# Can Surgeons Identify Appendicitis Macroscopically? Results From a Multicentre Prospective Study

*Da Wei Thong, MSc, MBBS,* Stephen Crouch, MMath, MD,* Samuel Morgan, BExSc, DPhty,† Thomas Arthur, BSc, MBBS,* and QUEST Collaboration*

**Purpose:** The primary outcome was to investigate the accuracy of intraoperative macroscopic diagnosis by the operating surgeon with the results of the subsequent histopathologic examination. The secondary outcome was to identify the predictors of discrepancies between these 2 groups.

**Materials and Methods:** A multicentre, prospective, observational study was conducted over a period of 2 months with a 30-day follow-up period. Patients who underwent surgery with the intention of appendicectomy were recruited in the study.

**Results:** A total of 1169 patients were recruited. False negatives (FNs) were defined as a normal macroscopic diagnosis but histopathologically appendicitis, whereas false positive otherwise. Overall, FN rates were 22.4%, whereas false positive rates were 8.2%. The seniority of the operating surgeons did not affect the ability to accurately diagnose appendicitis macroscopically ($P = 0.069$). However, consultant surgeons had the lowest FN rate of 15.6%. Females and preoperative ultrasound scan increased odds of FN, whereas preoperative computed tomography decreased the odds of FN appendicectomy.

**Conclusion:** Macroscopic identification intraoperatively is inaccurate with a FN rate of 22%.

**Key Words:** appendicectomy, laparoscopic, macroscopic identification, diagnostic errors, diagnostic imaging, audit

(*Surg Laparosc Endosc Percutan Tech* 2019;00:000–000)

Appendicitis carries an annual mortality rate of 0.2 per 100,000 in Australia and accounts for <1 in 100,000 global mortalities.[1,2] Appendicectomies continues to represent the most common general surgical procedure performed worldwide and with Australia recording 34,464 episodes of acute appendicitis in the 2014 to 2015 alone.[3] Globally, mortality from appendicitis has reduced by 52.5% since 1990.[4] Despite these improvements, macroscopic surgical identification of appendicitis remains speculative. Intraoperative accurate diagnosis of appendicitis continues to be a challenge and as a result, surgeons continue to debate whether to remove or leave a normal-appearing appendix during appendicectomies.[5–7] In Australia, appendicectomies are most commonly performed laparoscopically compared with the conventional open technique.[8] A recent study performed in 2015 in the United Kingdom demonstrated a moderate correlation with intraoperative macroscopic assessment when compared with microscopic histopathologic examination.[9]

Similarly, a multicentre trial was conducted across Australia, where operative identification of appendicitis by surgeons was recorded and compared with histopathologic reports; with the primary aim looking at correlations and agreements. The secondary aim was to ascertain potential predictors that reflect the outcomes as to further guide future Australian surgical practice.

## MATERIALS AND METHODS

We performed a multicentre, prospective, observational study between June and October 2016 across Australia. Data were collected in accordance with a protocol where patients who were suspected of having acute appendicitis and had appendicectomies were included in the study, whereas patients who underwent surgery for diagnostic reasons were excluded from the trial. The protocol was then disseminated to primary investigators of 27 participating institutions and the study was conducted over a period of 2 months with a 30-day follow-up. Each local principal investigator at each site was required to fill out a centralized database with prefilled variables.

Patient details were anonymized before central collation and analysis. Ethics approval was provided by Gold Coast Health and Hospital Service Human Research Ethics Committee, with the additions of individual sites approval provided by local governance at each site.

The primary outcome of the study was to investigate the accuracy of intraoperative diagnosis between surgeons and subsequent histopathologic examination. A subgroup analysis focused on laparoscopic technique and seniority of the surgeon was performed. The intraoperative diagnosis is based on the macroscopic identification of the appendix by the operating surgeon. Final histopathologic reports were used to obtain the pathologic diagnosis of the appendix specimen. The dichotomous variable of "appendicitis" and "not appendicitis" was used for the 2 variables stated. The secondary outcomes were to look for potential predictors of disagreements. The variables investigated include age, sex, location, American Society of Anesthesiologists, surgical method, level of supervision, and preoperative imaging.

Data were analyzed using SPSS, version 20 (SPSS Inc., Chicago, IL). Interobserver agreement was tested using Cohen κ and interpretation with a reliability coefficient of >0.8 indicating very good, 0.61 to 0.80 indicating good, 0.41 to 0.60 indicating moderate, and <0.4 indicating poor agreement.[10] A univariable logistic regression model was used to identify any significant predictors for false positive (FP) and false negative (FN) results. *P*-value <0.05 is regarded as significant in this study.

Received for publication December 12, 2018; accepted March 29, 2019.
From the *Department of General Surgery, Gold Coast University Hospital; and †School of Medical Science, Griffith University, Southport, QLD, Australia.
The author declares no conflicts of interest.
Reprints: Da Wei Thong, MSc, MBBS, Department of General Surgery, Gold Coast University Hospital, 1 Hospital Boulevard, Southport, QLD 4215, Australia (e-mail: thongdawei@gmail.com).

## RESULTS

A total of 1169 patients underwent appendicectomy in the study period across 27 centers with an average of 48 patients per center. Incomplete documentations were identified in 20 patients and they were excluded from this study, leaving 1169 patients for final analysis. There were slightly more female appendectomies performed compared with males (50.3% vs. 49.5%). Other demographics and clinical details of patients are as shown in Table 1.

### Surgeon and Histopathologic Agreement

Surgeon assessed appendix macroscopically as normal (n = 183), the subsequent histopathologic assessment revealed appendicitis in 41 patients (22.4%). Although, if appendicitis is established macroscopically during appendicectomy (n = 986), histopathology reported 82 patients to have a normal appendix (8.2%). There was an overall disagreement between surgeons and pathologist of 123 cases (10.5%), leading to a good agreement ($\kappa = 0.635$), as shown in Table 2.

### Surgeon and Histopathologic Agreement in Laparoscopic Technique

Overall, 95.8% of appendectomies in this study were conducted laparoscopically (n = 1121). Surgeon assessed appendix macroscopically as normal (n = 176) and further subsequent histopathologic assessment revealed appendicitis in 40 patients (22.7%). Although, if appendicitis is established macroscopically during appendicectomy (n = 945), histopathology reported 81 patients to have a normal appendix (8.6%). There was an overall disagreement between surgeon and pathologist of 121 (10.8%) cases, leading to a good agreement ($\kappa = 0.628$), as shown in Table 3.

**TABLE 1.** Demographics and Preoperative Information

| | N = 1169 [n (%)] |
|---|---|
| Age (y) | |
| 0-17 | 231 (19.8) |
| 18-35 | 555 (47.5) |
| 36-65 | 315 (26.9) |
| ≥ 65 | 65 (5.6) |
| Unknown | 3 (0.3) |
| Sex | |
| Male | 579 (49.5) |
| Female | 588 (50.3) |
| Unknown | 2 (0.2) |
| ASA | |
| 1-2 | 1088 (93.1) |
| 3-4 | 68 (5.8) |
| Unknown | 13 (1.1) |
| ATSI | |
| Yes | 21 (1.8) |
| No | 1100 (94.1) |
| Unknown | 48 (4.1) |
| BMI | |
| < 30 | 632 (54.1) |
| > 30 | 163 (13.9) |
| Unknown | 374 (32.0) |
| Imaging | |
| None | 411 (35.2) |
| USS | 369 (31.6) |
| CT | 340 (29.1) |
| Both | 38 (3.3) |
| Unknown | 11 (0.9) |
| Method | |
| Laparoscopic | 1121 (95.9) |
| Laparoscopic to open | 28 (2.4) |
| Open | 19 (1.6) |
| Unknown | 1 (0.1) |

ASA indicates American Society of Anesthesiologists; ATSI, Aboriginal and Torres Strait Islander; BMI, body mass index; CT, computed tomography; USS, ultrasound scan.

**TABLE 2.** Surgeons Macroscopic Assessment of Appendix Compared With Histopathologic Examination of Pathologists

| | Histopathology [n (%)] | | |
|---|---|---|---|
| | Appendicitis | Normal | κ |
| Surgeon (N = 1169) | | | |
| Appendicitis | 904 (91.7) | 82 (8.3) | 0.635 |
| Normal | 41 (22.4) | 142 (77.6) | |

### The seniority of Surgeon and Histopathologic Agreement

Table 4 below noted the seniority of the surgeon agreements with histopathologic diagnosis. The results show that consultant surgeons were least likely to call an appendix macroscopically normal that were subsequently reported histopathologically as appendicitis (FN) with a rate of 15.6%. Followed by Surgical Education and Training (SET) years I to II 20.0%, prevocational trainees 22.4% and SET years III to IV with a rate of 29.1%.

When a surgeon reported macroscopically appendicitis but histopathologically reported normal (FP); prevocational trainees at have the lowest rate of 6.0%, followed by fellows at 7.2%, consultant at 8.3%, SET III to IV at 9.0% and SET I to II at 14.7%.

There is an overall good agreement between prevocational and consultant group with a histopathologic diagnosis with the former higher at 0.697 and later at 0.683. However, SET trainee years I to II and II to IV only achieved a moderate agreement with κ score of 0.533 and 0.586.

### Predictor: FP and FN Macroscopic Diagnosis of Appendicitis

Univariable logistic regression was performed to investigate the predictors of FP and FN appendicectomy rates. The odds for females to have a FN macroscopic diagnosis for appendicectomy is 2.45 times more common than males, as shown in Table 5.

Furthermore, ultrasound scan (USS) carries an odd of 2.32 times more likely for FN macroscopic diagnosis post-appendicectomy, whereas computed tomography (CT) scan decreases the odds of FN macroscopic diagnosis post-appendicectomy by 91% when compared with no investigation performed.

**TABLE 3.** Surgeons Macroscopic Assessment of Appendix Compared With Histopathologic Examination of Pathologists in Patients Undergoing Laparoscopic Appendicectomy

| | Histopathology [n (%)] | | |
|---|---|---|---|
| | Appendicitis | Normal | κ |
| Surgeon (N = 1121) | | | |
| Appendicitis | 864 (91.4) | 81 (8.6) | 0.628 |
| Normal | 40 (22.7) | 136 (77.3) | |

**TABLE 4.** Different Seniority of Surgeons' Macroscopic Assessment of Appendix Compared With Histopathologic Examination of Pathologists

| | Histopathology [n (%)] | | |
|---|---|---|---|
| **Primary Operator** | **Appendicitis** | **Normal** | **κ** |
| Prevocational | | | |
| Appendicitis | 267 (94.0) | 17 (6.0) | 0.697 |
| Normal | 13 (22.4) | 45 (77.6) | |
| Surgical Education and Training trainee I-II | | | |
| Appendicitis | 110 (85.3) | 19 (14.7) | 0.533 |
| Normal | 5 (20.0) | 20 (80.0) | |
| Surgical Education and Training trainee III-V | | | |
| Appendicitis | 244 (91.0) | 24 (9.0) | 0.586 |
| Normal | 16 (29.1) | 39 (70.9) | |
| Fellow | | | |
| Appendicitis consultant | 11 (91.7) | 1 (8.3) | Nil |
| Appendicitis | 272 (92.8) | 21 (7.2) | 0.683 |
| Normal | 7 (15.6) | 38 (84.4) | |

## Females-Specific Subgroup Analysis

The study found that 52.8% (n = 316) of the females in the study had USS as a part of their diagnosis workup, whereas only 14.4% (n = 83) male did. In total, majority of the USS performed in the study also compromises exclusively to females, which is up to 77.9% (n = 289).

As noted before, females have a higher risk of a FN macroscopic diagnosis of appendicitis. Further subgroup analysis demonstrated that females who had USS have almost twice (7.3%) the probability of being FN macroscopic diagnosis compared with females who had no imaging modalities use (4.5%) and CT scans (0.7%). These differences seen in the imaging modalities tested are statistically significant with a *P*-value of 0.03.

## DISCUSSION

The study has successfully investigated the correlation and agreements between intraoperative macroscopic diagnosis and histopathologic conclusion of appendicectomy. Data analysis demonstrated a greater agreement in the overall group; both laparoscopic and open compared with laparoscopic alone. In comparison, higher rates of FP and FN in the laparoscopic only group suggests that laparoscopic identification of macroscopic appendicitis was not superior as it has been suggested. This could be confounded by the fact that the study has a low number of open method cases and 28 of the 47 (59.6%) open cases were a conversion from laparoscopic cases. These cases have a higher percentage of appendicitis on histologic diagnosis compared with laparoscopic cases alone (87.2% vs. 80.6%); indicative of a higher likelihood of a true positive result because of its difficulty and method conversion.

**TABLE 5.** Univariable Logistic Regression for Predictors of False Positive and False Negative Macroscopic Diagnosis of Appendicitis

| | False Positive | | | | False Negative | | | |
|---|---|---|---|---|---|---|---|---|
| | **Odds Ratio** | **95% CI (Upper)** | **95% CI (Lower)** | ***P*** | **Odds Ratio** | **95% CI (Upper)** | **95% CI (Lower)** | ***P*** |
| Age (y) | | | | | | | | |
| 0-17 | | | | | | | | |
| 18-35 | 0.63 | 0.38 | 1.07 | 0.85 | 1.03 | 0.49 | 2.19 | 0.94 |
| 35-65 | 0.38 | 0.19 | 0.74 | 0.00 | 0.51 | 0.51 | 1.37 | 0.18 |
| > 65 | 0.26 | 0.06 | 1.11 | 0.07 | * | * | * | 1.00 |
| Sex | | | | | | | | |
| Male | | | | | | | | |
| Female | 1.28 | 0.82 | 2.01 | 0.28 | 2.45 | 1.24 | 4.85 | 0.01 |
| Area | | | | | | | | |
| Metropolitan | | | | | | | | |
| Rural | 0.54 | 0.28 | 1.03 | 0.06 | 1.91 | 0.99 | 3.70 | 0.06 |
| Aboriginal and Torres Strait Islander | | | | | | | | |
| Yes | | | | | | | | |
| No | 1.55 | 0.21 | 11.68 | 0.67 | 0.72 | 0.09 | 5.47 | 0.75 |
| ASA | | | | | | | | |
| I/II | | | | | | | | |
| III/IV | 0.83 | 0.30 | 2.35 | 0.73 | * | * | * | 1.00 |
| Surgical method | | | | | | | | |
| Laparoscopy | | | | | | | | |
| Laparoscopic converted | * | * | * | 1.00 | 1.00 | 0.13 | 7.55 | 1.00 |
| Open | 0.71 | 0.94 | 5.41 | 0.74 | * | * | * | 1.00 |
| Consultant presence | | | | | | | | |
| Yes | | | | | | | | |
| No | 1.41 | 0.88 | 2.27 | 0.15 | 1.09 | 0.57 | 2.09 | 0.80 |
| Preoperative imaging | | | | | | | | |
| None | | | | | | | | |
| USS | 1.25 | 0.74 | 2.09 | 0.40 | 2.32 | 1.17 | 4.59 | 0.02 |
| CT | 0.55 | 0.28 | 1.05 | 0.07 | 0.09 | 0.01 | 0.69 | 0.02 |
| Both | 1.92 | 0.70 | 5.29 | 0.20 | 0.83 | 0.11 | 6.50 | 0.86 |

*Nil event of false positive/negative cases in that category.
ASA indicates American Society of Anesthesiologists; CI, confidence interval; CT, computed tomography; USS, ultrasound scan.

In general, laparoscopic surgery has been shown to offer more benefits over an open approach. This includes a short length of overall hospital stay, decreased outpatient costs, a faster return to function and less pain day 1 postoperatively.[11] Other studies also reported laparoscopic appendicectomy had fewer rates of postoperative ileus, lower mortality rates in complex cases and a shorter time to return to a normal diet.[12,13] Hence, it is not surprising as to why laparoscopic is the preferred primary modality for the management of appendicitis.[14] The small number of open approach in this study compared with the laparoscopic approach makes comparison impractical.

The study demonstrated that different seniority of the surgeon did not affect the ability to accurately diagnose appendicitis macroscopically with similar FP and FN rates ($P = 0.069$, Table 4). This is in concordance with the existing literature.[3,9] This is despite a larger cohort of consultants as primary operators in this study when compared with the UK multicentre trial (28.7% vs. 10.4%). However, an observation was noted; surgeons from the prevocational group demonstrated the highest interobserver agreement with a $\kappa$ of 0.697. Further, correlation tests showed the level of supervision during the surgery between different levels of seniority surgeon was significant ($P = 0.00$). SET I/II having the highest rate of supervision of 42.7%, whereas the prevocational group only had a supervision rate of 31.1%. Hence, if these surgeries were done under supervision, we speculate that the acumen of macroscopic diagnosis is preserved.

The study demonstrated that females have higher odds of FN and FP diagnosis when compared with male patients. However, only the FN rates were found to be statistically significant ($P = 0.01$). The propensity of a more clinically difficult diagnosis and management in females have been shown as contributing factors for higher FP and FN diagnosis rates and possibly explaining a higher negative appendicectomy rate.[15,16] These results were also in accordance with the existing literature.[9,17] Whether the incorrect diagnosis of negative intraoperative diagnosis of appendicitis has a correlation to the knowledge USS findings preoperatively in females; influencing the surgeons' intraoperative diagnosis remains uncertain although an association has been demonstrated in the results section previously.

The study demonstrated that the use of USS increases the odds of a FN diagnosis by 230%; which is surprising as the UK study reports lower odds of FN and FP diagnosis with the use of USS.[9] This also suggests that ultrasound is a modality with high variability as it is dependent on individual sonographer experience for an accurate diagnosis. However, USS is a valuable adjunct, especially in the younger age population.[18] In contrast, the use of CT decreased the odds of a FN diagnosis by 91% and FP by 45% when compared with no use of an imaging modality. This supports the use of CT as a valuable adjunct in the management of appendicitis due to its sensitivity, specificity, cost-effectiveness and results being available soon after scan is performed.[19,20] However, individual risk stratification has to be performed due to the risk of radiation exposure, particularly to younger and pregnant population.

On occasion, other inflammatory and gynecologic conditions were found to impact the accuracy macroscopic identification of appendicitis. In addition, other appendiceal pathology such as fibroepithelial changes, endoluminal inflammation, and lymphoid hyperplasia attests to be a diagnostic challenge macroscopically as well.[16,21,22] This is even more so with the lack of tactile sensation when the surgery is performed laparoscopically.

The study conducted was of a nationwide multicentre study. This allows generalization of our data to reflect most of the population groups and the different practices across Australia. Performance bias was also low given each histopathologic report and macroscopic identification during surgery was performed by different individuals and centers; not by a small group of individuals. Stratification of surgeons into different groups based on seniority rather than years of surgical practice is imperfect as this does not accurately reflect each individual's experience to macroscopic identification of appendicitis especially with the prevocational and SET I to II groups. Furthermore, an inability to correlate surgeons' seniority with the complexity of each case makes it difficult to explore it as a confounding factor.

Macroscopic identification intraoperatively is challenging and inaccurate with a FN rate of 22%. The authors of this study support the practice of removing the appendix upon clinical suspicion of appendicitis when no other alternative diagnosis is identified during surgery.

## REFERENCES


1. Australian Bureau of Statistics. Underlying causes of death, Australia; 2016. Canberra, Australia. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3303.02016?OpenDocument. Accessed December 5, 2018.
2. Australian Institute of Health and Welfare. Admitted patient care 2015–16: Australian hospital statistics. Health services series no.75. Cat. no. HSE 185; 2017. Canberra, Australia. Available at: http://www.aihw.gov.au/getmedia/3e1d7d7e-26d9-44fb-8549-aa30ccff100a/20742.pdf.aspx?inline=true. Accessed December 5, 2018.
3. Roberts JK, Behravesh M, Dmitrewski J. Macroscopic findings at appendicectomy are unreliable: implications for laparoscopy and malignant conditions of the appendix. *Int J Surg Pathol*. 2008;16:386–390.
4. D'Inca R, Neri L, Daperno M, et al. Poor adherence is associated with impaired quality of life and increased illness intrusiveness in patients with ulcerative colitis and Crohn disease. In: D'Inca R, Neri L, Daperno M, Comberlato M, eds. *Dig Liver Dis*. 2015;47:e183.
5. Jaunoo S, Hale A, Masters J, et al. An international survey of opinion regarding investigation of possible appendicitis and laparoscopic management of a macroscopically normal appendix. *Ann R Coll Surg Engl*. 2012;94:476–480.
6. National SRC. Multicentre observational study of performance variation in provision and outcome of emergency appendicectomy. *Br J Surg*. 2013;100:1240.
7. Turner EJH, Lightwood R. Management of the 'normal'-appendix during laparoscopy for right iliac fossa pain. *World J Laparoscopic Surg*. 2009;2:15–17.
8. Singhal V, Jadhav V. Acute appendicitis: are we over diagnosing it? *Ann R Coll Surg Engl*. 2007;89:766–769.
9. Strong S, Blencowe N, Bhangu A. How good are surgeons at identifying appendicitis? Results from a multi-centre cohort study. *Int J Surg*. 2015;15:107–112.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
11. Sauerland S, Jaschinski T, Neugebauer EAM. Laparoscopic versus open surgery for suspected appendicitis. *Cochrane Database Syst Rev*. 2010;10:CD001546.
12. Tiwari MM, Reynoso JF, Tsang AW, et al. Comparison of outcomes of laparoscopic and open appendectomy in management of uncomplicated and complicated appendicitis. *Ann Surg*. 2011;254:927–932.
13. Wei HB, Huang JL, Zheng ZH, et al. Laparoscopic versus open appendectomy: a prospective randomized comparison. *Surg Endosc*. 2010;24:266–269.

14. Di Saverio S, Birindelli A, Kelly MD, et al. WSES Jerusalem guidelines for diagnosis and treatment of acute appendicitis. *World J Emerg Surg*. 2016;11:34.
15. Bachur RG, Hennelly K, Callahan MJ, et al. Diagnostic imaging and negative appendectomy rates in children: effects of age and gender. *Pediatrics*. 2012;129:877–884.
16. Chandrasegaram MD, Rothwell LA, An EI, et al. Pathologies of the appendix: a 10-year review of 4670 appendicectomy specimens. *ANZ J Surg*. 2012;82:844–847.
17. Graff L, Russell J, Seashore J, et al. False-negative and false-positive errors in abdominal pain evaluation failure to diagnose acute appendicitis and unnecessary surgery. *Acad Emerg Med*. 2000;7:1244–1255.
18. Goske MJ. Doctor, is a CT scan safe for my child? *Br J Radiol*. 2014;87:20130517.
19. Hernanz-Schulman M. CT and US in the diagnosis of appendicitis: an argument for CT. *Radiology*. 2010;255:3–7.
20. Rosen MP, Ding A, Blake MA, et al. ACR Appropriateness Criteria® right lower quadrant pain—suspected appendicitis. *J Am Coll Radiol*. 2011;8:749–755.
21. Borgstein PJ, Gordijn RV, Eijsbouts QAJ, et al. Acute appendicitis—a clear-cut case in men, a guessing game in young women. *Surg Endosc*. 1997;11:923–927.
22. Pham H, Devadas M, Howle J. Effect of surgical experience on the macroscopic diagnosis of appendicitis: a retrospective cohort study. *Int J Surg*. 2015;16:78–82.