Reducing the negative appendectomy rate with the laparoscopic appendicitis score; a multicenter prospective cohort and validation study

Gelpke, Koen; Hamminga, Jenneke T. H.; van Bastelaar, James J.; de Vos, Bart; Bodegom, Maarten E.; Heineman, Erik; Hofker, H. Sijbrand; El Moumni, Mostafa; Haveman, Jan Willem

Published in:
International Journal of Surgery

DOI:
10.1016/j.ijsu.2020.04.041

International Journal of Surgery 79 (2020) 257–264

Contents lists available at ScienceDirect

International Journal of Surgery

journal homepage: www.elsevier.com/locate/ijsu

Original Research

Reducing the negative appendectomy rate with the laparoscopic appendicitis score; a multicenter prospective cohort and validation study

Koen Gelpke[a], Jenneke T.H. Hamminga[a], James J. van Bastelaar[b], Bart de Vos[c], Maarten E. Bodegom[d], Erik Heineman[a], H. Sijbrand Hofker[a], Mostafa El Moumni[a], Jan Willem Haveman[a,*]

[a] Department of Surgery, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
[b] Department of Surgery, Zuyderland Medical Center, Sittard-Geleen & Heerlen, the Netherlands
[c] Department of Surgery, Wilhelmina Hospital, Assen, the Netherlands
[d] Department of Surgery, Bethesda Hospital, Hoogeveen, the Netherlands

ARTICLE INFO

Keywords:
Appendicitis
Appendectomy
Laparoscopy
Score
Evaluation

ABSTRACT

Background: Approximately nine percent of all acute appendectomies are unintentionally performed on a normal appendix. Failure of treatment (negative appendectomy or missed appendicitis) is associated with higher morbidity and mortality when compared to appendectomy for uncomplicated appendicitis. The Laparoscopic APPendicitis (LAPP) score was developed in order to systematically evaluate the appendix for the presence of inflammation. This study aims to determine whether the LAPP score reduces the negative appendectomy rate without missing appendicitis.

Methods: From September 2013 through May 2016, 322 adult patients presenting with a clinical suspicion of acute appendicitis and an indication for diagnostic laparoscopy were included and analyzed in this multicenter prospective validation study. Depending on the LAPP score, the appendix was either removed (n = 300) or left in situ (n = 22). These patients were compared to a historical control group of 584 patients treated at the same hospitals. The appendix was examined by a pathologist and the negative appendectomy rate was calculated.

Results: The negative appendectomy rate was significantly lower when the LAPP score was used (4,7% vs. 8,4%; $P = 0{,}034$). None of the patients with a negative LAPP score, in which the appendix remained in situ, developed acute appendicitis within three months. There were no significant differences in operation time, complications, or readmissions. Using the LAPP score was associated with significantly higher rates of preoperative radiological imaging (98% vs. 70%; $P < 0{,}001$). After adjusting for covariables, including radiological imaging, use of the LAPP score led to fewer treatment failures when compared to not using the LAPP score (OR: 0,48, 95% C.I. 0,251 to 0,914; $P = 0{,}025$).

Conclusion: The LAPP score is a safe and simple tool to reduce the negative appendectomy rate during laparoscopic surgery without missing cases of acute appendicitis.

1. Introduction

Acute appendicitis is the most common intra-abdominal infection and its treatment requires emergency appendectomy [1]. Despite advanced imaging techniques and several validated clinical prediction rules, correctly diagnosing a patient with acute appendicitis remains challenging. Whilst performing diagnostic laparoscopy (DL), the macroscopic evaluation of the appendix by surgeons has been proven to be inaccurate and, moreover, it does not improve with higher levels of experience [2,3]. Histopathological examination after appendectomy occasionally describes the appendix as normal (negative appendectomy). Recently published papers report a wide variation in the negative appendectomy rate, ranging from 3% to 25%, mostly affecting women [4–9]. In the Netherlands and neighboring countries, this rate lies around 9% [10–14]. Performing a negative appendectomy is potentially harmful to patients, with a morbidity rate of 10–12%, including surgical site infections, intra-abdominal abscess formation, and postoperative adhesions [15,16]. Negative appendectomy is associated with a 30-day mortality of 1% [6]. Furthermore, it is associated with longer hospitalization and higher health care costs when compared to

* Corresponding author. Department of Surgery, University Medical Center Groningen, Hanzeplein 1, 9700 RB, Groningen, the Netherlands.
E-mail address: j.w.haveman@umcg.nl (J.W. Haveman).

https://doi.org/10.1016/j.ijsu.2020.04.041
Received 4 November 2019; Received in revised form 5 April 2020; Accepted 15 April 2020
Available online 06 May 2020
1743-9191/

appendectomies performed for cases of uncomplicated appendicitis [4,5]. The short-term mortality rate of a negative appendectomy is increased fourfold when compared to an appendectomy for uncomplicated appendicitis [8]. The mortality rate after negative appendectomy remains elevated for years after surgery, even when lacking an alternative diagnosis and when no other additional intra-abdominal operations were performed [17]. It seems doubtful to attribute these findings solely to undiagnosed co-morbidities.

Delayed time to appendectomy in cases with acute appendicitis is associated with a higher appendiceal perforation rate [18–20]. Conversely, a negative DL, in which the appendix remains in situ, is associated with a very low morbidity rate of 0.06% [21].

The surgical removal of a macroscopically normal appendix during laparoscopy for clinically suspected appendicitis is considered unnecessary, even when no other intra-abdominal pathology is found [22,23]. Hence, the guideline of the Dutch Society of Surgeons states that a normal-looking appendix should be left in situ [24].

In a prospective pilot study conducted on 134 patients in 2012, the Laparoscopic Appendicitis (LAPP) score was introduced. This was developed in order to macroscopically evaluate the appendix during DL [25]. The LAPP score consists of five items regarding the presence of appendiceal inflammation (Table 1). When all five items are negative, the appendix can safely be left in situ. In any other case, the appendix should be removed. In this pilot study, use of the LAPP score would have led to a positive predictive value of 99% and a negative predictive value of 100%. The present multicenter prospective validation study aims to determine whether the LAPP score can reduce the negative appendectomy rate without increasing the rate of missed appendicitis.

2. Methods

2.1. Patients

From September 2013 through May 2016, adult patients presenting with a clinical suspicion of acute appendicitis with an indication for DL in one of five public hospitals (one university and four non-university) across the Netherlands were recruited. These patients were compared to a historical control group that was treated in the same hospitals from January 2011 through December 2012. In the university hospital, patients in this control group underwent surgery between January 2008 and June 2009 as preparations for the pilot study on the LAPP score started hereafter in this hospital. The study was registered at www.clinicaltrials.gov (NCT02029781). Ethical approval was obtained from the local Medical Ethics Committee, reference number: NL43820.042.13.

Exclusion criteria were as follows: younger than 18 years, interval appendectomy (after prior conservative treatment), incidental appendectomy (during 'non-appendiceal' surgery), primarily chosen for an appendectomy with open approach, inability to determine the LAPP-score due to insufficient laparoscopic exposure, and patients that were unable or unwilling to give informed consent. Insufficient laparoscopic exposure was defined as the inability to completely assess the appendix for the presence of all five items of the LAPP score (Table 1).

Table 1
The laparoscopic appendicitis (LAPP) score [25].

Item
1. Is there a perforation or necrosis of the appendix?
2. Is the appendix thickened?
3. Is the mesentery of the appendix thickened?
4. Is there presence of injected vessels on the serosa of the appendix?
5. Are there any adhesions around the appendix?

An appendectomy should be performed when the LAPP score is positive (one or more questions answered with yes). The appendix should be left in situ when the LAPP score is negative (all questions answered with no).

The study was approved by the local institutional ethical committees. All data were extracted from the patients' electronic medical records in the participating hospitals. For the intervention group, data were prospectively collected after written informed consent had been obtained. In the control group, data were retrospectively collected from patients that had undergone an appendectomy during DL. The pre-operative work-up for each patient was not standardized in this study. Each hospital followed their own protocol, all of which are based on the national guideline on appendicitis. In all patients, a thorough history was obtained and clinical examination, standard blood tests with white blood cell count, C-reactive protein (CRP), and urinary analysis were performed. If necessary, radiological imaging by abdominal ultrasound (US), computed tomography (CT), or Magnetic Resonance Imaging (MRI) was undertaken. The modified Alvarado score [26,27] was calculated. Postoperative complications were recorded according to the Clavien-Dindo classification [28].

The LAPP score was assessed and recorded intra-operatively by the surgeon immediately after exposing the appendix. Depending on the calculated score, the surgeon either proceeded to perform an appendectomy or the procedure was terminated without removing the appendix (Table 1). Care was taken to list the LAPP score in the patients' operating reports.

This study was reported in line with the STROCSS criteria [29].

2.2. Endpoints

The primary endpoint of the study was the negative appendectomy rate. The gold standard was histological examination of the appendix by a pathologist after appendectomy. This was performed in all cases in which an appendectomy had been performed. At least three slices of the appendix were examined: two transverse transections, one at the base and one approximately halfway, and a longitudinal transection through the tip of the appendix. Histological inflammation of the appendix was defined by the presence of at least local ulceration with infiltration of polymorphonuclear neutrophils in the mucosa or submucosa (endo-appendicitis), either or not in combination with transmural inflammation, necrosis, perforation, and/or peri-appendicular inflammation [30,31]. In cases where inflammation without mucosal ulceration was found, the contralateral part of the appendix and more transverse slices were examined. A normal appendix was defined by the absence of inflammation or other rare pathology like a neuroendocrine tumor (NET), adenocarcinoma, endometriosis, Crohn's disease, or parasitic infection. These diagnoses were scored as a separate entity. Fibrous obliteration and lymphoid hyperplasia were considered to represent a histologically normal appendix in this context [32].

The secondary endpoint was the number of missed appendicitis diagnoses; a maximum of 1% was considered acceptable in daily practice. Missed appendicitis was defined as a confirmed appendicitis or appendicular infiltrate or abscess on imaging or during re-operation within three months of initial laparoscopy when the appendix was intentionally left in situ (LAPP score < 1).

Failure of treatment was defined either as a negative appendectomy or a missed diagnosis of appendicitis.

2.3. Sample size

Reducing the negative appendectomy rate with 50% or more was considered a clinically relevant outcome. The sample size calculation was performed according to the methods by William D. Dupont and Walton D. Plummer (PS Power and Sample Size Calculations). For the sample size calculation, we performed a retrospective cohort analysis in 2008 and 2009 from five hospitals in the Netherlands (data not published). During this period, 843 patients underwent laparoscopic surgery for suspected acute appendicitis; 9% of the removed appendices turned out to be normal. In order to halve the negative appendectomy rate with a power of 0.8 and an alpha of 0.05, we needed to include 490 patients in both groups. Our previous pilot study and the retrospective data showed that approximately 20% of the included patients would be excluded because of the inability to determine the LAPP-score due to insufficient laparoscopic exposure. Therefore, 613 patients in the intervention group were required.

Despite previous commitments, not all of the initial hospitals used in the sample size calculation were willing to participate in the study. Three new hospitals agreed to participate, and the data extracted from the control group concerned patients that were operated on in the participating hospitals.

After almost three years of prospectively recruiting patients, the study was stopped with the inclusion of 331 patients in the intervention group. Accrual was prematurely terminated due to a lower than expected inclusion rate in the participating hospitals.

2.4. Statistical analysis

Patient data were compared between the intervention group and the historical control group. Normally distributed variables are presented as means with standard deviations and were tested using the independent samples T test. Not normally distributed variables are presented as medians with interquartile ranges and were tested using the Mann-Whitney *U* test. Categorical data are presented as absolute numbers with percentages; differences were tested with Pearson's Chi-square test or Fisher's exact test. Multiple logistic regression analyses were performed to determine the relationship between the use of the LAPP score and failure of treatment. We adjusted for the following covariables: sex, age, temperature, leukocyte count, CRP, Modified Alvarado Score ≤ 3, and radiological imaging. Odds ratios are shown with 95% confidence intervals. Outcomes with a *P*-value less than 0,05 were considered to be statistically significant. All statistical analyses were performed using IBM SPSS Statistics for Windows (Released 2013, Version 23.0, Armonk, NY: IBM Corp.)

3. Results

3.1. Patients

In the intervention group, 331 patients with a clinical suspicion of acute appendicitis and an indication for DL were included prior to surgery. Eight patients were excluded from analysis after intra-operative conversion due to insufficient laparoscopic exposure of the appendix. One patient was excluded from analysis due to deviation from the study protocol (appendix left in situ after positive LAPP score). The case reported an inflammatory reaction of the appendix secondary to severe infiltration of the caecum (as a result of diverticulitis).

A total number of 322 patients were included for statistical analysis, of which 300 had a positive LAPP score and 22 had a negative LAPP score.

In the control group, 662 patients were identified that had undergone DL after a clinical suspicion of acute appendicitis. After exclusion of 78 patients due to insufficient laparoscopic exposure resulting in conversion to an open appendectomy, a total number of 584 patients remained for statistical analysis (Fig. 1).

3.2. Baseline characteristics

Baseline characteristics of both groups are presented in Table 2. There was no significant difference in sex, age, temperature, leukocyte count, and CRP and ASA classification. The modified Alvarado score was significantly lower in the intervention group: 6 (IQR 5–7) vs. 7 (IQR 6–8); $P < 0{,}001$. There was no significant difference in patients with a score of 3 or lower, but there was a significantly lower number of patients with a score of 7 or higher in the intervention group (44,4% vs. 55,5%; $P = 0{,}002$). There was a significantly higher use of preoperative radiological imaging (US, CT, MRI, or a combination) in the intervention group (98,4% vs. 70,0%; $P < 0{,}001$).

3.3. Intra-operative and post-operative characteristics

The intra-operative and post-operative characteristics are presented in Table 3. The median LAPP score in the intervention group was 4 (IQR 3–4). There were 22 patients (6,8%) with a negative LAPP score, seven of which were men. In all 22 patients with a negative LAPP score, the appendix remained in situ. None of these patients developed appendicitis in the following three months. One of the patients required a laparoscopic cholecystectomy due to suspected symptomatic cholecystolithiasis; the histology report however was uneventful. In six cases with a negative LAPP score, another abdominal condition was identified (epiploic appendagitis, omental necrosis, endometriosis, terminal ileitis, gastric ulcer perforation, and colonic perforation). In the other sixteen cases, no intra-operative diagnosis was found. There were 300 patients (93,2%) with a positive LAPP score, and an appendectomy was performed in all of these patients.

There was no significant difference in operating time between both groups. The median length of stay during primary admission was significantly lower in the intervention group (1 day, IQR 1–3) compared to the control group (2 days, IQR 1–4; $P = 0{,}021$). The number of patients in whom one or more complications occurred during admission or readmission was not significantly different between both groups. An overview of the total number of complications with the corresponding grade is presented in Table 4. The rate of readmission due to complications did not significantly differ, nor did the rate of readmission due to other reasons (e.g. pain with unknown cause, elective surgery for a neoplasm, related to pre-existent conditions). Furthermore, there was no significant difference in the admission rate to the Intensive Care Unit (ICU). One patient in our cohort died within three months after surgery. This patient was included in the control group and died during an ICU admission due to a supraventricular tachycardia that could not be successfully converted. Upon the patient's request, no further treatment was instituted, and death ensued.

3.4. Histopathological examination

The number of normal appendices was significantly lower in the intervention group (4,7% vs. 8,3%; $P = 0{,}037$) (Table 5). The number of patients with an inflamed appendix was significantly higher in the intervention group (94,3% vs. 90,4%; $P = 0{,}045$). In the intervention group, one case presented with both an inflamed appendix and an NET, and one case presented with both an inflamed appendix and an adenocarcinoma. In the control group, there was one case with both an inflamed appendix and endometriosis, and one case with both an inflamed appendix and pinworms (Enterobius vermicularis). There were no significant differences with regard to other pathologies found in the appendix. In both groups there were three patients with an NET, adenocarcinoma, or both. Additionally, in the control group there was one patient diagnosed with Crohn's disease, one patient with an appendicular diverticulitis, and one patient with appendicular endometriosis.

In the intervention group, all 14 patients (100,0%) with a normal appendix after appendectomy had undergone preoperative imaging. All 14 cases had been considered radiologically positive for acute appendicitis. In the control group, 35 out of 50 patients (70%) with a negative appendectomy had undergone preoperative imaging, of which 23 appeared radiologically positive. A notable finding was that in the intervention group, nine out of these 14 patients (64,3%) with a normal appendix after appendectomy were male, compared to 16 out of 50 patients (32,0%) in the control group.

3.5. Application of the LAPP score

In the present study, application of the LAPP score resulted in a sensitivity of 100% and a specificity of 61,1%, with a positive and

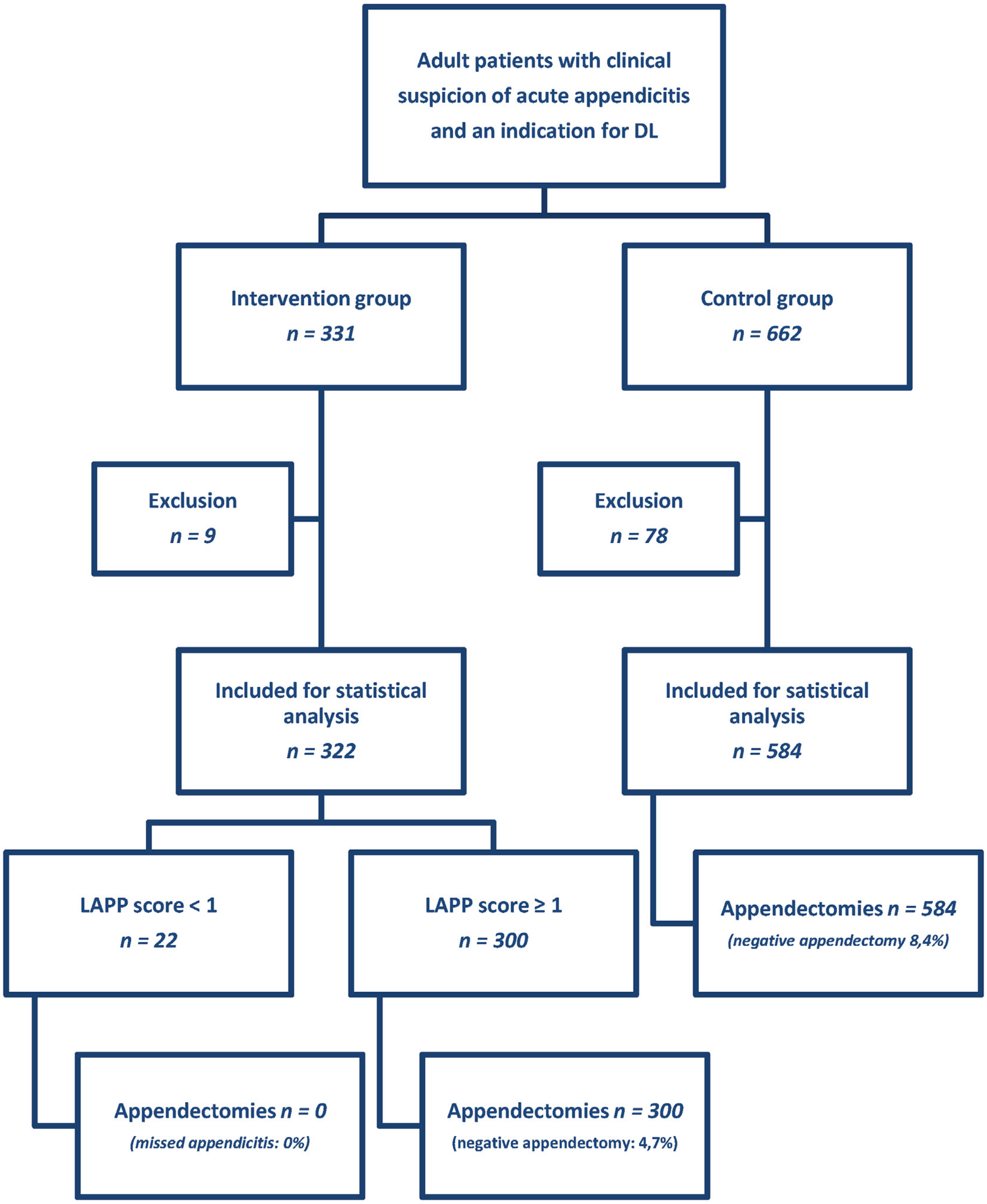

Fig. 1. Flowchart of patients included in the study.
Eight patients in the intervention group and 78 patients in the control group were excluded from analysis after intra-operative conversion due to insufficient laparoscopic exposure of the appendix. One patient in the intervention group was excluded from analysis due to deviation from the study protocol.

negative predictive value of 95,3% and 100%, respectively (Table 6).

Logistic regression analysis showed that the LAPP score was associated with significantly less failure of treatment when compared to not using the LAPP score (OR: 0,49, 95% C.I. 0,264 to 0,893; $P = 0{,}020$). After adjusting for covariables, the effect of the LAPP score on failure of treatment remained significant (OR: 0,48, 95% C.I. 0,251 to 0,914; $P = 0{,}025$).

This resulted in the odds of failure of treatment being 2.08 (95% CI: 1.09–3.98) times higher when the LAPP score was not used.

4. Discussion

4.1. Main findings

This study shows that use of the LAPP score, a tool developed to quantify the laparoscopic assessment of the appendix for inflammation, reduced the negative appendectomy rate from 8,4% to 4,7% ($P = 0{,}034$), without missing an episode of appendicitis. The LAPP score has a positive predictive value of 95,3% and a negative predictive value of 100%. The baseline characteristics of both groups in this study were not completely comparable. A significantly higher number of patients underwent preoperative radiological imaging in the intervention group (98,4% vs. 70,0%; $P < 0{,}001$). Ultrasound was most commonly used in both groups. When adjusted for covariables –

Table 2
Baseline characteristics presented per group.

	Intervention group (n = 322)	Control group (n = 584)	*P*-value
Sex, male	147 (45,7%)	265 (45,4%)	0,936[a]
Age, years	41,3 (± 17,0 SD)	39,3 (± 16,1 SD)	0,074[b]
Temperature, ⁰ C	37,3 (± 0,78 SD)	37,3 (± 0, 76 SD)	0,698[b]
Leukocyte count, 10^9/L	14,2 (± 4,7 SD)	14,6 (± 4,3 SD)	0,161[b]
C-reactive protein, mg/L	43 (IQR 13–93)	37 (IQR 13–84)	0,401[c]
Modified Alvarado score	6 (IQR 5–7)	7 (IQR 6–8)	**< 0,001**[c]
Score ≤ 3	22 (6,8%)	27 (4,6%)	0,159[a]
Score ≥ 7	143 (44,4%)	321 (55,0%)	**0,002**[a]
ASA classification:			0,291[a]
I	227 (70,5%)	438 (75,0%)	
II	88 (27,3%)	131 (22,4%)	
III	7 (2,2%)	13 (2,2%)	
IV	0 (0,0%)	2 (0,3%)	
Radiological imaging	317 (98,4%)	409 (70,0%)	**< 0,001**[a]
US	272 (84,5%)	338 (57,9%)	**< 0,001**[a]
CT	115 (35,7%)	264 (29,1%)	**0,001**[a]
MRI	3 (0,9%)	4 (0,7%)	0,685[a]

[a] Pearson's Chi-square test or Fisher's exact test.
[b] Independent samples T test.
[c] Mann-Whitney *U* test.

Table 3
Intra-operative and post-operative characteristics presented per group.

	Intervention group (n = 322)	Control group (n = 584)	*P*-value
LAPP score total	4 (IQR 3–4)	–	N/A
LAPP score < 1	22 (6,8%)	–	N/A
Operation time, min	52,5 (± 22,0 SD)	54,8 (± 22,5 SD)	0,082[b]
Length of stay, days	1 (IQR 1–3)	2 (IQR 1–4)	**0,019** [c]
Complicated cases	37 (11,5%)	81 (13,9%)	0,308[a]
Readmission (complication)	15 (4,7%)	47 (8,0%)	0,053[a]
Readmission (other)	8 (2,5%)	12 (2,1%)	0,674[a]
ICU admission	0 (0,0%)	5 (0,9%)	0,096[a]

[a] Pearson's Chi-square test or Fisher's exact test.
[b] Independent samples < SUP > T < /SUP > [test].
[c] Mann-Whitney < SUP > U < /SUP > [test].

Table 4
Total number of complications presented per group, according to the Clavien-Dindo classification [28].

	Intervention group	Control group	Total
Grade I	5 (1,6%)	13 (2,2%)	18 (2,0%)
Grade II	22 (6,9%)	19 (3,3%)	41 (4,5%)
Grade III	13 (4,0%)	49 (8,4)	62 (6,9%)
Grade IV	0 (0,0%)	4 (0,7%)	4 (0,4%)
Grade V	0 (0,0%)	1 (0,3%)	1 (0,1%)
Total	40 (12,5%)	86 (14,7%)	126 (13,9%)

Grade I: Any deviation from the normal postoperative course without the need for pharmacological treatment.
Grade II: Complication requiring pharmalogical treatment.
Grade III: Complication requiring surgical, endoscopic or radiological intervention.
Grade IV: Life-threatening complications requiring ICU management.
Grade V: Death of the patient.
In a single patient, multpiple separate complications could occur. Only when one complication leads directly to another, it is accounted as a single complication belonging to the highest grade within the classification.

Table 5
Diagnosis after histopathological examination presented per group.

Diagnosis	Positive LAPP score (n = 300)	Control group (n = 584)	*P*-value
Normal appendix	14 (4,7%)	50 (8,6%)	**0,034**[a]
Inflamed appendix	283 (94,3%)*	528 (90,4%)†	**0,045**[a]
Other pathology	3 (1,0%)	6 (1,0%)	1,000[a]

* Including two patients with both inflammation and a neoplasm.
† Including one patient with both inflammation and endometriosis and one patient with both inflammation and parasites.
[a] Pearson's Chi-square test or Fisher's exact test.

including radiological imaging – use of the LAPP score was associated with significantly fewer treatment failures (negative appendectomy or missed appendicitis) when compared to not using the LAPP score (OR: 0,48, 95% C.I. 0,251 to 0,914; P = 0,025).

4.2. Strengths and weaknesses

This study is the first to prospectively evaluate the effect of the LAPP score. The high number of included patients, combined with the multicenter approach, assures that the results are representative of general practice. It has been shown that surgeons are generally not very accurate in evaluating the macroscopic appearance of the appendix during laparoscopy [2,3]. The LAPP score has proven to be simple in use, readily available, free of charge, and more importantly, it does not increase the rate of missed appendicitis. Gomes et al. [33] proposed a laparoscopic grading system in order to allow uniform clinical stratification of patients with acute appendicitis. They found a laparoscopic sensitivity and specificity in the diagnosis of acute appendicitis that was similar to our study. They did not, however, provide exact criteria to be used in deciding whether or not to remove the appendix. To our knowledge, there are no other studies describing a laparoscopic score in order to reduce the negative appendectomy rate.

A limitation of this study is the lack of uniformity in the groups. In theory, a randomized controlled trial would have been more suitable. However, we felt that creating a situation in which surgeons had to use the LAPP score in one patient and not in the next would not be feasible. Furthermore, after publication of the results of the pilot study, surgeons may have started using the LAPP score, changing the way that they macroscopically evaluated the appendix.

In the intervention group, more patients underwent preoperative radiological imaging. In the Netherlands it has become increasingly common to perform imaging since the introduction of the Dutch Guideline on Appendicitis in 2010 [24]. Previous studies suggested that an increase in radiological imaging leads to a decrease in the negative appendectomy rate [34]. In countries like the United States and Korea, where nearly all patients undergo preoperative CT when suspected of acute appendicitis, the negative appendectomy rate has dropped to as low as 3–4% [4,5,35]. Introduction of CT in the United States did in fact drastically reduce the negative appendectomy rate, however, extension of use did not result in a further decline [36]. Furthermore, it has been shown that a decrease in the percentage of negative appendectomies is mainly due to the use of CT; US does not seem to have a significant effect [37–39]. In the Netherlands, it is less common to use preoperative CT as a primary imaging modality in suspected acute appendicitis. US is readily available and affordable and is therefore generally regarded as imaging of choice. A study by Schok et al. states that negative appendectomy rates have not declined in the Netherlands since the introduction of the Dutch Guideline on Appendicitis [11]. Boonstra et al., in contrast, note that an increase in preoperative imaging has led to a decrease in negative appendectomies in their hospital from 19% in 2008 to 5% in 2011 [40]. Both studies included children and part of the appendectomies were performed by a primarily open technique. In the study by Boonstra et al. a negative DL was also considered to be a

Table 6
Results on the Laparoscopic Appendicitis Score (LAPP) score.

	Pathological appendix	Normal appendix	Total
Positive LAPP score	286	14	300
Negative LAPP score	0	22	22
Total	286	36	322

Sensitivity (proportion positive LAPP scores among total pahtological appendices) and specificity (porportion negative LAPP scores among total normal appendices) of the LAPP score in this study are 100% and 61,1% respectively. Positive predictive value (proportion pathological appendices among total positive LAPP scores) is 95,3% and negative predictive value (proportion of normal appendices among total negative LAPP scores) is 100%.

negative appendectomy. In the present study, the use of preoperative imaging in the intervention group was much higher compared to the control group. Consequently, there might have been a selection bias due to more diagnostic certainty in the intervention group. Therefore, the reduction in the negative appendectomy rate may have been caused partially by an increase in radiological imaging. However, our multiple logistic regression analyses showed that the increase in imaging did not change the effect of the LAPP score on failure of treatment.

Further research is required to define the exact effect of the LAPP score in areas with variable negative appendectomy rates and different use of preoperative imaging modalities.

Another consequence of the lack of uniformity is the significantly lower modified Alvarado score in the intervention group. With equal values in both groups, one would hypothetically expect an even lower negative appendectomy rate in the intervention group or a higher negative appendectomy rate in the control group; either way resulting in a bigger difference between both groups. The original Alvarado score, however, is validated as a diagnostic tool to rule out appendicitis in patients with a low score [41–43] and the modified Alvarado has shown to have even lower specificity [44]. There was no significant difference in the number of patients with a score of 3 or less.

In order to calculate the negative appendectomy rate, we relied on the histopathological examination reports. Due to the multicenter design, the reports were made by different pathologists, which might have led to inter-observer variation. Moreover, the histological finding of endo-appendicitis was not regarded as a normal finding. This diagnosis and its clinical significance are controversial. Some authors believe it is an early form of appendicitis while others see it as a self-limiting disease that cures spontaneously without surgical intervention. In the context of endo-appendicitis, another possibility is that intramural inflammation was missed upon examination of the transected appendix [32,45].

4.3. Clinical relevance

Reducing the negative appendectomy rate is relevant. A negative appendectomy is associated with higher morbidity and mortality, longer hospital stay, and higher health care costs when compared to appendectomies performed for uncomplicated appendicitis [4–6,15,16]. In our study, 22 patients (6,8%) in the intervention group did not undergo a negative appendectomy due to a negative LAPP score. We cannot determine whether or not the appendix would have been removed if the LAPP score would not have been used. Only six of these patients, however, showed another intra-operative diagnosis that could explain their symptoms. We presume that it is likely that in the other 16 cases (5,0%), a negative appendectomy was prevented by using the LAPP score. This supposition is strengthened by a study conducted by Sørensen et al. [23] in which 271 patients were selected with a negative DL after presenting with clinically suspected appendicitis. Criteria equal to those of the LAPP score were used during surgery. Likewise, no other intra-abdominal pathology was found, and all appendices were left in situ. After follow-up, only one case of missed appendicitis was identified, for which subsequent laparoscopic appendectomy ensued.

Furthermore, it is known that negative appendectomy occurs more frequently in women, which was also a finding in our control group. Surprisingly, the majority of patients in the intervention group with a negative appendectomy comprised males. In contrast, the majority of patients where the appendix safely stayed in situ was female, suggesting that the LAPP score may be especially useful in preventing negative appendectomies in women.

In the absence of a standardized and validated procedure to macroscopically evaluate the appendix, both the 'World Society of Emergency Surgeons' and the 'European Association of Endoscopic Surgery' still recommend the removal of a normal looking appendix when there is no other intra-abdominal pathology found during laparoscopy [46,47]. The main reason for removing an apparently normal appendix is the risk of missing an inflamed appendix. Our study shows that an appendix that is regarded uneventful during laparoscopy can be safely left in situ when the LAPP score is used.

In conclusion, the LAPP score is a safe and simple tool that can help surgeons reduce the negative appendectomy rate during laparoscopy without increasing the rate of missed appendicitis.

Ethical approval

Medical Ethics Committee – University Medical Center Groningen (METc UMCG).

Reference number: NL43820.042.13.

Sources of funding

None for all authors.

Trial registry number

1. Name of the registry: clinicaltrials.gov

The Laparoscopic Appendicitis Score; a Multicenter Validation Study (APPLE).

2. Unique Identifying number or registration ID:NCT02029781
3. Hyperlink to the registration (must be publicly accessible): https://clinicaltrials.gov/ct2/show/NCT02029781?term=NCT02029781

Guarantor

Jan Willem Haveman.
Koen Gelpke.

Data statement

To foster transparency, our research data is available to access upon request.

Disclosures

Authors Gelpke, Hamminga, Van Bastelaar, De Vos, Bodegom, Heineman, Hofker, El Moumni and Haveman have no conflicts of interest or financial ties to disclose.

Provenance and peer review

Not commissioned, externally peer-reviewed.

CRediT authorship contribution statement

Koen Gelpke: Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Visualization, Validation. **Jenneke T.H. Hamminga:** Conceptualization, Methodology. **James J. van Bastelaar:** Investigation, Resources. **Bart de Vos:** Investigation, Resources. **Maarten E. Bodegom:** Investigation, Resources. **Erik Heineman:** Conceptualization, Methodology, Investigation, Resources. **H. Sijbrand Hofker:** Conceptualization, Methodology, Investigation, Resources. **Mostafa El Moumni:** Formal analysis, Resources. **Jan Willem Haveman:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Supervision, Project administration.

Declaration of competing interest

None for all authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijsu.2020.04.041.

References

[1] M. Sartelli, A. Chichom-Mefire, F.M. Labricciosa, T. Hardcastle, F.M. Abu-Zidan, A.K. Adesunkanmi, et al., The management of intra-abdominal infections from a global perspective: 2017 WSES guidelines for management of intra-abdominal infections, World J. Emerg. Surg. 12 (2017) 29, https://doi.org/10.1186/s13017-017-0141-6.

[2] S. Strong, N. Blencowe, A. Bhangu, National Surgical Research Collaborative, How good are surgeons at identifying appendicitis? Results from a multi-centre cohort study, Int. J. Surg. 15 (2015) 107–112, https://doi.org/10.1016/j.ijsu.2015.01.032.

[3] H. Pham, M. Devadas, J. Howle, Effect of surgical experience on the macroscopic diagnosis of appendicitis: a retrospective cohort study, Int. J. Surg. 16 (2015) 78–82, https://doi.org/10.1016/j.ijsu.2015.02.019.

[4] K. Mock, Y. Lu, S. Friedlander, D.Y. Kim, S.L. Lee, Misdiagnosing adult appendicitis: clinical, cost, and socioeconomic implications of negative appendectomy, Am. J. Surg. 212 (2016) 1076–1082, https://doi.org/10.1016/j.amjsurg.2016.09.005.

[5] B.G. Jeon, Predictive factors and outcomes of negative appendectomy, Am. J. Surg. 213 (2017) 731–738, https://doi.org/10.1016/j.amjsurg.2016.05.020.

[6] S.A. Seetahal, O.B. Bolorunduro, T.C. Sookdeo, T.A. Oyetunji, W.R. Greene, W. Frederick, et al., Negative appendectomy: a 10-year review of a nationally representative sample, Am. J. Surg. 201 (2011) 433–437, https://doi.org/10.1016/j.amjsurg.2010.10.009.

[7] S. Charfi, A. Sellami, A. Affes, K. Yaich, R. Mzali, T.S. Boudawara, Histopathological findings in appendectomy specimens: a study of 24,697 cases, Int. J. Colorectal Dis. 29 (2014) 1009–1012, https://doi.org/10.1007/s00384-014-1934-7.

[8] S. Kotaluoto, M. Ukkonen, S.L. Pauniaho, M. Helminen, J. Sand, T. Rantanen, Mortality related to appendectomy; a population based analysis over two decades in Finland, World J. Surg. 41 (2017) 64–69, https://doi.org/10.1007/s00268-016-3688-6.

[9] M.D. Chandrasegaram, L.A. Rothwell, E.I. An, R.J. Miller, Pathologies of the appendix: a 10-year review of 4670 appendicectomy specimens, ANZ J. Surg. 82 (2012) 844–847, https://doi.org/10.1111/j.1445-2197.2012.06185.x.

[10] T. Slotboom, J.T. Hamminga, H.S. Hofker, E. Heineman, J.W. Haveman, Intraoperative motive for performing a laparoscopic appendectomy on a postoperative histological proven normal appendix, Scand. J. Surg. 103 (2014) 245–248, https://doi.org/10.1177/1457496913519771.

[11] T. Schok, P.C. Simons, M.L. Janssen-Heijnen, N.A. Peters, J.L. Konsten, Prospective evaluation of the added value of imaging within the Dutch National Diagnostic Appendicitis Guideline - do we forget our clinical eye? Dig. Surg. 31 (2015) 436–443, https://doi.org/10.1159/000369587.

[12] J. Liese, T.M. Halbinger, F. Ulrich, W.O. Bechstein, C.W. Strey, Appendicitis-the balance between cost effectiveness and safety remains challenging, Langenbeck's Arch. Surg. 399 (2014) 493–501, https://doi.org/10.1007/s00423-014-1179-5.

[13] M.A. Abdelhalim, J.D. Stuart, G.A. Nicholson, Augmenting the decision making process in acute appendicitis: a retrospective cohort study, Int. J. Surg. 17 (2015) 5–9, https://doi.org/10.1016/j.ijsu.2015.03.004.

[14] A.O. Omiyale, S. Adjepong, Histopathological correlations of appendectomies: a clinical audit of a single center, Ann. Transl. Med. 3 (2015) 119, https://doi.org/10.3978/j.issn.2305-5839.2015.05.02.

[15] M. Lee, T. Paavana, F. Mazari, T.R. Wilson, The morbidity of negative appendicectomy, Ann. R. Coll. Surg. Engl. 96 (2014) 517–520, https://doi.org/10.1308/003588414X13946184903801.

[16] National Surgical Research Collaborative, Multicentre observational study of performance variation in provision and outcome of emergency appendicectomy, Br. J. Surg. 100 (2013) 1240–1252, https://doi.org/10.1002/bjs.9201.

[17] R.E. Andersson, Short and long-term mortality after appendectomy in Sweden 1987 to 2006. Influence of appendectomy diagnosis, sex, age, co-morbidity, surgical method, hospital volume, and time period. A national population-based cohort study, World J. Surg. 37 (2013) 974–981, https://doi.org/10.1007/s00268-012-1856-x.

[18] J. Li, R. Xu, D. Hu, Y. Zhang, T. Gong, X. Wu, Effect of delay to operation on outcomes in patients with acute appendicitis: a systematic review and meta-analysis, J. Gastrointest. Surg. 23 (2019) 210–223, https://doi.org/10.1007/s11605-018-3866-y.

[19] B.A. Fair, J.C. Kubasiak, I. Janssen, J.A. Myers, K.W. Millikan, D.J. Deziel, et al., The impact of operative timing on outcomes of appendicitis: a National Surgical Quality Improvement Project analysis, Am. J. Surg. 209 (2015) 498–502, https://doi.org/10.1016/j.amjsurg.2014.10.013.

[20] M. Elniel, J. Grainger, E.J. Nevins, N. Misra, P. Skaife, 72 h is the time critical point to operate in acute appendicitis, J. Gastrointest. Surg. 22 (2018) 310–315, https://doi.org/10.1007/s11605-017-3614-8.

[21] P. Harkki-Siren, J. Sjoberg, T. Kurki, Major complications of laparoscopy: a follow-up Finnish study, Obstet. Gynecol. 94 (1999) 94–98.

[22] W.T. Van den Broek, A.B. Bijnen, P. de Ruiter, D.J. Gouma, A normal appendix found during diagnostic laparoscopy should not be removed, Br. J. Surg. 88 (2001) 251–254, https://doi.org/10.1046/j.1365-2168.2001.01668.x.

[23] A.K. Sørensen, A. Bang-Nielsen, K. Levic-Souzani, H.C. Pommergaard, A.B. Jørgensen, M.B. Tolstrup, et al., Readmission and reoperation rates following negative diagnostic laparoscopy for clinically suspected appendicitis: the "normal" appendix should not be removed - a retrospective cohort study, Int. J. Surg. 64 (2019) 1–4, https://doi.org/10.1016/j.ijsu.2019.02.001.

[24] O.J. Bakker, P.M. Go, J.B. Puylaert, G. Kazemier, H.A. Heij, [Guideline on diagnosis and treatment of acute appendicitis: imaging prior to appendectomy is recommended], Ned. Tijdschr. Geneeskd. 154 (2010) A303.

[25] J.T. Hamminga, H.S. Hofker, P.M. Broens, P.M. Kluin, E. Heineman, J.W. Haveman, Evaluation of the appendix during diagnostic laparoscopy, the laparoscopic appendicitis score: a pilot study, Surg. Endosc. 27 (2013) 1594–1600, https://doi.org/10.1007/s00464-012-2634-4.

[26] A. Alvarado, A practical score for the early diagnosis of acute appendicitis, Ann. Emerg. Med. 15 (1986) 557–564, https://doi.org/10.1016/s0196-0644(86)80993-3.

[27] M. Kalan, D. Talbot, W.J. Cunliffe, A.J. Rich, Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study, Ann. R. Coll. Surg. Engl. 76 (1994) 418–419.

[28] D. Dindo, N. Demartines, P.A. Clavien, Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey, Ann. Surg. 240 (2004) 205–213, https://doi.org/10.1097/01.sla.0000133083.54934.ae.

[29] R. Agha, A. Abdall-Razak, E. Crossley, N. Dowlut, C. Iosifidis, G. Mathew, et al., STROCSS 2019 Guideline: strengthening the reporting of cohort studies in surgery, Int. J. Surg. 72 (2019) 156–165, https://doi.org/10.1016/j.ijsu.2019.11.002.

[30] N.J. Carr, The pathology of acute appendicitis, Ann. Diagn. Pathol. 4 (2000) 46–58, https://doi.org/10.1053/adpa.2000.0046.

[31] R.H. Riddell, D. Jain, C.N. Bernstein, S. Guha Lewin, Weinstein, and Riddell's Gastrointestinal Pathology and its Clinical Implications, second ed., Wolters Kluwer Health, Philadelphia, 2014.

[32] J.R. Goldblum, L.W. Lamps, J.K. McKenney, J.L. Myers, L.V. Ackerman, J. Rosai. Rosai and Ackerman's Surgical Pathology, eleventh ed., Elsevier, Philadelphia, 2018.

[33] C.A. Gomes, T.A. Nunes, J.M. Fonseca Chebli, C.S. Junior, C.C. Gomes, Laparoscopy grading system of acute appendicitis: new insight for future trials, Surg. Laparosc. Endosc. Percutaneous Tech. 22 (2012) 463–466, https://doi.org/10.1097/SLE.0b013e318262edf1.

[34] S.A. Kabir, S.I. Kabir, R. Sun, S. Jafferbhoy, A. Karim, How to diagnose an acutely inflamed appendix; a systematic review of the latest evidence, Int. J. Surg. 40 (2017) 155–162, https://doi.org/10.1016/j.ijsu.2017.03.013.

[35] F.T. Drake, N.E. Mottey, A.A. Castelli, M.G. Florence, M.G. Johnson, S.R. Steele, et al., Time-of-day and appendicitis: impact on management and outcomes, Surgery 161 (2017) 405–414, https://doi.org/10.1016/j.surg.2016.06.052.

[36] J.G. Mariadason, W.N. Wang, M.K. Wallack, A. Belmonte, H. Matari, Negative appendicectomy rate as a quality metric in the management of appendicitis: impact of computed tomography, Alvarado score and the definition of negative appendicectomy, Ann. R. Coll. Surg. Engl. 94 (2012) 395–401, https://doi.org/10.1308/003588412X13171221592131.

[37] SCOAP Collaborative, J. Cuschieri, M. Florence, D.R. Flum, G.J. Jurkovich, P. Lin, et al., Negative appendectomy and imaging accuracy in the Washington state surgical care and outcomes assessment program, Ann. Surg. 248 (2008) 557–563, https://doi.org/10.1097/SLA.0b013e318187aeca.

[38] F.T. Drake, M.G. Florence, M.G. Johnson, G.J. Jurkovich, S. Kwon, Z. Schmidt, et al., Progress in the diagnosis of appendicitis: a report from Washington state's surgical care and outcomes assessment program, Ann. Surg. 256 (2012) 586–594, https://doi.org/10.1097/SLA.0b013e31826a9602.

[39] J. Lee, Y. Ko, S. Ahn, J.H. Park, H.J. Kim, S.S. Hwang, et al., Comparison of US and CT on the effect on negative appendectomy and appendiceal perforation in adolescents and adults: a post-hoc analysis using propensity-score methods, J. Clin. Ultrasound 44 (2016) 401–410, https://doi.org/10.1002/jcu.22351.

[40] P.A. Boonstra, R.N. van Veen, H.B. Stockmann, Less negative appendectomies due to imaging in patients with suspected appendicitis, Surg. Endosc. 29 (2015) 2365–2370, https://doi.org/10.1007/s00464-014-3963-2.

[41] R. Ohle, F. O'Reilly, K.K. O'Brien, T. Fahey, B.D. Dimitrov, The Alvarado score for predicting acute appendicitis: a systematic review, BMC Med. 9 (2011), https://doi.org/10.1186/1741-7015-9-139 139-139.

[42] I.J. Dahabreh, G.P. Adam, C.W. Halladay, D.W. Steele, L.A. Daiello, L.S. Wieland, A. Zgodic, B.T. Smith, T.W. Herliczek, N. Shah, T.A. Trikalinos, Diagnosis of Right Lower Quadrant Pain and Suspected Acute Appendicitis, [Internet] Agency for

Healthcare Research and Quality, Rockville, 2015.

[43] M.H. Ebell, J. Shinholser, What are the most clinically useful cutoffs for the Alvarado and Pediatric Appendicitis Scores? A systematic review, Ann. Emerg. Med. 64 (2014) 365–372, https://doi.org/10.1016/j.annemergmed.2014.02.025.

[44] M. Kularatna, M. Lauti, C. Haran, W. MacFater, L. Sheikh, Y. Huang, et al., Clinical prediction rules for appendicitis in adults: which is best? World J. Surg. 41 (2017) 1769–1781, https://doi.org/10.1007/s00268-017-3926-6.

[45] L.W. Lamps, Beyond acute inflammation: a review of appendicitis and infections of the appendix, Diagn. Histopathol. 14 (2008) 68–77, https://doi.org/10.1016/j.mpdhp.2007.12.002.

[46] S. Di Saverio, A. Birindelli, M.D. Kelly, F. Catena, D.G. Weber, M. Sartelli, et al., WSES Jerusalem guidelines for diagnosis and treatment of acute appendicitis, World J. Emerg. Surg. 11 (2016) 34, https://doi.org/10.1186/s13017-016-0090-5.

[47] R.R. Gorter, H.H. Eker, M.A. Gorter-Stam, G.S. Abis, A. Acharya, M. Ankersmit, et al., Diagnosis and management of acute appendicitis. EAES consensus development conference 2015, Surg. Endosc. 30 (2016) 4668–4690, https://doi.org/10.1007/s00464-016-5245-7.