Antibiotics vs Appendectomy for Uncomplicated Acute Appendicitis

Lillian S Kao, MD, FACS, Darrell Boone, MD, FACS, Rodney J Mason, MD, FACS, for Members of the Evidence-Based Reviews in Surgery Group

The term *evidence-based medicine* was first coined by Sackett and colleagues[1] as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients." The key to practicing evidence-based medicine is applying the best current knowledge to decisions in individual patients. Medical knowledge is continually and rapidly expanding, and it is impossible for an individual clinician to read all the medical literature. For clinicians to practice evidence-based medicine, they must have the skills to read and interpret the medical literature so that they can determine the validity, reliability, credibility, and utility of individual articles. These skills are known as critical appraisal skills. Generally, critical appraisal requires that the clinician have some knowledge of biostatistics, clinical epidemiology, decision analysis, and economics as well as clinical knowledge.

The Canadian Association of General Surgeons (CAGS) and the American College of Surgeons (ACS) jointly sponsor a program titled, "Evidence-Based Reviews in Surgery" (EBRS), supported by an educational grant from Ethicon Inc and Ethicon Endo Surgery Inc. The primary objective of this initiative is to help practicing surgeons improve their critical appraisal skills. During the academic year, 8 clinical articles are chosen for review and discussion. They are selected not only for their clinical relevance to general surgeons, but also because they cover a spectrum of issues important to surgeons; for example, causation or risk factors for disease, natural history or prognosis of disease, how to quantify disease (measurement issues), diagnostic tests and the diagnosis of disease, and the effectiveness of treatment. Both methodologic and clinical reviews of the article are performed by experts in the relevant areas and posted on the EBRS website. A listserve discussion is held where participants can discuss the monthly article. Fellows and candidates of the College can access Evidence-Based Reviews in Surgery through the American College of Surgeons website (www.facs.org). All journal articles and reviews are available electronically through the website. Currently we have a library of 50 articles and reviews, which can be accessed at any time. Each October, a new set of articles will be available each month until May. Surgeons who participate in the current (modules) packages can receive CME credits by completing a series of multiple choice questions. Additional information about EBRS is on the ACS website or by email to the administrator, Marg McKenzie at mmckenzie@mtsinai.on.ca.

In addition to making the reviews available through the ACS and CAGS websites, 4 of the reviews are published in condensed versions in *the Canadian Journal of Surgery*, 4 in the *Journal of the American College of Surgeons*, and 4 in *Diseases of Colon and Rectum* each year.

REFERENCE

1. Evidence-Based Medicine Working Group. Evidence-based medicine. JAMA 1992;268:2420–2425.

SELECTED ARTICLES

Use of antibiotics alone for treatment of uncomplicated acute appendicitis: a systematic review and meta-analysis

Liu K, Fogg L. Surgery 2011;150:673–683.

Surgery versus conservative antibiotic treatment in acute appendicitis: a systematic review and meta-analysis of randomized controlled trials

Ansaloni L, Catena F, Coccolini F, et al. Dig Surg 2011;28:210–221.

Objective: To compare the effectiveness and safety of antibiotics and appendectomy in the treatment of uncomplicated acute appendicitis.

Data Sources: Medline, Cochrane Controlled Trials Register, Cochrane Library, and Embase.

Study Selection: Studies that compared treatment with antibiotics alone to appendectomy in patients with a diagnosis of acute appendicitis without known abscess. The study by Liu and Fogg included both retrospective and prospective nonrandomized and randomized studies. Only randomized controlled trials were included

Published by Elsevier Inc.

ISSN 1072-7515/12/$36.00
http://dx.doi.org/10.1016/j.jamcollsurg.2012.12.009

in the meta-analysis performed by Ansaloni and associates.

Outcome Measures: Liu and Fogg reported antibiotic failure (defined as lack of improvement or clinical progression within 24 to 48 hours of antibiotic initiation) and recurrent appendicitis (defined as diagnosis of appendicitis up to 1 year postdischarge) treatment complications including and excluding the need for appendectomy and need for antibiotics. No outcomes for the appendectomy group were reported. Ansaloni and associates defined antibiotic efficacy as definitive improvement without need for surgery at a median of 1-year follow-up, and surgical efficacy as confirmation of acute appendicitis or other appropriate indication for surgery at the time of appendectomy.

Results: Liu and Fogg reported that the mean antibiotic failure rate was 6.9% (range 0% to 11.8%) and the mean recurrent appendicitis rate was 14.2% (range 5.3% to 35%). A mean of 7.3% of patients (range 3.2% to 10%) had a normal appendix at appendectomy. In the analysis by Ansaloni and coworkers, the odds ratio for treatment efficacy of surgery over antibiotics was 6.01 (95% CI 4.27 to 8.47) and the odds ratio for complications was 1.92 (95% CI 1.30 to 2.85).

Conclusions: Liu and Fogg concluded that antibiotic treatment may fail and there is a risk of recurrence, but surgically treated patients are subjected to the risks of operative morbidity and mortality. Antibiotic therapy incurs significantly fewer complications. Ansaloni and associates concluded that although a nonsurgical approach in acute appendicitis can reduce the complications rate, the lower efficacy prevents antibiotic treatment from being a viable alternative to surgery.

Commentary: Appendectomy has not been challenged until recently as the treatment of choice for acute appendicitis, in part because of age-old concerns of progression to perforation and the resultant complications.[1,2] However, multiple case series from remote or resource-poor areas have suggested that antibiotic therapy can be successful in resolving acute appendicitis.[3-7] On the other hand, advances in diagnostic capability, antibiotic therapy, and surgical technique have resulted in laparoscopic appendectomy for uncomplicated acute appendicitis safely being offered as day surgery in certain settings.[8]

In 1995, Eriksson and Granstrom[9] published the first randomized controlled trial (RCT) of antibiotics vs appendectomy in 40 adult patients. Since then, 4 additional RCTs have been published,[10-13] 1 of which was subsequently retracted by the journal editors 2 years after publication.[14] In addition, there have been at least 6 systematic reviews and meta-analyses addressing this question.[15-20] Two of the meta-analyses argue for further high quality RCTs.[17,18] One meta-analysis argues that appendectomy is still the gold standard therapy for acute appendicitis.[15] The remaining meta-analyses vary in their recommendations from considering antibiotics as an option in selected patients only,[16] to offering antibiotics as an alternative to appendectomy when appropriate counseling of patients has occurred,[19] to using antibiotics as a primary treatment option for early acute appendicitis.[20] Methodologic differences between meta-analyses, along with differences in included studies, may account for the discrepancies in these recommendations.

Systematic reviews and meta-analyses are used to guide clinical decisions[21]and health care policy;[22] performance of a systematic review of the literature is the recommended first step in translating evidence into guidelines.[23] Systematic reviews are comprehensive summaries of the available evidence on a clinical topic that are performed using rigorous methodology; they require transparency and reproducibility in terms of the search strategy, inclusion and exclusion criteria, study evaluation, and synthesis of information. Meta-analyses are systematic reviews that quantitatively pool the data to obtain summary statistics. So, in deciding whether to apply the results of systematic reviews and meta-analyses to clinical care, several issues must be considered: 1. Internal validity—what are the sources and direction of bias? 2. Appropriateness of quantitatively pooling the data (for meta-analyses)—is there significant heterogeneity between studies? and 3. External validity of the results—can the results be applied to my patients? These issues are considered for 2 of the 6 meta-analyses addressing antibiotics vs appendectomy for uncomplicated acute appendicitis: the Liu and Fogg [17] and the Ansaloni and associates[18] meta-analyses.

What are the sources and direction of bias?

Sources of bias in a meta-analysis can result from publication bias and from inclusion of poor quality studies. Publication bias occurs when relevant published or unpublished studies are not identified and included in a systematic review or meta-analysis; a skewed representation of positive studies will result in an overestimation of treatment benefit. When there are only a few studies available for inclusion in a meta-analysis, the risk of publication bias increases. Furthermore, because positive studies are often more likely to be published, it is important to search more than just popular bibliographic databases. Strategies to ensure that all of the pertinent literature is reviewed include scanning the reference lists of retrieved articles, translating non-English studies, contacting content experts regarding knowledge of unpublished studies, checking for registered trials (ie, at clinicaltrials.gov), and searching the "grey literature." Research that is not published in traditional journals (ie, dissertations), abstracts from conference proceedings, and reports from governmental agencies are examples of grey literature.

Publication bias can first be assessed by reviewing the cited sources of studies. Liu and Fogg [17] searched only Medline; Ansaloni and coworkers[18] additionally searched the Cochrane Controlled Trials Register, Cochrane Library, and EMBASE. Liu and Fogg [17] excluded non-English trials; Ansaloni and colleagues[18] did not exclude studies based on publication status or language. Ansaloni and colleagues[18] also performed a manual search of reference lists, but neither Liu and Fogg [17] nor Ansaloni and associates[18] described methods for identifying the grey literature. Although Liu and Fogg[17] did not describe a comprehensive search strategy, all of the studies identified by Ansaloni and coworkers[18] are in the Liu and Fogg[17] meta-analysis.

Publication bias can also be assessed using tools such as the publication bias assessment. Liu and Fogg[17] did not formally assess for publication bias. Ansaloni and associates[18] used the publication bias assessment, described by Klein and coworkers,[24] which calculates the number of similar-sized unpublished studies that would be required to render the results of the meta-analysis not statistically significant at a p value of 0.05 (assuming that when averaged, the studies show no effect of the treatment). For the outcome of treatment efficacy, 86 additional trials would have been required, and for the outcome of complications, only 4 additional trials would have been necessary. The higher the number of trials, the less likely that publication bias has altered the results of the meta-analysis. So, publication bias is unlikely to have affected the results with regard to treatment efficacy but may have affected the assessment of complications.

A more commonly used method that the authors could have used to visually assess the risk of publication bias is a funnel plot: the horizontal axis represents the magnitude of the treatment effect (ie, the odds ratio, [OR]) and the vertical axis represents the weight of the study (ie, the sample size) or the precision of the estimate of treatment effect (ie, the standard error of the logarithm of the odds ratio). The resulting plot of studies should resemble a symmetrical inverted funnel such that the most precise study is at the top. Asymmetry in the funnel plot suggests publication bias, although there can be other explanations such as poor methodologic quality. In the Ansaloni and colleagues[18] analysis there appeared to be no publication bias for treatment efficacy, but there are only 4 trials included.

Regarding the methodologic quality of the included studies, the Liu and Fogg[17] analysis included 2 observational studies in addition to the 4 RCTs; the studies were assessed using the Newcastle-Ottawa Quality Assessment Scale for Cohort Studies.[25] This scale evaluates 8 characteristics of cohort studies relating to patient selection, comparability, and outcome. Liu and Fogg[17] considered the study quality to be good if 5 of 8 points were achieved; they stated that all included studies achieved at least 5 points. Despite these objective criteria, one should be cautious of unavoidable biases associated with observational studies. Randomized controlled trials are considered the gold standard for studies of therapeutic interventions because they minimize differences in baseline characteristics, including both known and unknown confounders, between treatment groups. Observational studies can only adjust for known confounders. So, observational studies tend to overestimate the treatment effect as compared with randomized trials.[26]

The Liu and Fogg[17] and Ansaloni and associates[18] meta-analyses included the same 4 RCTs. Only 2 trials reported adequate generation of the randomization sequence, 1 reported allocation concealment, none were blinded, and only 1 used an intention-to-treat analysis. One of the trials, the trial by Hansson and coworkers[12] not only used an inadequate method to generate the randomization sequence (birth dates), but also had just over a 50% crossover rate in the group assigned to antibiotic treatment; 52.5% of patients in the antibiotic group underwent surgery. And as previously mentioned, one of the trials was retracted.[14] So, overall, the trials had an unclear to high risk of bias.

The high crossover rate in the Hansson and colleagues[12] trial deserves additional comment. An intention-to-treat analysis is performed when patients are analyzed in the group to which they were randomized, regardless of whether or not the treatment was received. Intention-to-treat analysis is recommended for superiority trials because of preservation of the balance of baseline variables achieved by randomization and prevention of bias due to differential (or nonrandom) loss of patients.[27] However, intention-to-treat analysis may also underestimate the benefit of a therapy. Given that receipt of appendectomy was classified as a failure of antibiotic treatment by Hansson and associates,[12] one should ask whether this trial, which contributed 50% of the total patients in the Ansaloni and colleagues[18] meta-analysis (369 of 741), resulted in an underestimate of antibiotic efficacy for acute appendicitis. Methods for addressing this issue include performing a sensitivity analysis, ie, calculating treatment efficacy with and without the Hansson and colleagues[12] trial, and performing both per-protocol (as-treated) and intention-to-treat analyses. A subsequent meta-analysis by Mason and associates,[19] which included an additional RCT, used both approaches and found that the results for complications and treatment efficacy were not significantly changed.

Is there significant heterogeneity between studies?

In mathematically combining studies into a meta-analysis, one must assess how similar the studies are to each other. There may be clinical, methodologic, or statistical heterogeneity that preclude data synthesis. Clinical heterogeneity can be determined by examining the details of the studies. For example, the RCTs included in both the Liu and Fogg[17] and the Ansaloni and associates[18] analyses had differences between them in terms of patient populations (1 RCT included only males), the method of patient selection and in diagnosing acute appendicitis, the intervention or antibiotic regimens, and the comparison or type of operation performed (1 RCT included only 8 of 124 laparoscopic appendectomies). Methodologic heterogeneity was present in that in 1 trial, there was a high crossover rate, as already described.

Statistical heterogeneity can be determined by evaluating the forest plots of individual studies, which show the point estimates of treatment effect and the 95% CIs. Studies are homogenous if the point estimates are similar and the 95% CIs overlap. Visual examination of the forest plots in the Ansaloni and coworkers[18] analyses of treatment efficacy and complications shows no obvious heterogeneity. Alternatively, statistical tests such as the Cochran's Q or I^2 statistics can be performed, although meta-analyses are often underpowered to detect heterogeneity. If the Cochran's Q test is significant ($p < 0.05$), then the differences observed between studies are unlikely to be due to chance and there is heterogeneity present. In the Liu and Fogg[17] analysis, the test was nonsignificant, with a p value of 0.189 for the analysis of complications. In the Ansaloni and colleagues[18] analysis, no significant statistical heterogeneity was noted for treatment efficacy, perforated appendix, or complications.

Can the results be applied to my patients?

The external validity, or generalizability of the results to patients in the United States and Canada or to other settings, is limited. The patients were primarily from European countries. In addition, the Liu and Fogg[17] analysis excluded studies of only pediatric patients; the youngest age in the included RCTs was 18 years old. One of the studies excluded women and overall, there was a preponderance of men (ie, 70% [830 of 1,201] in the Liu and Fogg[17] analysis). Tinidazole was prescribed in half of the studies; however this agent is not commonly used in the US or Canada for treatment of intra-abdominal infections. The surgical approach (laparoscopic or open appendectomy) was not described in the majority of included studies, but open appendectomy appears to have been mainly used. In contrast, the majority of appendectomies in the US and Canada are performed laparoscopically, which is associated with fewer wound infections and shorter hospital stays.[28,29] Routine CT scan was not performed in most of the included studies, which may have resulted in patients with complicated appendicitis being included.

Ultimately, in deciding whether or not to use a therapy, risks and benefits as well as other important outcomes measures should be considered. Based on these meta-analyses, there is insufficient evidence on which to make a decision weighing benefits and burdens. Surgery has greater treatment efficacy than antibiotic treatment, but a higher risk of complications. However, treatment efficacy was variably defined. Complications were also not standardized across included studies, were not consistently reported, and were not attributed to the treatment groups in a similar manner. Furthermore, other important factors such as length of stay, costs, and patient satisfaction were assessed in only 1 of the included RCTs.

In conclusion, the evidence to date is not definitive enough to change routine practice. However, evidence-based medicine does not require Level I evidence, but merely that clinicians judiciously and conscientiously apply the best available evidence in making decisions about the care of individual patients.[30] So, current evidence provides support for the feasibility and safety of antibiotic therapy in patients with uncomplicated acute appendicitis. However, individual patient factors such as surgical risk, compliance with medical therapy, values placed on time to return to work or pain, anxiety about the risk of recurrence, and patient preferences should factor into the decision about which patients to offer antibiotic therapy alone. Further high quality RCTs are indicated, but need to have outcomes that are clinically relevant and patient-centered and to address unanswered questions such as which patients are most likely to benefit from (or fail) antibiotic therapy alone.

The Evidence-Based Reviews in Surgery Group Comprises:

Members of the EBRS Steering Committee

Nancy N Baxter, MD, FACS, Toronto, ON, Canada
Karen J Brasel, MD, FACS, Milwaukee, WI
Carl J Brown, MD, Vancouver, BC, Canada

REFERENCES

1. Fitz RH. Perforating inflammation of the vermiform appendix. Am J Med Sci 1886;92:321−346.
2. McBurney C. Experience with early operative interference in cases of disease of the vermiform appendix. NY Med J 1889; 50:676−684.
3. Adams ML. The medical management of acute appendicitis in a nonsurgical environment: a retrospective case review. Mil Med 1990;155:345−347.
4. Gurin NN, Slobodchuk Iu S, Gavrilov Iu F. [The efficacy of the conservative treatment of patients with acute appendicitis on board ships at sea]. Vestn Khir Im I I Grek 1992;148:144−150.
5. Bowers WF, Hughes CW, Bonilla KB. The treatment of acute appendicitis under suboptimal conditions. U S Armed Forces Med J 1958;9:1545−1557.
6. Foraker AG. A reluctant surgeon at sea. JAMA 1981;245: 2302−2303.
7. Campbell MR, Johnston SL 3rd, Marshburn T, et al. Nonoperative treatment of suspected appendicitis in remote medical care environments: implications for future spaceflight medical care. J Am Coll Surg 2004;198:822−830.
8. Alkhoury F, Malvezzi L, Knight CG, et al. Routine same-day discharge after acute or interval appendectomy in children: a prospective study. Arch Surg 2012;147:443−446.
9. Eriksson S, Granstrom L. Randomized controlled trial of appendicectomy versus antibiotic therapy for acute appendicitis. Br J Surg 1995;82:166−169.
10. 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.
11. Malik AA, Bari SU. Conservative management of acute appendicitis. J Gastrointest Surg 2009;13:966−970.
12. 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.
13. 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.
14. Retraction note to: Conservative management of acute appendicitis. J Gastrointest Surg 2011;15:2302.
15. Varadhan KK, Humes DJ, Neal KR, Lobo DN. Antibiotic therapy versus appendectomy for acute appendicitis: a meta-analysis. World J Surg 2010;34:199−209.
16. Wilms IM, de Hoog DE, de Visser DC, Janzing HM. Appendectomy versus antibiotic treatment for acute appendicitis. Cochrane Database Syst Rev 2011:CD008359.
17. Liu K, Fogg L. Use of antibiotics alone for treatment of uncomplicated acute appendicitis: a systematic review and meta-analysis. Surgery 2011;150:673−683.
18. Ansaloni L, Catena F, Coccolini F, et al. Surgery versus conservative antibiotic treatment in acute appendicitis: a systematic review and meta-analysis of randomized controlled trials. Dig Surg 2011;28:210−221.
19. Mason RJ, Moazzez A, Sohn H, Katkhouda N. Meta-analysis of randomized trials comparing antibiotic therapy with appendectomy for acute uncomplicated (no abscess or phlegmon) appendicitis. Surg Infect 2012;13:74−84.
20. Varadhan KK, Neal KR, Lobo DN. Safety and efficacy of antibiotics compared with appendicectomy for treatment of uncomplicated acute appendicitis: meta-analysis of randomised controlled trials. BMJ 2012;344:e2156.
21. Mulrow CD, Cook DJ, Davidoff F. Systematic reviews: critical links in the great chain of evidence. Ann Int Med 1997;126: 389−391.
22. Bero LA, Jadad AR. How consumers and policymakers can use systematic reviews for decision making. Ann Int Med 1997; 127:37−42.
23. Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. BMJ 2008;336:1049−1051.
24. Klein S, Simes J, Blackburn GL. Total parenteral nutrition and cancer clinical trials. Cancer 1986;58:1378−1386.
25. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed June 6, 2012.
26. Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.
27. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010;63:e1−37.
28. Hemmila MR, Birkmeyer NJ, Arbabi S, et al. Introduction to propensity scores: A case study on the comparative effectiveness of laparoscopic vs open appendectomy. Arch Surg 2010; 145:939−945.
29. Sauerland S, Jaschinski T, Neugebauer EA. Laparoscopic versus open surgery for suspected appendicitis. Cochrane Database Syst Rev 2010:CD001546.
30. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. BMJ 1996;312:71−72.