International Journal of Surgery 15 (2015) 107–112

Contents lists available at ScienceDirect

International Journal of Surgery

journal homepage: www.journal-surgery.net

Original research

How good are surgeons at identifying appendicitis? Results from a multi-centre cohort study

Sean Strong [a, b, c, *], Natalie Blencowe [a, b, c], Aneel Bhangu [d], On behalfof the National Surgical Research Collaborative

[a] Centre for Surgical Research, School of Social and Community Medicine, University of Bristol, Bristol, UK
[b] Division of Surgery Head and Neck, University Hospitals Bristol NHS Foundation Trust, Bristol, UK
[c] Severn and Peninsula Audit and Research Collaborative for Surgeons, UK
[d] West Midlands Research Collaborative, Academic Department of Surgery, Queen Elizabeth Hospital, Birmingham, B15 2TH, UK

HIGHLIGHTS

- Arguments have been made to support both removing and leaving in situ a macroscopically normal appendix.
- Treatment strategies however rely on the inability of surgeons to assess pathology.
- This multi-centre study suggests that surgeons' judgements of the intra-operative macroscopic appearance of the appendix is inaccurate.

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form
13 January 2015
Accepted 26 January 2015
Available online 31 January 2015

Keywords:
Appendicectomy
Emergency surgery

ABSTRACT

Background: Convincing arguments for either removing or leaving in-situ a macroscopically normal appendix have been made, but rely on surgeons' accurate intra-operative assessment of the appendix. This study aimed to determine the inter-rater reliability between surgeons and pathologists from a large, multicentre cohort of patients undergoing appendicectomy.

Materials and methods: The Multicentre Appendicectomy Audit recruited consecutive patients undergoing emergency appendicectomy during April and May 2012 from 95 centres. The primary endpoint was agreement between surgeon and pathologist and secondary endpoints were predictors of this disagreement.

Results: The final study included 3138 patients with a documented pathological specimen. When surgeons assessed an appendix as normal (n = 496), histopathological assessment revealed pathology in a substantial proportion (n = 138, 27.8%). Where surgeons assessed the appendix as being inflamed (n = 2642), subsequent pathological assessment revealed a normal appendix in 254 (9.6%). There was overall disagreement in 392 cases (12.5%), leading to only moderate reliability (Kappa 0.571). The grade of surgeon had no significant impact on disagreement following clinically normal appendicectomy. Females were at the highest risk of false positives and false negatives and pre-operative computed tomography was associated with increased false positives.

Conclusions: This multi-centre study suggests that surgeons' judgements of the intra-operative macroscopic appearance of the appendix is inaccurate and does not improve with seniority and therefore supports removal at the time of surgery.

1. Introduction

Appendicectomy is the most common general surgery emergency across America and Europe, with some 350,000 procedures performed annually [1–3]. Traditionally, appendicectomy was performed using open techniques, during which the appendix is routinely removed regardless of its macroscopic appearance at the

* Corresponding author. Centre for Surgical Research, School of Social and Community Medicine, Canynge Hall, 39 Whatley Road, University of Bristol, Bristol, BS8 2PS, UK.
E-mail address: sean.strong@bristol.ac.uk (S. Strong).

http://dx.doi.org/10.1016/j.ijsu.2015.01.032
1743-9191/

time of surgery. The introduction of laparoscopy, however, has introduced a dilemma concerning intra-operative decision making, with convincing arguments for both removing and leaving in-situ a macroscopically normal appendix, with opinion divided amongst surgeons [4–6]. Proponents of removing a macroscopically normal appendix maintain that this prevents future episodes of appendicitis and avoids missing other pathology [7]. Conversely, others suggest that it is safe to leave a normal looking appendix in place [8–10], although adoption of this latter strategy relies on surgeons' accurate intra-operative assessment of the appendix, for which current evidence is limited.

The largest published study investigating surgeons' ability to identify appendicitis intra-operatively correlated operation notes with histological reports of 876 appendicectomies. Normal appendixes were correctly identified in fewer than 70% of cases and this did not appear to be influenced by training grade [11]. However, despite the large number of included cases, this study has several limitations. All operations were undertaken in a single centre using mainly open techniques (97.2%), which are now decreasing in popularity [4]. Patients were identified between 2003 and 2006 which may limit the applicability of findings to modern day practice. Recent restructuring of surgical training has reduced working hours and re-branded training grades, with potential implications for operative ability and confidence with diagnostic and decision-making skills [13]. Although appendicectomy was traditionally viewed as an index surgical procedure, it is increasingly performed by trainees of a higher level [14]. In view of the limitations of previous studies, the increasing number of appendicectomies being completed laparoscopically and changes to surgical training, this multi-centre study was therefore undertaken to determine the inter-rater reliability between surgeons and pathologists from a large, multicentre cohort of patients undergoing appendicectomy.

2. Materials and methods

2.1. Patients and data collection

The Multicentre Appendicectomy Audit collected data on prospective patients undergoing appendicectomy from 89 centres within the United Kingdom and 6 international centres. Permission to perform the audit was granted from each individual site's Clinical Audit Department and complete methodology has been described previously [4]. Briefly, design and data collection was trainee-led, protocol driven, prospective and multi-centred. The study period included May and June 2012, with 30-day follow-up for the last patient to the end of July. Results regarding the use of laparoscopy, normal appendicectomy rates and adverse events have been previously published, without reference to clinical and pathological reliability [4].

2.2. Patient eligibility

Each local trainee-level principal investigator was responsible for identifying patients, entering information into the pre-specified database and ensuring completeness of data. Patients of any age or sex were included if they underwent an emergency appendicectomy between 1 May and 30 June 2012 (inclusive). Excluded were patients whose appendicectomy was planned, or part of another procedure (such as incidental appendicectomy during elective right hemicolectomy for colonic cancer).

2.3. Outcome measures

The primary outcome measure was agreement between surgeon and pathologist. The surgeon's intra-operative clinical judgement was recorded at the time of surgery. Original histopathology reports were used to obtain the pathological diagnosis. False positive results were defined when the surgeon judged appendicitis but pathology was normal; false negatives occurred when the surgeon judged a normal appendix, but pathology showed inflammation. Secondary endpoints were predictors of disagreement, including the seniority of operating surgeon and level of supervision required. Operating surgeons were classified as consultants, senior registrars (Specialty Trainee 6 (ST6) or above), junior registrars (ST3-5), or core trainees (CT 1–3). Foundation Year 1 (FY1) and 2 (FY2) doctors were included within the latter category.

2.4. Data collection

Data were collected using a specially constructed Microsoft® Access (Microsoft, Redmond, Washington, USA) database, locally held and fully anonymised before centralisation.

2.5. Statistical analysis

Results are reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for observational studies. Pre-planned comparisons of surgeon opinion to pathological finding were performed for all patients, and then by subgroups (laparoscopy and grade of surgeon). To ensure data accuracy, parsimonious data fields were used, including a composite of complex appendicitis encompassing phlegmonous, gangrenous and perforated appendicitis. Differences between demographic groups were compared using the Chi-squared test. Inter-observer agreement was tested using Cohen's Kappa, with a reliability coefficient of 0.70 or higher indicating good correlation, 0.50–0.69 indicating moderate correlation, and below 0.50 indicating poor correlation. To further investigate associations with disagreement, multiple binary logistic regression was used to produce adjusted odds ratios for false positive and false

Table 1
Characteristics of patients undergoing appendicectomy.

	n = 3138 (%)
Age (years)	
<16	690 (22.0)
16–50	1990 (63.4)
>50	452 (14.4)
Unrecorded	6 (0.2)
Gender	
Male	1583 (50.4)
Female	1548 (49.3)
Unrecorded	7 (0.2)
ASA	
I/II	2951 (94.0)
III/IV	80 (2.5)
Unrecorded	107 (3.4)
Surgical method	
Open	1022 (32.6)
Laparoscopic	1897 (60.5)
Converted to open	219 (7.0)
Pre-operative imaging	
None	2099 (66.9)
CT scan	633 (20.2)
USS scan	406 (12.9)
Grade of surgeon	
Consultant	328 (10.4)
ST6+	1199 (38.3)
ST3-ST5	1120 (35.8)
FY1-CT2	394 (12.6)
Unrecorded	97 (3.1)