Epidemiological Similarities Between Appendicitis and Diverticulitis Suggesting a Common Underlying Pathogenesis

Edward H. Livingston, MD; Thomas B. Fomby, PhD; Wayne A. Woodward, PhD; Robert W. Haley, MD

Background: Nonperforating appendicitis is primarily a disease of children, and nonperforating diverticulitis affects mostly older adults. Apart from these age differences, the diseases share many epidemiological features, such as association with better hygiene and low-fiber diets.

Hypothesis: Nonperforating appendicitis and nonperforating diverticulitis are different manifestations of the same underlying colonic process and, if so, should be temporally related.

Design: Data from the National Hospital Discharge Survey were analyzed to investigate the incidence of admissions for appendicitis in children and diverticulitis in adults between 1979 and 2006.

Setting: Statistical sampling of all US hospitals.

Patients: Children admitted for appendicitis and adults with diverticulitis.

Main Outcome Measures: Time trends were assessed for stationarity using unit root analysis, and similarities between time trends were tested using cointegration analysis.

Results: The incidence rates of nonperforating appendicitis and nonperforating diverticulitis exhibited U-shaped secular trends. The rates of perforating appendicitis and perforating diverticulitis rose slowly across all the study years. Cointegration analysis demonstrated that the rates of nonperforating and perforating diverticulitis did not cointegrate significantly over time. The rates of nonperforating and perforating appendicitis did not vary together. Nonperforating appendicitis and nonperforating diverticulitis rates were significantly cointegrated over time.

Conclusions: Childhood appendicitis and adult diverticulitis seem to be similar diseases, suggesting a common underlying pathogenesis. Secular trends for their nonperforating and perforating forms are strikingly different. At least for appendicitis, perforating disease may not be an inevitable outcome from delayed treatment of nonperforating disease. If appendicitis represents the same pathophysiologic process as diverticulitis, it may be amenable to antibiotic rather than surgical treatment.

Arch Surg. 2011;146(3):308-314

APPENDICITIS AND DIVERTICULITIS are common diseases of the colon. They are responsible for numerous hospitalizations in the United States. In 2007, there were 320,000 hospital discharges for colonic diverticular disease, accumulating more than $8.9 billion in hospital charges. Appendicitis was responsible for 295,000 discharges and $7.4 billion in hospital charges, with almost all these expenses related to the performance of appendectomies. Despite their frequency, little is known about the pathogenesis of these diseases. Both have been associated with elevated intraluminal pressure, but definitive studies demonstrating a causal relationship between high pressures and the disease entities are lacking. Why the diseases occur when they do and what initiates the disease process remain unknown.

There are many similarities and some differences between appendicitis and diverticulitis. Both diseases are rare where hygiene is poor and diets are high in fiber. Both diseases have increased in incidence as cleanliness in the Western world has improved. Both diseases are more common in populations with a higher socioeconomic status. With improved grain-processing technologies culminating in lower dietary fiber content, both diseases have become more common.

See Invited Critique at end of article

Whereas older adults do get appendicitis, it is rare for it to be nonperforating disease. Similarly, children almost never experience nonperforating diverticulitis. We hypothesized that appendicitis and diverticulitis represent differing manifesta-
tions of the same underlying pathophysiological process at different ends of the age spectrum. To investigate this hypothesis, we performed an epidemiological study using cointegration analysis to examine the secular trends of nonperforating and perforating appendicitis and diverticulitis.

METHODS

The annual National Hospital Discharge Survey (NHDS) databases for 1979 to 2006 were acquired from the Centers for Disease Control and Prevention (CDC) (http://www.cdc.gov/nchs/about/major/hdads/nhds.htm) and the Inter-University Consortium for Political and Social Research (http://www.icpsr.umich.edu/icpsrweb/ICPSR) Web sites. The NHDS is the principal database used by the US government for monitoring hospital use. Each year, approximately 300 000 hospital discharges are selected for the NHDS from the 35 million total discharges nationally using a complex, multistage design to ensure that the database is representative of the US hospitalized population. Using US Census information, the CDC provides statistical weighting factors for each patient entry in the NHDS database so that incidence and prevalence estimates of hospitalized disease can be made for the entire US population. These weighting factors were used to determine the national incidence of appendicitis. The estimated US population for each year of the study was obtained from the US Census Bureau as accessed through the CDC Web site (http://wonder.cdc.gov/population.htm).

The NHDS converted from the International Classification of Diseases Adapted for Use in the United States, Eighth Revision (ICDA-8), coding to the International Classification of Diseases, Ninth Revision (ICD-9), in 1979. Although the codes before 1978 codes for the various appendicitis diagnoses were the same in the ICDA-8 system as in the ICD-9, this was not the case for diverticular disease. In the ICDA-8 system, all forms of diverticular disease were coded as 562.1. This diagnostic code was clarified by the creation of 2 codes in the ICD-9: 562.10 for diverticulosis and 562.11 for diverticulitis. Because of the coding changes that occurred with adoption of the ICD-9 in 1979, we started the sampling frame then.

Appendicitis was defined as a patient having any of the following 7 NHDS discharge diagnostic codes: 540.9 (acute appendicitis), 541.0 (appendicitis-unqualified), 542.0 (other appendicitis), 543.0 (other diseases of the appendix), and 543.9 (unspecified disease of the appendix); perforated appendicitis (540.0 or 540.1). The diagnostic codes for diverticular disease were as follows: colonic diverticulosis (562.1), diverticulitis (562.11), peritonitis (567.99), perforation (569.4), and abscess (569.5).

Data extraction was performed using a statistical software package (SAS, version 9.1; SAS Institute Inc, Cary, North Carolina). Because the hypothesis was that the diseases manifest themselves differently in children and adults, we limited the diverticulitis analysis to patients older than 40 years. Similarly, we included only patients younger than 20 years in the appendicitis cohort.

Cointegration Analysis

Similarity in the long-term trends of 2 disease rates was tested using cointegration analysis. For a review of cointegration analysis, see the appendix in the article by Adler et al. Two disease rates are considered cointegrated (varying together across time) if each rate is nonstationary (ie, individual data points do not turn about a mean value) over time and there exists a linear combination of the 2 rates that is stationary over time. We tested the null hypothesis that the time series had a unit root and, therefore, was nonstationary using the Augmented Dickey-Fuller (ADF) test. Rejecting the null hypothesis (P .05) supports stationarity (ie, the data points wander around the mean value), whereas failing to reject the null hypothesis (P > .05) supports nonstationarity.

TESTING FOR STATIONARITY

The Dickey-Fuller method tests the stationarity of an autoregressive model of order p [AR(p)] that has the following form:

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \ldots + \phi_p(X_{t-p} - \mu) + w_t,$$

where $$X_t$$ specifies the variable of interest evaluated at time t, and â is the mean of $$X_t$$. Thus, equation 1 expresses $$X_t - \mu$$ as a linear combination of observations at the previous p periods plus a noise component $$w_t$$, which is assumed to be uncorrelated (zero mean white noise). An AR(p) time series is stationary (ie, in equilibrium about a constant mean level) if all the roots of the characteristic equation

$$1 - \phi_1 \cdot \ldots - \phi_p \cdot \mu = 0$$

lie outside the unit circle and nonstationary otherwise. A special case of a nonstationary AR(p) process is one for which the characteristic equation 2 has a root of 1 (ie, a unit root). The ADF method specifically tests the null hypothesis of a unit root (nonstationarity). Following the recommendation of Ng and Perron, the lag order of the ADF test was determined by minimizing the modified Akaike information criterion. The choice of the order of the augmenting terms is important for ensuring that the ADF test statistics have the correct size and are powerful as possible for detecting the alternative hypothesis of stationarity. These tests calculate t ratios. However, under the null hypothesis of a unit root, this test statistic does not follow the conventional t distribution. One-sided P values were determined from distributions calculated by MacKinnon. Failure to reject the null hypothesis (P > .05) suggests that a time series was nonstationary (ie, slow turning or tending to wander without an attraction to a constant mean value). Visual inspection of the time trends for nonperforating appendicitis and diverticulitis revealed a U-shaped configuration. Consequently, the unit root model was fit with intercept terms only. The perforating disease entities displayed progressively increasing disease incidences. These were fit with trend and intercept terms.

TESTING FOR COINTEGRATION

The Johansen error correction model (ECM) was used to determine whether 2 time series were cointegrated, that is, wandered in time together. The order of the lagged differences in the Johansen ECM multivariate test equation was determined by first choosing the optimal lag length by performing vector autoregression on the data and then reducing that optimal lag length by 1 to accommodate the differencing imposed by the Johansen ECM equation. The optimal lag length of the level’s vector autoregression was chosen by minimizing the systemwide Akaike information criterion and Schwarz Bayesian goodness-of-fit criteria. When there was a split decision (the Schwarz criterion choosing the lesser lag length and the Akaike choosing the greater lag length), both lag lengths minus 1 were used in the specification of the lag length for the ECM to implement the Johansen tests. The lag length is important for ensuring that the test statistics have the correct size and are as
powerful as possible for potentially detecting the relevance of alternative hypotheses. Given the sample size of the data, we chose to examine the robustness of the results by using both lag lengths when there was a split decision rendered by the goodness-of-fit criteria.

Results of the Johansen ECM are assessed using trace and maximum eigenvalue statistics. These tests examine the number of cointegrating relationships that exist. They step through the various possibilities starting with the hypothesis that there are no cointegrating relationships. The P values for these hypothesis tests are reported herein. The trace test examines the null hypothesis against an alternative hypothesis that there are as many cointegrating relationships as endogenous variables. In the present case, there were 2 endogenous variables: nonperforating appendicitis and nonperforating diverticulitis. The maximum eigenvalue test is a test of the alternative hypothesis that there is 1 more cointegrating relationship than is tested in the null hypothesis.

SUMMARY ANALYSIS ACROSS REGIONS

Because we had a relatively small sample size in terms of the number of years of data available for time series analysis, we increased the power to detect cointegration by using a summary analysis across regions. The data were grouped into regions because the lowest level of geographic segregation available in the public-use NHDS files is regional information. Panel analysis is used to simultaneously assess trend data that exist in various strata. If the trends are similar across strata, the likelihood that these trends really exist is substantially increased. The NHDS provides information by region, and we used the 4 major regions of the United States as strata for the purpose of cointegration panel analysis. The panel cointegration tests proposed by Pedroni19,20 were used to assess cointegration between diverticulitis and appendicitis. All time series computations in this study were performed using an econometrics software package (EViews 6; Quantitative Micro Software LLC, Irvine, California).

STATIONARITY OF RATES OVER TIME

The trends in the nationwide incidence rates of nonperforating diverticulitis and nonperforating appendicitis followed similar U-shaped curves with nadir in 1995 (Figure 1). In contrast, those of perforating diverticulitis and perforating appendicitis increased steadily without the U-shaped pattern observed for nonperforating disease. The ADF testing of no unit root indicated nonstationarity for the rates of nonperforating diverticulitis (P = .53, by ADF test), nonperforating appendicitis (P = .25, by ADF test) (Table 1), and perforating diverticulitis (P = .27, by ADF test) but suggested stationarity for the rates of perforating appendicitis (P = .002, by ADF test) (Table 1). Stratification by region confirmed the consistency of these stationarity results across the country for the nonperforating forms of these diseases. Perforating diverticulitis was nonstationary for national data but was stationary in all regions of the country except the South. Perforating appendicitis was also nonstationary in the South (Table 1).

WITHIN-DISEASE TESTING FOR COINTEGRATION

Although the rates of nonperforating and perforating diverticulitis were nonstationary over time, Johansen ECM testing failed to reject the null hypothesis of no cointegration (P = .86, by trace test; P = .75, by maximum eigenvalue test), indicating that the 2 forms of diverticulitis were not moving together over time.

The finding previously herein that the time series of nonperforating appendicitis rates was nonstationary whereas that of perforating appendicitis rates was stationary indicates that the 2 time series were not moving together over time. When 2 time series differ on stationarity, there is sufficient evidence against cointegration, rendering a test of cointegration unnecessary.

BETWEEN-DISEASE TESTING FOR COINTEGRATION

When the Johansen ECM test of cointegration with lag length of either 1 or 2 was applied to U-shaped time series of the nationwide incidence rates of nonperforating

Figure 1. Annual incidence rates of nonperforating and perforating diverticulitis in adults (A) and of nonperforating and perforating appendicitis in children (B) between 1979 and 2006. A, The rate of nonperforating diverticulitis displayed a slow-turning wandering behavior over a U-shaped course, whereas perforating diverticulitis increased slowly and monotonically with time. Testing confirmed that both time series were nonstationary (ie, a unit root was demonstrated), but they did not move together over time (ie, negative test of cointegration). B, The rate of nonperforating appendicitis was likewise nonstationary, but that of perforating appendicitis was stationary (ie, a unit root was not demonstrated). Demonstration that one process is nonstationary and the other stationary indicates that they are not moving together over time, and a test of cointegration is not applicable.

Results of the Johansen ECM are assessed using trace and maximum eigenvalue statistics. These tests examine the number of cointegrating relationships that exist. They step through the various possibilities starting with the hypothesis that there are no cointegrating relationships. The P values for these hypothesis tests are reported herein. The trace test examines the null hypothesis against an alternative hypothesis that there are as many cointegrating relationships as endogenous variables. In the present case, there were 2 endogenous variables: nonperforating appendicitis and nonperforating diverticulitis. The maximum eigenvalue test is a test of the alternative hypothesis that there is 1 more cointegrating relationship than is tested in the null hypothesis.

SUMMARY ANALYSIS ACROSS REGIONS

Because we had a relatively small sample size in terms of the number of years of data available for time series analysis, we increased the power to detect cointegration by using a sum-
diverticulitis in adult and nonperforating appendicitis in children (Figure 2), the null hypothesis of no cointegration between the 2 time series was rejected (Table 2).

When the analysis was stratified by region, the null hypothesis of no cointegration was strongly rejected for the Northeast and South regions but equivocally rejected for the West and not for the Midwest (Table 2).

Suspecting that the inconsistency of test results across regions was due to having relatively few data points (only 27 years) in the time series to test, we performed the more powerful panel test of cointegration, analyzing all 4 regions simultaneously for consistency of effect. The panel test strongly rejected the null hypothesis of no cointegration (Pedroni group rho statistic = −3.025971, \( P = .001 \)), providing strong evidence that the cointegration between the time series of rates of nonperforating diverticulitis and nonperforating appendicitis was similar across the 4 regions.

**COMMENT**

We recently demonstrated large secular changes in the annual incidence rates of appendicitis.21,22 The present analysis of nationally representative data demonstrates that the annual incidence rates of nonperforating diverticulitis changed greatly during the past 25 years, following the same secular pattern as nonperforating appendicitis. These secular changes were significantly cointegrated, meaning that the incidence rates changed in time together, suggesting that nonperforating appendicitis and nonperforating diverticulitis could be different manifestations of the same underlying process. These findings support the hypothesis that childhood nonperforating appendicitis may be the same disease process as adult nonperforating diverticulitis, an observation supported by the many epidemiological similarities between these 2 disease entities.2–6

Perforating appendicitis and diverticulitis behaved in fundamentally different ways than their nonperforating counterparts. In contrast to the nonstationarity of time trends for nonperforating diseases, perforating appendicitis and diverticulitis were stationary over time and, consequently, are not temporally related. These findings seem incompatible with the long-held view that perforating appendicitis is merely the progression of nonperforating disease where surgical intervention was delayed too long. If perforating appendicitis was simply a manifestation of nonperforating appendicitis not treated in a timely man-
ner, the secular trends should have been statistically similar, which they were not.21

Appendicitis and diverticulitis may have epidemiologic similarities, but they are treated differently, possibly because the therapeutic strategies for these entities evolved differently. Treatments for these diseases were developed over many years and were greatly affected by the state of clinical medicine in the era when the diseases were initially described. The first description of appendicitis resulted from a detailed pathologic analysis performed by Reginald Fitz in 1886.23 In establishing the relationship between appendiceal inflammation and right lower-quadrant sepsis (ie, typhlitis), Fitz noted that the appendix developed ulcerations that seemed to cause inflammation, gangrene, and eventual perforation. Part of Fitz’s argument in favor of the appendix as the source of pelvic sepsis was the observation that one-third of patients examined in autopsies in the era before appendectomy had evidence of periappendiceal inflammation with apparent spontaneous resolution of appendicitis. Fitz hypothesized that appendicitis progressed from acute disease to perforation, an idea that was supported by Murphy in 190424 but never proved experimentally. Nevertheless, mortality from pelvic sepsis was high, and at the turn of the 20th century, appendectomy could be performed with relatively little morbidity. Because of the potential for reducing pelvic sepsis-related complications, emergency appendectomy became the accepted approach for patients with presumed appendicitis.

Treatment for diverticulitis evolved differently. Whereas appendectomy was a procedure associated with relatively low morbidity, this was not the case for colon

---

Table 2. Johansen Error Correction Model Tests for Cointegration of the Incidence Rates of Nonperforating Diverticulitis and Nonperforating Appendicitis, 1979 to 2006

<table>
<thead>
<tr>
<th>Region</th>
<th>Trace Test</th>
<th>P Valuea</th>
<th>Maximum Eigenvalue Test</th>
<th>P Valuea</th>
<th>Lag Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>National</td>
<td>25.0014</td>
<td>.01</td>
<td>17.0948</td>
<td>.03</td>
<td>1</td>
</tr>
<tr>
<td>Northeast</td>
<td>21.6206</td>
<td>.03</td>
<td>16.7201</td>
<td>.04</td>
<td>0</td>
</tr>
<tr>
<td>South</td>
<td>23.3392</td>
<td>.02</td>
<td>16.2778</td>
<td>.04</td>
<td>1</td>
</tr>
<tr>
<td>West</td>
<td>21.9989</td>
<td>.03</td>
<td>16.9652</td>
<td>.03</td>
<td>0</td>
</tr>
<tr>
<td>Midwest</td>
<td>13.0231</td>
<td>.36</td>
<td>10.6624</td>
<td>.26</td>
<td>1</td>
</tr>
</tbody>
</table>

a The P values (1-sided) test the null hypothesis of no cointegrating relationship between the 2 time series. P < .05 indicates cointegration, that is, that the disease incidence rates are “tethered together” across time. (The panel summary test rejected the null hypothesis of no cointegration across all regions. Panel tests have greater power to detect cointegration than do tests of individual regions, lessening the risk of type II error.) These reported P values were obtained from MacKinnon.14
resection in the early 1900s. Any colon resection was associated with a high complication rate. Surgery for diverticulitis disease was generally limited to patients with perforations and involved 3 separate sequential procedures: colostomy placement, colon resection, and colostomy takedown. Consequently, the morbidity associated with the conservative management of diverticulitis was far less than that of surgical treatment, even if the disease worsened while a patient was being observed. Resistance to pursue high-risk surgical therapy for diverticulitis disease was heightened by the difficulty in establishing the diagnosis. As late as in the 1960s, operations for suspected diverticulitis were associated with nearly a one-third false-positive rate.

The differing treatment approaches for appendicitis and diverticulitis have more to do with these historical differences in the management approach than any fundamental difference in the pathogenesis of these diseases. Both diseases begin with localized inflammation that is thought to have the potential for culminating in perforation if left untreated. Diverticulitis is thought to have an increased risk of perforation after multiple episodes of the acute disease entity. It is assumed, but has never been proved, that appendicitis always perforates unless appendectomy is performed early in its course. There is a growing body of evidence to suggest that this is not the case. Fitz noted that one-third of patients undergoing autopsies before appendectomy was recognized as a disease had evidence of resolved appendicitis, providing evidence that, oftentimes, appendicitis is a relatively benign condition that does not require any intervention at all. Even the early advocates of appendectomy performed the operation on patients who had multiple episodes of recurrent appendicitis.

There is a potential biological basis explaining how perforations occur. Studies of complicated appendicitis have revealed polymorphisms of the interleukin 6 (IL-6) promoter. These result in diminished IL-6 production, which is associated with a lower incidence of complicated appendicitis. Interleukin 6 is a strong procoagulant, and its presence in the face of appendicitis probably induces thrombosis, leading to gangrene and perforation. Polymorphisms of the IL-6 promoter are more common in patients with simple (ie, nonperforating) appendicitis, leading to lower levels of IL-6 and fewer perforations. A biological basis for perforation might explain why some patients perforate and others do not. Patients with a propensity to perforate might develop severe disease from a minor insult, and those who do not perforate could have recurrent episodes of appendicitis that present atypically and do not get treated. Varying presentation of the disease may result in an apparent dissociation of the incidence of perforating and nonperforating appendicitis from an epidemiological perspective.

More recent data suggest that appendicitis can be treated with antibiotics in much the same way that diverticulitis is managed. In the 1950s, a large series of patients with appendicitis treated with antibiotics only was reported. In the 1960s and 1970s, the Navy realized that treatment of sailors with appendicitis on submarines was best done by administering antibiotics while at sea and deferring appendectomy until the ship surfaced, often times weeks after the initial episode of appendicitis occurred. Perforations have only rarely been observed with this treatment approach. Because deferred surgery does not increase perforation rates, it is now the standard for pediatric surgeons to not perform nighttime emergency appendectomy and defer appendectomies admitted at night until the next day. There have been 3 randomized controlled trials comparing antibiotic treatment of appendicitis with surgery and a nonrandomized study demonstrating that acute appendicitis can be safely treated with antibiotics alone.

There are limitations to this study. We evaluated the incidence of appendicitis and diverticulitis using hospital admission diagnoses. These could be affected by the method used to establish the diagnosis. Potentially, the increased incidence for these diseases observed in the past decade could reflect increased use of imaging technologies that may have led to increased recognition of these disorders over time. However, the previously cited studies demonstrating effective treatment with antibiotics without surgery suggests that the conclusion we have drawn that acute appendicitis can be treated with antibiotics as diverticulitis is treated may be valid.

A growing body of literature suggests that the relationship between nonperforating and perforating appendicitis is not as clear as previously thought. Given the vastly improved abilities to establish these disease diagnoses and detect complications coupled with more effective antibiotic therapies, the treatment of acute appendicitis should be revisited. The present data suggest that nonperforating appendicitis behaves similarly to nonperforating diverticulitis. Clinical trials of nonperforating appendicitis with antibiotic therapy without surgery, as is done for diverticulitis, seem justified.

Accepted for Publication: February 11, 2010.
Correspondence: Edward H. Livingston, MD, Division of Gastrointestinal and Endocrine Surgery, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Room E7-126, Dallas, TX 75390-9156 (edward.livingston@utsouthwestern.edu).
Author Contributions: Study concept and design: Livingston, Fomby, and Haley. Acquisition of data: Livingston. Analysis and interpretation of data: Livingston, Fomby, Woodward, and Haley. Drafting of the manuscript: Livingston. Critical revision of the manuscript for important intellectual content: Livingston, Fomby, Woodward, and Haley. Statistical analysis: Livingston, Fomby, Woodward, and Haley. Obtained funding: Livingston. Administrative, technical, and material support: Livingston.
Financial Disclosure: None reported.
Funding/Support: This study was supported in part by the Hudson-Penn Endowment.