeared to be more common in the appendectomy group than in the antibiotics group. Factors associated with these outcomes in both treatment groups included a mix of sociodemographic and clinical factors. Given similar outcomes, patients with these characteristics should be given the opportunity to consider both antibiotics and appendectomy, and this information should be used to support more informed decision-making.

REFERENCES

1. Eriksson S, Granstrom L. Randomized controlled trial of appendicectomy versus antibiotic therapy for acute appendicitis. *Br J Surg*. 1995; 82:166–169.

2. Styrud J, Eriksson S, Nilsson I, et al. Appendectomy versus antibiotic treatment in acute appendicitis. A prospective multicenter randomized controlled trial. *World J Surg*. 2006;30:1033–1037.

3. Turhan AN, Kapan S, Kutukcu E, et al. Comparison of operative and non operative management of acute appendicitis. *Ulus Travma Acil Cerrahi Derg*. 2009;15:459–462.

4. Hansson J, Korner U, Khorram-Manesh A, et al. Randomized clinical trial of antibiotic therapy versus appendicectomy as primary treatment of acute appendicitis in unselected patients. *Br J Surg*. 2009;96:473–481.

5. Vons C, Barry C, Maitre S, et al. Amoxicillin plus clavulanic acid versus appendicectomy for treatment of acute uncomplicated appendicitis: an open-label, non-inferiority, randomised controlled trial. *Lancet*. 2011;377:1573–1579.

6. Salminen P, Paajanen H, Rautio T, et al. Antibiotic therapy vs appendectomy for treatment of uncomplicated acute appendicitis: the APPAC randomized clinical trial. *JAMA*. 2015;313:2340–2348.

7. Talan DA, Saltzman DJ, Mower WR, et al. Antibiotics-first versus surgery for appendicitis: a US pilot randomized controlled trial allowing outpatient antibiotic management. *Ann Emerg Med*. 2017;70:1–11.

8. O'Leary DP, Walsh SM, Bolger J, et al. A Randomised clinical trial evaluating the efficacy and quality of life of antibiotic only treatment of acute uncomplicated appendicitis: results of the COMMA trial. *Ann Surg*. 2021;274:240–247.

9. Ceresoli M, Pisano M, Allievi N, et al. Never put equipoise in appendix! Final results of ASAA (antibiotics vs. surgery for uncomplicated acute appendicitis in adults) randomized controlled trial. *Updates Surg*. 2019;71:381–387.

10. CODA Collaborative, Flum DR, Davidson GH, et al. A Randomized trial comparing antibiotics with appendectomy for appendicitis. *N Engl J Med*. 2020;383:1907–1919.

11. Sippola S, Haijanen J, Viinikainen L, et al. Quality of life and patient satisfaction at 7-year follow-up of antibiotic therapy vs appendectomy for uncomplicated acute appendicitis: a secondary analysis of a randomized clinical trial. *JAMA Surg*. 2020;155:283–289.

12. EuroQol GroupEuroQol–a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199–208.

13. Jacobs D. Antibiotics for appendicitis - proceed with caution. *N Engl J Med*. 2020;383:1985–1986.

14. Flum DR, Alfonso-Cristancho R, Devine EB, et al. Implementation of a "real-world" learning health care system: Washington State's Comparative Effectiveness Research Translation Network (CERTAIN). *Surgery*. 2014;155:860–866.

15. Davidson GH, Flum DR, Talan DA, et al. Comparison of Outcomes of antibiotic Drugs and Appendectomy (CODA) trial: a protocol for the pragmatic randomised study of appendicitis treatment. *BMJ Open*. 2017;7:e016117.

16. The Ottawa Hospital. Decision Regret Scale. Patient Decision Aids Web site. https://decisionaid.ohri.ca/eval_regret.html. Accessed September 9, 2021.

17. van Buuren S. Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1–67.

18. Wilson A, Ronnekleiv-Kelly SM, Pawlik TM. Regret in surgical decision making: a systematic review of patient and physician perspectives. *World J Surg*. 2017;41:1454–1465.

19. Lyu H, Wick EC, Housman M, et al. Patient satisfaction as a possible indicator of quality surgical care. *JAMA Surg*. 2013;148:362–367.

20. Oz MC, Zikria J, Mutrie C, et al. Patient evaluation of the hotel function of hospitals. *Heart Surg Forum*. 2001;4:166–171.

21. Lieser MJ, Watts DD, Cooper T, et al. Critical role of trauma and emergency surgery physicians in patient satisfaction: an analysis of consumer assessment of healthcare providers and systems, hospital version data from 186,779 patients and 168 hospitals in a national healthcare system. *J Am Coll Surg*. 2021;232:656–663.

22. Buckius MT, McGrath B, Monk J, et al. Changing epidemiology of acute appendicitis in the United States: study period 1993–2008. *J Surg Res*. 2012;175:185–190.

23. Augustin T, Cagir B, Vandermeer TJ. Characteristics of perforated appendicitis: effect of delay is confounded by age and gender. *J Gastrointest Surg*. 2011;15:1223–1231.

24. Becerra Perez MM, Menear M, Brehaut JC, et al. Extent and predictors of decision regret about health care decisions: a systematic review. *Med Decis Making*. 2016;36:777–790.

25. Crisp CC, Book NM, Smith AL, et al. Body image, regret, and satisfaction following colpocleisis. *Am J Obstet Gynecol*. 2013;209:473.

26. Lorenzo AJ, Pippi Salle JL, Zlateska B, et al. Decisional regret after distal hypospadias repair: single institution prospective analysis of factors associated with subsequent parental remorse or distress. *J Urol*. 2014;191 (5 Suppl):1558–1563.

27. CODA Collaborative, Davidson GH, Flum DR, et al. Antibiotics versus appendectomy for acute appendicitis - longer-term outcomes. *N Engl J Med*. 2021;385:2395–2397.