Surgery versus Conservative Antibiotic Treatment in Acute Appendicitis: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Luca Ansaloni a Fausto Catena b Federico Coccolini b Giorgio Ercolani b Filippo Gazzotti b Eddi Pasqualini b Antonio Daniele Pinna b

a Unit of General Surgery, Ospedali Riuniti, Bergamo, and b Unit of General, Emergency and Transplant Surgery, St. Orsola-Malpighi University Hospital, Bologna, Italy

Key Words
Acute appendicitis · Appendectomy · Antibiotic treatment · Systematic review · Meta-analysis

Abstract
Background/Aims: Although standard treatment typically consists of an early appendectomy, there has recently been an increase in the use of antibiotic therapy as primary treatment for acute appendicitis (AA). The aim of this analysis is to systematically evaluate the evidence available in relevant literature in order to compare the relative effectiveness of antibiotic therapy as a viable alternative to appendectomies in the treatment of AA. Methods: Literature was searched for randomized clinical trials (RCTs) comparing the efficacy of surgery versus antibiotic therapy. Differences in pooled odds ratios (OR) for outcomes within 95% confidence intervals (CI) were calculated. Results: Four RCTs were identified including 741 patients. Efficacy was significantly higher for surgery (OR = 6.01, 95% CI = 4.27–8.46). No differences were found in the numbers of perforated appendices (OR = 0.73, 95% CI = 0.29–1.84) and patients treated with antibiotics (OR = 0.04, 95% CI = 0.00–3.27). Complication rates were significantly higher for surgery (OR = 1.92, 95% CI = 1.30–2.85). Conclusion: Although a nonsurgical approach in AA can reduce the complications rate, the lower efficacy prevents antibiotic treatment from being a viable alternative to surgery. Since only a small number of RCTs of poor methodological quality are available, well-designed RCTs are needed for further investigation.

Introduction
More than 120 years have passed since the earliest appendectomies were first advocated as viable surgeries following the classic paper on 247 patients with perforated acute appendicitis (AA) published by Fitz [1] in 1886. Later, in 1889, McBurney [2] reported his study of 8 patients with AA, making a special reference to the use of early appendectomy. Since the time of these historic publications, there have been several reports of conservatively treating AA using antibiotics. In 1953, Harrison [3] reported 42 of 47 cases of AA being successfully treated using antibiotic therapy. Additionally, in 1959, Coldrey [4] reported 471 cases of AA treated conservatively with only 1 death, 9 patients requiring abscess drainage, and only 48 cases requiring a subsequent appendectomy; the idea was as controversial as it is today. Another report from China in 1977...
described 425 of 500 patients with suspected AA who were treated conservatively with either antibiotics or traditional Chinese medicine; upon follow-up examination, only 7% presented with symptoms of recurrence [5]. In both of these studies, patients were assessed by individual history and clinical examination. The treatment method differed without standardization and there was no consecutive follow-up. Moreover, conservative antibiotic treatment has also been described in the cases of 9 US submariners [6] and in 247 patients aboard Soviet ships at sea [7]. The latter is a review of conservative treatment outcomes in 252 patients with AA on vessels of the Kalingrad fishing industry from 1975 to 1987, where Gurin et al. [7] reported a recovery rate of 84.1% with only the use of antibiotics. The authors found no difference in outcomes based on presenting symptoms or age, but found that conservative treatment was most effective when administered within 12 h of symptom onset, ideally within the first 6 h. More recently, Kaminski et al. [8] reported a 5% recurrence rate within a median follow-up time of 4 years in 864 patients with AA treated solely with antibiotics. Furthermore, other researchers have reported a high success rate for the nonoperative medical treatment of AA as compared to the high complication rates of appendectomy performed in suboptimal conditions during situations of remote medical care [9–11].

Furthermore, the use of preoperative antibiotic treatment in AA can be implemented as a means of delaying an appendectomy, particularly during twilight hours, in which case the incidence of perforation, complications, and hospitalization in children operated within 6 h was equivalent to the rates of those who underwent surgery between 6 and 18 h after admission [12]. Even the conservative management of appendiceal mass has been advocated [13–15], although some recommend interval appendectomy [13] in the case of a cecal neoplasm or recurrent AA supervenes.

Generally speaking, many of these experiences did not receive much attention and, consequently, early appendectomy remained the standard treatment for AA in order to avoid perforations, even though population-based evaluations have indicated significant long-term risks following surgical intervention for AA [16]. It has been shown that small bowel obstruction requiring surgery occurs in 1.3% of individuals by 30 years and that the 30-day mortality is 0.24% with an increased standard mortality ratio [17, 18]. A negative appendectomy is particularly encumbered with problems. Using the standardized mortality ratio to describe the excess mortality in the operated patients compared with the general population in Sweden, Blomqvist et al. [18] found a 3.5-fold excess mortality after an appendectomy for nonperforated appendicitis and 6.5-fold excess mortality after perforated appendicitis. Surprisingly, the standardized mortality ratio after negative appendectomy with a discharge diagnosis of nonspecific abdominal pain as the only diagnosis that could motivate an exploration was increased 9.1-fold. This result is in accordance with the report by Flum and Koepsell [19] who found a 3-fold increase in mortality after negative appendectomy compared with appendectomy for appendicitis. Due to this excess mortality associated with appendectomy for nonperforated appendicitis and with negative appendectomy with a discharge diagnosis of nonspecific abdominal pain, suggesting that appendectomy is not a harmless operation, antibiotic treatment for AA has received increased attention in recent years [20, 21]. Several studies have indicated that even perforated AA in children can be effectively treated with antibiotics [22]. In addition, retrospective studies analyzing adults with conservatively treated perforated AA suggest that late recurrences were often very mild and treatable [23, 24].

More recently, four randomized controlled trials (RCTs) compared the respective efficacies of surgery and conservative antibiotic treatment in AA [25–28]. The aim of this systematic review (SR) and meta-analysis (MA) was to evaluate the available evidence, comparing the effectiveness and safety of appendectomy versus conservative antibiotic regimens in the treatment of AA.

**Methods**

**Literature Search and Selection of Randomized Trials**

Eligible RCTs were those comparing surgery (either laparoscopic or laparoscopic) with antibiotic therapies alone for the treatment of adult patients with AA, regardless of the publication status or language of the article. All other clinical investigations, such as observational studies, cases series, case-control studies, and non-RCTs, were excluded. A computerized bibliographic search from the year 1966 to August 15, 2009, was performed on MEDLINE, the Cochrane Controlled Trials Register, the Cochrane Library, and EMBASE databases using the following key words: ‘acute’, ‘appendicitis’, and ‘antibiotic’. This computer search was supplemented with a manual search of reference lists for all available review articles, primary studies, and books to find other studies not found during the computer search. Decisions regarding which trials to include were made independently by two reviewers (L. Ansaloni and F. Catena). Disagreements were resolved by discussion. Trial data, patient characteristics, technical methods, and treatment outcome (tables 1–3) were abstracted by two independent reviewers (L. Ansaloni and F. Catena) and discrepancies were resolved by discussion. All reports and analyses were in accordance with the PRISMA statement [29].
Table 1. Treatment methods of each RCT included in the meta-analysis

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surgery (open/ laparoscopic; antibiotics administered in the perioperative period)</strong></td>
<td>Not stated; antibiotics administered for a period of 24 h, but only in the event of bowel perforation or abdominal spillage</td>
<td>116/8; not stated</td>
<td>Not stated; antibiotics administered for 48 h, but only in the event of bowel perforation or abdominal spillage</td>
<td>Not stated; single-dose antibiotic prophylaxis and antibiotic therapy in the event of a gangrenous or perforated appendix</td>
</tr>
<tr>
<td><strong>Antibiotics</strong></td>
<td>2 g of cefotaxime every 12 h, 0.8 g of tinidazole administered intravenously every 24 h for a period of 2 days, 200 mg of ofloxacin twice a day, and 500 mg of tinidazole per os twice a day for a period of 8 days following discharge; i.v. fluids administered during first 48 h</td>
<td>2 g of cefotaxime every 12 h, 0.8 g of tinidazole administered intravenously every 24 h for a period of 2 days, 200 mg of ofloxacin twice a day, and 500 mg of tinidazole per os twice a day for a period of 10 days following discharge; i.v. fluids administered during first 24 h and eating allowed by the second day of hospitalization</td>
<td>500 mg of ciprofloxacin every 12 h, 500 mg of metronidazole administered intravenously every 8 h for a period of 2 days, 500 mg of ciprofloxacin twice a day, and 600 mg of tinidazole per os twice a day for 7 days following discharge; i.v. fluids administered during admission</td>
<td>1 g of cefotaxime twice a day, 1.5 g of metronidazole administered intravenously every 24 h (this i.v. therapy was sustained when the patient’s clinical condition did not show improvement), 500 mg of ciprofloxacin twice a day, and 400 mg of metronidazole per os three times a day for 10 days following discharge; i.v. fluids administered with no oral intake during admission</td>
</tr>
</tbody>
</table>

**Primary Outcome**

*Treatment Efficacy.* Treatment efficacy was considered as a primary end point. Given that the treatments being compared are so qualitatively different (surgery vs. conservative medical therapy), the definition of efficacy is also different for the two distinct treatment groups. Therefore, although this approach could be criticized for having natural biases and being inherently skewed [30], the definitions reported by Hansson et al. [28] were chosen for the purposes of this study. Efficacy for antibiotic treatment was defined as definitive improvement without requiring surgery within a median follow-up of 1 year. Efficacy for surgical treatment was defined as AA confirmed at the time of the surgical operation or another appropriate indication that the surgical procedure was required. This information was available for each of the four RCTs considered.

**Secondary Outcomes**

*Perforated Appendix.* In three studies [25–27] involving a total of 372 patients, the number of patients undergoing surgery to treat perforated AA was reported.

*Number of Patients Treated with Antibiotics.* This information was available in three studies [25, 27, 28], involving a total of 489 patients. In the surgical group, the patients treated with antibiotics include those who refused the operation as well as those who underwent surgery to address a gangrenous or perforated appendix. In the conservative treatment group, the patients not treated with antibiotics were those who preferred the appendectomy minus those who underwent operations with a gangrenous or perforated appendix.

*Complications.* Although this information was readily available in all four studies, only in a single article [28] were they clearly reported and subdivided into major and minor complications.

*Pain and Analgesic Consumption.* In two studies [25, 27] involving a total of 120 patients, pain evaluation was reported as the mean of the VAS (0–100 mm) score as registered by patients every 6 h during their hospital stay (first 42 postoperative hours) and by a surgeon during follow-up analyses. Even analgesic consumption, expressed as the mean of morphine dose in milligram [25] and of diclofenac sodium dose in milligram [27], was reported.

*Inflammatory Laboratory Tests and Body Temperature.* In the same two studies [25, 27], the mean concentration of C-reactive protein (CRP), the mean total white blood cell (WBC) count, and the mean body temperature were recorded during hospitalization and during the 30-day follow-up period.

*Hospital Stay, Sick Leave and Time Off of Work.* The patients’ hospital stay, expressed as the mean number of hospitalized days, was reported in all studies [25–28]. However, the average sick leave was reported by only two studies [26, 28] and the average recovery time spent ‘off of work’ was reported in only one study [26].

*Costs.* The total costs for the primary hospital stay (including hospital supplies, drugs, radiology, surgical resources, postoperative surveillance, laboratory tests, and pathology analyses) were only reported in the study of Hansson et al. [28], expressed as the mean of overall expenditures in Swedish krona (SEK).

*Patient Experience.* The patients’ individual experiences were evaluated in the study of Hansson et al. [28]; the mean number of days with abdominal pain after hospital discharge, the number of patients still experiencing symptoms 1 month after treatment, and the number of patients still experiencing abdominal pain 1 year after treatment were all recorded in this study.

*Trials Review and Data Extraction* The methodological integrity of the trials was assessed according to four major criteria [31–35]: adequately randomized sequences, appropriate concealment of treatment allocation, blinded outcome assessment, and intention-to-treat analysis. Each criterion was rated as ‘yes’, ‘unclear’, or ‘no’. The trial integrity was reported individually for each of the aforementioned criteria [32].
**Table 2.** Study characteristics of each RCT included in the meta-analysis

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of centers</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Inclusion criteria</td>
<td>Adult patients with typical AA history and clinical signs, positive findings following US, and either increased WBC and CRP values or elevated WBC and CRP levels as measured on two separate occasions within a 4-hour interval</td>
<td>Male patients 18–50 years of age admitted for suspected AA with a CRP level &gt;10 mg/l and having no clinical signs of perforation</td>
<td>Adult patients with typical AA history and clinical signs (Alvardo's score), positive findings following US, and either increased WBC and CRP values or elevated WBC and CRP levels as measured on two separate occasions within a 4-hour interval</td>
<td>Patients 18 years of age or older with a diagnosis of AA based on disease history, clinical signs, laboratory tests, and, in some cases, US, CT and gynecological examination</td>
</tr>
<tr>
<td>Exclusion criteria</td>
<td>NR</td>
<td>Patients with suspected appendix perforation, a CRP level &lt;10 mg/l, and an allergy to the protocol's antibiotics</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Planned follow-up</td>
<td>6, 10, and 30 days</td>
<td>1 week, 6 weeks, 1 year</td>
<td>7, 12, and 30 days</td>
<td>1 month, 1 year</td>
</tr>
<tr>
<td>Primary end points</td>
<td>NS (hospital stay, complications, pain, analgesic consumption, inflammatory laboratory tests, and body temperature)</td>
<td>NS (hospital stay, sick leave, time taken off of work, and complications)</td>
<td>NS (hospital stay, complications, pain, analgesic consumption, inflammatory laboratory tests, and body temperature)</td>
<td>Treatment efficacy and major complications</td>
</tr>
<tr>
<td>Secondary end points</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>Minor complications, duration of antibiotic therapy, abdominal pain after discharge from hospital, length of hospital stay and sick leave, and total in-hospital cost</td>
</tr>
<tr>
<td>Patients screened, n</td>
<td>45</td>
<td>Reported for only one treatment center</td>
<td>NR</td>
<td>369</td>
</tr>
<tr>
<td>Patients randomized, n (antibiotics/surgery)</td>
<td>40 (20/20)</td>
<td>252 (128/124)</td>
<td>80 (40/40)</td>
<td>369 (202/167)</td>
</tr>
<tr>
<td>Sample size calculation</td>
<td>NR</td>
<td>NR</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ethical committee approval</td>
<td>Yes</td>
<td>Yes</td>
<td>NR</td>
<td>Yes</td>
</tr>
<tr>
<td>Patients lost to follow-up evaluation, n</td>
<td>Antibiotic group: 1 (30 days)</td>
<td>NR</td>
<td>NR</td>
<td>Antibiotic group: 11 (1 month) and 16 (1 year); surgical group: 24 (1 month) and 22 (1 year)</td>
</tr>
<tr>
<td>Mean follow-up, months</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Publication status</td>
<td>Full paper</td>
<td>Full paper</td>
<td>Full paper</td>
<td>Full paper</td>
</tr>
<tr>
<td>Year of publication</td>
<td>1995</td>
<td>2006</td>
<td>2009</td>
<td>2009</td>
</tr>
</tbody>
</table>

NR = Not reported; CRP = C-reactive protein; WBC = white blood count; AA = acute appendicitis; US = ultrasonography; CT = computed tomography; NS = not stated.

**Statistical Methods**

Formal statistical tests for heterogeneity of the odds ratios (ORs) were performed using the Cochrane Q-test, assuming homogeneity with a $p$ value >0.05. Pooled ORs with 95% confidence intervals (CI) were calculated using either the fixed effects model (FEM) or random effects model (REM) as needed. The Peto method was used to calculate the odds [36]. A statistically significant result was assumed when the 95% CI of the pooled OR did not include 1. The publication bias assessment (PBA), calculated using the method of Klein et al. [37], expresses how many unpublished studies with negative or null results are needed to influence the results of the MA. The number needed to treat (NnT) was also calculated as 1/pooled risk difference. All procedures and calculations were made according to methods reported elsewhere [36].
Results

Computerized bibliographic searches retrieved 511 articles from the MEDLINE database, 77 articles from the Cochrane Controlled Trials Register, 75 articles from the Cochrane Reviews Library, and 880 articles from the EMBASE database. Among these articles, five studies were found comparing surgery to conservative antibiotic treatment in adult patients with AA. Only one trial was excluded [20] due to a lack of randomization. Four RCTs, all published as full papers, fulfilled the inclusion criteria and were included in the MA [25–28]. The four RCTs included a total of 741 patients, 351 treated with surgery and 390 treated with conservative antibiotic treatment.

The methodological quality of the RCTs included in the SR is rather poor according to the criteria considered previously (table 4). Three main items were variably considered in the four trials. The only item not respected by

| Table 3. Characteristics of patients included in each RCT (antibiotics/surgery) included in the meta-analysis |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Patients randomized, n | 20/20 | 128/124 | 40/40 | 202/167 |
| Mean age (range), years | 27.8 (18–53)/5.0 (19–75) | NR | 28.7 (17–56)/32.6 (18–64) | 38 (1)/38 (1) |
| Previous abdominal surgery | NR | NR | NR | 20/27 |
| Suspect of previous AA | NR | NR | NR | 18/10 |
| Mean duration of pain, h | 21.0 ± 14.7/18.4 ± 11.8 | NR | 23.0 ± 16.4/21.3 ± 14.3 | NR |
| Mean total WBC count (× 10^9/l) upon admission | 13.8 ± 4.4/13.9 ± 4.1 | 12.5 ± 3.8/12.4 ± 3.5 | 14.2 ± 4.9/14.7 ± 4.4 | 12.7 (0.3)/13.6 (0.3)* |
| Mean CRP concentration, mg/l, upon admission | 41 ± 30/40 ± 38 | 55 ± 44/54 ± 49 | 43 ± 29/42 ± 34 | 55 (4)/54 (4) |
| Mean body temperature upon admission, °C | 37.2 ± 0.7/37.1 ± 0.7 | 37.5 ± 0.7/37.4 ± 0.8 | 37.4 ± 0.6/37.6 ± 0.7 | 37.3 (0.1)/37.5 (0.1) |
| Mean hospital stay, days | 3.1 ± 0.3/3.4 ± 1.9 | 3.0 ± 1.4/2.6 ± 1.2 | 2.3 ± 0.3/1.2 ± 2.1 | 3 (0.1)/3 (0.3) |
| Mean sick leave, days | NR | 5.3 ± 4.1/6.0 ± 4.4 | NR | 7 (1)/11 (1)** |
| Mean time off work, days | NR | 8.0 ± 8.0/10.1 ± 7.6 | NR | NR |

NR = Not reported; CRP = C-reactive protein; WBC = white blood count; AA = acute appendicitis.

* p < 0.05, ** p < 0.01.

| Table 4. Methodological quality of the RCTs included in the review |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Adequate generation of the randomization sequence | Unclear | Yes | Yes | No |
| Adequate concealment of treatment allocation | Unclear | Yes | Unclear | Yes |
| Blinded outcome assessment | No | No | No | No |
| Intention-to-treat analysis | Unclear | Unclear | Unclear | Yes |
all four RCTs was the blinding outcome assessment, which was impossible considering the drastically different levels of invasiveness of the two procedures. It is also worth noting that only two studies [26, 27] used adequately generated randomization sequences, one study [25] did not clearly discuss its methods, and another study [28] used a random pattern to generate the sequences (i.e., grouping patients' based on an even or uneven date of birth), which was an inadequate randomization method. Although Hansson et al. [38] stated that the patient assignments based on date of birth were found to be sufficiently randomized when compared to the conventional envelope procedure, it should be pointed out that such allocation schemes using date of birth, date of admission, hospital numbers, or basic alternations are generally considered unsatisfactory [31]. Adequate concealment of treatment allocation was satisfactorily reported in two studies [26, 28], but was unclear in the remaining two [25, 27]. An ITT analysis was declared in only one study [28]; in the other studies [25–27], it was not clear whether or not the outcomes were analyzed on an ITT basis.

The clinical and methodological characteristics of the four RCTs are shown in table 2, and the patient characteristics are shown in table 3.

### Treatment Methods

The treatment methods vary greatly within the four studies; these methods are summarized in table 1.

#### Primary Outcome

**Treatment Efficacy.** The treatment efficacy was significantly higher in patients treated with surgery than in those treated with antibiotics: FEM OR = 6.01 (4.27–8.46); heterogeneity not present; PBA = 86; NnT = 4 (fig. 1).

#### Secondary Outcomes

**Perforated Appendix.** No differences were found between the two groups with respect to the number of perforated appendices that were treated surgically: FEM OR = 0.73 (0.29–1.84); heterogeneity not present; PBA = –; NnT = –72 (fig. 2).

**Number of Patients Treated with Antibiotics.** No significant difference was found between the two groups re-
Regarding the number of patients treated with antibiotics: REM OR = 0.04 (0.00–3.27); heterogeneity present p < 0.001; PBA = –3; NnT = –2 (fig. 3). The presence of heterogeneity is due to the reported findings of Hansson et al. [28], in which a high percentage of patients in the antibiotics group received surgery and were treated with antibiotics only in the event of a gangrenous or perforated appendix.

Complications. The number of patients who developed complications was significantly higher in the surgery group: FEM OR = 1.92 (1.30–2.85); heterogeneity not present; PBA = 4; NnT = 11 (fig. 4).

Pain and Analgesic Consumption. Both studies that evaluated the VAS score for pain [25, 27], including 120 patients, reported significantly less pain observed after 12 h of conservative treatment (p < 0.001). Significantly lower pain scores were also noted during follow-up examinations. It should be pointed out that only in the study by Eriksson and Granström [25] were the mean VAS values for the two groups actually reported. In the RCT of Malik and Bari [27], a significant difference in the mean VAS values was merely noted.
A significant decrease in analgesic consumption in patients treated with antibiotics was reported in both studies. Eriksson and Granström [25] reported mean morphine doses of 0.9 mg (SD 2.5) and 17.8 mg (SD 15.3) for the antibiotics and surgery groups, respectively (p < 0.001). Malik and Bari [27] reported mean diclofenac sodium doses of 75 mg (SD 50) and 200 mg (SD 100) for the antibiotics and surgery groups, respectively (p < 0.001).

**Inflammatory Laboratory Tests and Body Temperature.** In the same two studies [25, 27], the WBC count tended to decline significantly faster in patients treated with antibiotics; additionally, these patients’ mean temperatures were significantly lower on days 1 and 2 (p < 0.05), differing by no more than 0.5°C. However, the patterns of CRP levels in both groups were very similar. The quantitative values for mean concentration of CRP and mean total WBCs were only reported in the study by Eriksson and Granström [25].

**Hospital Stay, Sick Leave and Time Off of Work.** As reported in table 3, the differences in hospital stay between the two groups were not significant in any of the four studies [25–28]. Only Hansson et al. [28] found a significant difference in the mean number of sick leave days.

**Costs.** A significant difference in hospital costs was reported by Hansson et al. [28]: SEK 26,300 (SEM 1,200) and 36,400 (SEM 3,300) for the antibiotics and surgery groups, respectively (p < 0.01).

**Patient Experience.** Again, a statistically significant difference was only found by the study of Hansson et al. [28] regarding the mean number of days in which patients experienced abdominal pain after hospital discharge: 6 days (SEM 1) and 9 days (SEM 1) for the antibiotics and surgery groups, respectively (p < 0.05). No difference was reported for the number of patients still experiencing symptoms 1 month after treatment or for the number of patients experiencing abdominal pain during the first year after treatment.

**Discussion**

Although in the past AA was considered a fatal condition, improvements in surgical techniques and the advent of anesthesia in the 1920s made the appendectomy a commonplace procedure widely accepted by the medical community [39]. Emergency appendectomies are still considered the primary means of treating AA, with mortality rates of 0.5–2.4 and 0.07–0.7% for patients with and without perforation, respectively [40, 41]. Overall postappendectomy complication rates are typically around 10–19% for AA without perforation and reach 12–30% for perforated AA [42–44]. The surgical morbidity rate for the removal of a normal appendix in patients suspected of having AA is the same as that for resection in nonperforated AA [19, 45]. Although far less demanding than the immediate health concerns regarding AA, the substantial financial burden associated with this condition should also be kept in mind. In 1997, overwhelming amounts of resources were spent treating AA in the USA, including roughly 1 million hospitalization days costing nearly 3 billion dollars, with postoperative complications accounting for only half the hospital charges [46]. Additionally, it is worth mentioning that the costs of laparoscopic appendectomies [41, 43, 47] have already far surpassed these alarming figures.

For this reason, attempts to treat straightforward and otherwise uncomplicated cases of AA with a conservative antibiotic therapy has become a very attractive alternative to surgery. Effective monotherapy antibiotics have recently become increasingly available [48], having been used successfully to treat complicated cases, such as acute diverticulitis [49] and AA with phlegmon or abscess formation [50, 51]. However, despite its success with more advanced infections, antibiotic treatment alone has not yet been accepted as a standardized treatment for non-complicated cases of AA.

Although antibiotics following an appendectomy may reduce wound infections and intra-abdominal abscesses when compared to surgery alone, very little evidence exists to suggest that conservative antibiotic treatment of AA is an effective treatment in and of itself [52]. Although the protocol of a Cochrane SR [53] and an SR without MA [54] have been published recently on the same topic, to our knowledge this appears to be the first MA comparing conservative antibiotic treatment to surgery in adult patients with uncomplicated cases of AA. Even the present SR shows a significantly lower clinical effectiveness of conservative antibiotic treatment when compared to appendectomies in the treatment of AA. According to the designated definition of ‘clinical effectiveness’, which is very limited from an experimental standpoint, antibiotic regimens are considered to have low efficacy due primarily to the high percentage of patients requiring an appendectomy during their first AA episode as well as those patients experiencing recurrences after receiving conservative treatments. Among the patients in the antibiotic therapy group of the different studies included in the SR, the percentages of those undergoing surgery during the first 48 h and of those experiencing an AA recurrence within the first year follow-
On the other hand, no statistically significant differences were found between the two treatment groups with respect to incidences of perforated appendices in those patients who had undergone appendectomies. These results support the hypothesis that perforated and nonperforated cases of AA represent different conditions altogether, suggesting that one form of AA rapidly results in perforation, whereas the other milder form does not experience perforation at all and may spontaneously resolve itself [55, 56]. Moreover, the MA did not show any significant differences between the two groups regarding the number of patients treated with antibiotics. This rather surprising result is due to the fact that, in the surgery group, the patients with gangrenous or perforated AA invariably received a full antibiotic course and, in the conservative treatment group, a high percentage of patients underwent surgery and were treated with antibiotics only in the event of a gangrenous or perforated appendix.

The number of patients who developed posttreatment complications was significantly higher in the surgery group. It should be pointed out that the complications reported in three studies [25–27] were not subdivided into major and minor categories for further evaluation. In the sole study in which these subdivision categorizations were made [28], major complications were significantly higher in patients who had undergone appendectomies (5.4 vs. 10.8%, antibiotic vs. surgery, p < 0.05); contrastingly, the relative number of minor complications was similar in the two groups.

Although an MA could not be performed for the other outcomes, the SR of these studies demonstrates that antibiotic treatment, in lieu of surgery, significantly reduces postoperative pain and analgesic consumption. This could be due to the fact that certain indicators of inflammation, such as WBC count and body temperature, subside much faster in patients treated with antibiotics. Additionally, even though a significant difference in the length of hospitalization was not found between the two groups, antibiotic treatment as an alternative to an appendectomy may reduce the amount of time needed for recovery; in fact, one of the RCTs included in the SR [28] showed a significant difference in the mean number of days of abdominal pain after hospital discharge and number of days of sick leave.

The same paper [28] showed a significant reduction in the overall costs associated with the antibiotic treatment group. Although limited to the in-hospital costs, this result suggests that the nonsurgical approach may result in significant cost-savings. Given the improved bioavailability in the form of oral antibiotics, it would not be too far fetched to imagine uncomplicated appendicitis being treated in an outpatient setting in the foreseeable future. Additional cost-savings would result from such streamlined treatment considering that the use of antibiotics often excludes the effects of morbidity and mortality commonly associated with the surgical approach.

However, despite the exciting implications of the aforementioned results, the four RCTs [25–28] feature several glaring inaccuracies, sharing certain experimental flaws as pointed out by the available literature [30, 52, 57–65]. Generally speaking, the quality of the four studies was quite poor (table 4). The small number of RCTs found in the available literature did not permit us to exclude the trials that featured an inadequate [28] or otherwise unclear [25] randomization method. Furthermore, it should be pointed out that one of the larger studies [28] used a highly inappropriate randomization method, relying on the patients’ even or uneven dates of birth to divide them into treatment groups. Satisfactorily concealing the allocation of treatments was clearly discussed in only in two studies [26, 28] and was unclear in the remaining two [25, 27]. Not unexpectedly, due to the noticeable qualitative differences in the two treatments (one being conservative antibiotics, the other invasive surgery) none of the four trials could take measures to effectively blind the investigator and/or patient to the administered treatment. Furthermore, protocol violations occurred in all four studies. For the three RCTs in which ITT analysis is unclear [25–27], 5–11.7% of patients randomly assigned to conservative antibiotic treatment also underwent surgery. In the one remaining study [28], 47.5% of the patients randomly assigned to conservative antibiotic treatment underwent surgery and 7.8% of those randomly assigned to surgery did not undergo surgery due to an objection made by either the patient or the surgeon. Considering that the rate of protocol violations in this study was so high, the ITT and protocol analyses give entirely different results.

Additionally, it should be noted that the results of this MA are subject to further limitations. Differences in the baseline characteristics of the populations, the patients’ respective levels of inflammation, and any variations in technical procedures, either for antibiotic therapy or surgical treatment (open or laparoscopic technique), all serve as factors limiting the accuracy of these RCTs (table 1). Only two studies [25, 28] mention how many patients presented with suspected AA during the trial period, how
many of these patients were primarily eligible for the trial, and how many subsequently refused trial participation. In the RCT by Eriksson and Granström [25], 45 patients were eligible for the study, but 5 (11.1%) refused to participate leaving the remaining 40 to be randomized. In the study by Hansson et al. [28], it is stated that all of the 369 patients deemed eligible for the study were subsequently randomized; however, as mentioned previously, 109 patients (29.5%), due to personal objections or medical considerations, opted not to undergo the treatment of the group to which they were randomly assigned. Finally, in the study by Styrud et al. [26], the figures from one fiscal year were reported for only one hospital among the six involved in the RCT. A total of 196 men with AA presented in this hospital and 67 were excluded from the study for the following reasons: suspicion of perforation (41 patients, 61%), CRP <10 mg/l (25 patients, 37%), and allergy to the antibiotics selected for the study (1 patient, 2%). Among the remaining 129 eligible patients, 79 refused to participate or were not informed of the study, suggesting that only 50 (25.5%) of 196 patients with AA were enrolled. These numbers should be kept in mind when extrapolating results and drawing conclusions for the general population. Furthermore, the sample size calculation is reported in only two studies [27, 28], although only in one of them [28] is the exact number explicitly stated. Approval by the ethical committee is clearly mentioned in only three of the studies [25, 26, 28]. Consequently, a word of caution is offered before drawing generalized conclusions from the results of this SR: further RCTs are needed to better resolve some of the ambiguities present in the available literature.

The main strength of the present SR is that it focuses exclusively on randomized prospective trials, which can potentially reduce bias when reporting the resulting data. Furthermore, we incorporated an ITT analysis of the data from individual trials, a more accurate assessment of efficacy than the per-protocol analysis carried out in one of the original studies [28]. In this way, the results of our MA are, methodologically speaking, more conservative but more robust.

**Reviewer’s Conclusions**

**Implications for Practice.** Although treating uncomplicated cases of AA with a conservative antibiotic regimen alone has the potential to greatly reduce the human and financial costs associated with surgery, it nevertheless exhibits significantly lower efficacy when compared to a classic appendectomy. Hence, although the present SR provides preliminary results suggesting that the nonsurgical approach can be safe and practical, thereby reducing complications, decreasing posttreatment pain and recovery time, and curbing costly expenses, the low efficacy made evident by the subsequent treatment failures and the AA recurrences means that antibiotic regimens cannot be recommended as a viable alternative to surgery. Conservative antibiotic therapy for AA should continue to be considered within the limitations imposed by its inherent advantages and disadvantages [52]; surgery remains the gold standard for treating AA despite the clinical challenges involved. However, the results presented in this review are useful in their capacity as preliminary findings that have revealed clinical and experimental inconsistencies.

**Implications for Research.** Due to the fact that only a few studies of poor experimental quality were available for the analyses in this review, future research should focus on well-designed and scrupulously conducted randomized trials with larger numbers of participating subjects in order to definitively compare the efficacy and risks of the nonsurgical and surgical approaches. Better study designs should be used, focusing more on the specifics of the study, such as definite categorized outcomes or precise patient subgroups. Regarding these study designs, particular attention should be made to identify a clear and concise definition of efficacy to be used for both the conservative and surgical treatments, standardizing the different treatment procedures as much as possible. In addition, a clear sample size calculation must be implemented. Finally, the randomization methods must be improved (including a standardized eligibility evaluation and an adequate randomization method).

**Acknowledgment**

We thank Mr. Daniel Lazzareschi for revising the English.

**Disclosure Statement**

This systematic review was not supported by any company, private individual, or grant. The entire cost of the systematic review and meta-analysis was subsidized by the authors’ institutions. The authors have no potential competing interests to declare.
References


