Diagnostic Validity of Abdominal Sonography in the Early Course of Paediatric Appendicitis

Inauguraldissertation
zur Erlangung des Grades eines Doktors der Medizin
des Fachbereichs Medizin
der Justus-Liebig-Universität Gießen

vorgelegt von

Philipp Klocke

aus Göttingen

Gießen 2021

Aus dem Fachbereich Medizin der Justus-Liebig-Universität
Gießen

Gutachter: Professor Dr. med. F. K. Tegtmeyer

Gutachter: Professor. Dr. med. K. Zimmer

Tag der Disputation: 20. September 2021

Table of Contents

1.0 Introduction

1.1 Problem statement

Appendicitis is the most common paediatric surgical emergency, accounting for ca. 5% of urgent paediatric outpatient visits for abdominal pain alone[1]. Although appendicitis is common and its symptoms are generally recognised by paediatricians and general practitioners, young children are still overlooked due to their atypical presentation[2,3]. Hence, diagnosing appendicitis in children still continues to pose a challenge even for experienced clinicians[4]. In the past few decades, imaging modalities such as ultrasound sonography (USS), computer tomography (CT) and magnetic resonance imaging (MRI) have reinforced the diagnostic repertoire and helped to decrease the incidence of negative appendectomies performed in children[5].

Within the United States, CT scans and their reported sensitivities of close to 100% of appendicitis detection are routinely used to aid physicians in their clinical diagnosis[6]; its utilisation, however, constitutes a significant health risk to children and adolescents[7,8]. Due to increasing concerns over exposure to ionising radiation from medical diagnostics and long-term cancer risk, efforts have been made to reduce the use of CT in favour of increased reliance on ultrasonography[5,6,9]. Nonetheless, it might be reasonable to assume that ultrasound investigations with a reported sensitivity of 'only' 88% will therefore detect appendicitis less frequently, yet sparing children the increased risk of heightened malignancy rates decades later[6].

Based on the presumption that the clinical manifestation of appendicitis is subject to an age-dependent dynamic[10,11] and closely correlates with the duration of abdominal pain[5], Narsule *et al.* (2011) have shown that the appendiceal perforation rate amounts to 7.7% within the first 24 hours after onset of abdominal pain and increases in a linear fashion with the duration of the symptoms[12]. At the same time, USS sensitivity improves as abdominal pain duration increases. Thus, ultrasound performed in the early phase of appendicitis, when the disease is still 'less macroscopic' can lead to a false-negative diagnosis compared to investigations taken further down the course of the disease[5].

With perforation being a major determinant of disease-related morbidity, efforts to improve early detection and hereby reducing the appendiceal perforation rate are necessary[10]. It remains unclear as to what point in time after the onset of abdominal pain

USS investigations demonstrate the highest positive predictive value (PPV) for the presence of appendicitis in a child presenting with acute abdominal pain. Clinicians should therefore determine the optimal timing of diagnostic imaging in order to detect acute appendicitis at an early stage, before it may progress to perforation.

1.2 Literature Review: Appendicitis

Epidemiology and demographics of paediatric appendicitis

Many years have passed since R. H. Fitz first described the perforating inflammation of the appendix in 1886[13]. After all these years, the epidemiology and aetiology of the underlying condition still remain poorly comprehensible. One of the most cited papers regarding the epidemiology of appendicitis stems from a review article from 1990 titled 'the epidemiology of appendicitis and appendectomy in the United States' by Addiss *et al.* In this paper, it was estimated that appendicitis affects approximately 250,000 patients each year, including 77,000 children alone. In both males and females, the highest rates were noted in persons aged 10-19 years of age (peak incidences in males: 10-14 years; females: 15-19 years of age). Moreover, a male dominance of 1.4:1 and an overall lifetime risk of developing appendicitis of 8.6% for men and 6.7% for females was noted, making the incidence of appendicitis higher in males compared to females[1].

Younger children often present with appendiceal perforation as the diagnosis is difficult to make in this age group[14]. In a study by Nance *et al.* (2000), observing 132 children younger than 5 years of age, researchers revealed an appendiceal perforation rate of 74.2% upon surgery[10]. This finding is within the range of other reported perforation rates of 40-89% seen in pre-school children[3,15-17]. Their inability to communicate to parents, atypical presentation and other associated conditions may delay the diagnosis, leading to higher perforation rates seen in this group of patients[3].

Aetiology and pathogenesis

Acute appendicitis most likely underlies a multifactorial aetiology and has been attributed to a variety of possible causes including but not limited to mechanical obstruction, foreign bodies, lymphoid hyperplasia, inadequate dietary fibre consumption, familial susceptibility, factors associated with improved socioeconomic conditions, bacterial-, viral- or parasitic pathogens as well as malignancy[1,18-20]. In normal healthy persons, the

appendix ranges in length from below 5 to more than 25 cm and obstruction can occur at any point from the tip to the junction between the caecum and the appendix[18].

The primary event is believed to be obstruction of the appendiceal lumen with distension following impaired drainage. As a consequence, the 8th to 10th visceral afferent thoracic nerves are being stimulated, resulting in mild to moderate peri-umbilical pain that typically lasts 4-6 hours. As appendiceal distension progresses, intraluminal pressure increases which leads to decreased appendiceal wall perfusion and arterial insufficiency. Tissue ischemia and mucosal compromise follow. Bacteria are then able to invade the luminal wall, followed by transmural inflammation. Once inflammatory processes have reached this point, the inflammation will extend beyond the original structure to the parietal peritoneum and adjacent structures will become inflamed as well. This stage marks the shift whereby the initial peri-umbilical pain is perceived to be 'wandering' to the right lower quadrant (RLQ) of the abdomen. At this point, the pain is typically more severe, continuous and often accompanied by constitutional symptoms, such as anorexia, nausea and vomiting[18,19]. The most common bacterium found in appendicitis is *Escherichia coli* (E. coli)[21].

In young children, this disease model might not always be fully applicable as infants and toddlers present with anatomical differences compared to older children or adults. Young children may have little omentum and intraabdominal fat which facilitates peritoneal spread more easily[10,12]. Additionally, children tend to present later than adults, delaying their diagnosis as well as contributing to higher perforation rates[3,10].

Why the concern for a timely diagnosis of appendicitis?

In general, two main factors correlate with the rate of appendiceal perforation[12]: Observational research has shown that time to and therefore delay in treatment - the interval between first noted symptoms of abdominal pain and surgery - was the most predictive factor for perforation[18,22]. Particularly in one study, Brennan *et al.* (2006) demonstrated that a delay of more than 36 hours was associated with a 65% or greater incidence of perforation[18]. It is worth mentioning that time to first medical contact is primarily dependent on parents or other caretakers. Their level of awareness, availability of transportation, relationships with a primary medical provider, work restrictions,

financial and insurance status are just a few factors among others that all influence the parental decision to have their child evaluated for potential appendicitis[22,23].
Nonetheless, even after prompt onset of abdominal pain and subsequent evaluation, false-negative diagnoses can further contribute to the development of appendiceal perforation. Nance *et al.* and Graham *et al.* (1980) were both able to demonstrate increased rates of cases where care was sought prior to definitive therapy but appendicitis was still missed in pre-school children[10,15]. The child's inability to communicate to parents, atypical presentation and concurrent respiratory tract infections, diarrhoea, otitis media or dysuria can confuse the diagnosis of appendicitis upon presentation and further delay necessary surgery[3].

The other main factor is young age. A reported perforation rate of 74% in children younger than five years of age was found by Nance et al. [10]. Similarly, higher rates of abscess formation were found in children ≥10 years of age[11,24]. Given the anatomical immaturity and lack of an adequate omental barrier, rapid progression to perforation and peritonitis is being facilitated more easily as the omentum is unable to contain or localise infection in young patients[3,14]. Thus, the appendiceal perforation rate in children was shown to double every 6 hours for the first 24 to 48 hours after the onset of symptoms[12], compared to an adult population study by Bickell *et al.* (2006), suggesting that the risk of appendiceal rupture is almost nil within the first 36 hours after the onset of symptoms and remains at 5% thereafter[25]. Although no definitive time range from presentation to appendix rupture has been established, it is believed to be anywhere from 12 to 24 hours in younger children and more than 24 hours in older children[2].

Given the aforementioned evidence on paediatric perforation rates, this would emphasize the importance of rapid intervention and the possibility that a substantial number of perforations occur both before and in the hospital while patients await further testing and surgery[12]. As delay in diagnosis ultimately leads to higher perforation rates, imaging modalities therefore need to be sensible enough to reliably 'rule in' cases of appendicitis while 'ruling out' those that simply mimic the disease to minimise delay. For this reason, the presumptive time-dependent progression from appendiceal inflammation to rupture and consequent abscess formation has served as the primary justification for prompt surgical intervention[26].

Extra-appendiceal conditions mimicking appendicitis

Conditions that mimic appendicitis and present with right iliac fossa pain are broad and are summarised in table 1 below. Special consideration should be given to the female adolescent, who is at risk of ectopic pregnancy, ovarian pathology and other gynaecological conditions. Likewise, perforation in the young child presenting to the emergency department (ED) should be considered as well[14,27].

Table 1. – Differential diagnosis in children with suspected appendicitis:

Less emergent	More emergent
➢ Constipation	➢ Nephrolithiasis
➢ Functional abdominal pain	➢ Right-sided pyelonephritis
➢ Ovarian cyst	➢ Psoas abscess
➢ Mittelschmerz	➢ Pelvic inflammatory disease
➢ Mesenteric adenitis	➢ Bowel obstruction
➢ Gastroenteritis	➢ Ectopic pregnancy
➢ Meckel's diverticulum	➢ Gonadal torsion
➢ Inflammatory bowel disease	➢ Diabetic ketoacidosis
➢ Henoch-Schönlein Purpura	➢ Intussusception
➢ Pneumonia	➢ Septic hip
➢ Urinary Tract infection	

Adapted from Lipsett SC and Bachur RG (2017)[27]

Diagnosis of appendicitis in children

Clinical examination, signs and symptoms

First noted symptoms in classic textbook descriptions of paediatric appendicitis are periumbilical pain followed by nausea, migration of pain to the right lower quadrant, vomiting and low-grade fever[27]. This description however only matches roughly 50% of adult cases and is seen even less frequently in paediatric cases of appendicitis. This makes diagnosing paediatric appendicitis difficult leading to missed diagnoses with significant morbidity and even mortality[2]. Table 2 below summarises the frequency of specific signs and symptoms in children with suspected appendicitis:

Table 2. – Frequency of signs and symptoms in children with suspected appendicitis:

Finding	Appendicitis (%)	No Appendicitis (%)
➢ Anorexia	59.6	47.4
➢ Nausea and/or vomiting	71.1	55.7
➢ Migration of pain	50.2	27.5
➢ Pain duration >48h	82.2	74.1
➢ Absence of diarrhoea	82.9	78.4
➢ Decreased bowel sounds	36.3	14.3
➢ RLQ tenderness	67.8	53.4
➢ Rovsing's sign	31.9	15.9
➢ Rebound pain	48.5	24.7
➢ Temperature ≥38°C	17.3	19.7

Adapted from Becker *et al.* (2007)[2]

Revisiting the different anatomical setup in children, it is worth mentioning that many children present differently than adults[12,14]. Based on a study by Becker *et al.* (2007), table 3. outlines the age-dependent distribution of typical and atypical features found in paediatric appendicitis[2].

Table 3. – Typical and atypical features of paediatric appendicitis

Feature	Typical	Atypical
➢ Age	≥5 years	<5 years
➢ Anorexia	Present	Absent
➢ Nausea/Vomiting	Present	Absent
➢ Migration of pain	Present	Absent
➢ Diarrhoea	Absent	Present
➢ Pyrexia	Present	Absent
➢ Bowel sounds	Decreased	Normal or Increased
➢ RLQ tenderness	Present	Absent
➢ Rovsing's sign	Present	Absent

Adapted from Becker *et al.* (2007)[2]

Becker *et al.* (2007) studied 270 children aged <5 years, looking at atypical features of appendicitis upon presentation. They found that among patients with appendicitis, the most common atypical features included absence of pyrexia (83%), absence of Rovsing's sign (68%), normal or increased bowel sounds (64%), absence of rebound pain (52%), lack of migration of pain (50%) and absence of maximal pain in the right lower quadrant

(RLQ) (32%). Approximately 44% of patients with proven appendicitis had ≥6 atypical characteristics upon presentation which is concordant with Sivit *et al.* (2003), reporting a range of 33-50% of infants and young children presenting atypically[2,18,28]. Likewise, a white blood cell count (WBC) of <10,000/mm^3 and an absolute neutrophil count (ANC) of <7,500/mm^3 were found to be the strongest negative predictors of appendicitis[2].

Physical examination should start with the general assessment of the overall child's appearance and subsequently cover all organ systems that might mimic the symptoms of appendicitis. Key components of the examination include but are not limited to examining the lung fields to evaluate for pneumonia as well as the genitourinary system to rule out testicular torsion, hernia or a haematocolpos in amenorrhoeic pubescent girls. Urine pregnancy test should be obtained in all postpubertal females to rule out an ectopic pregnancy. When examining the abdomen, focus on the location of maximal tenderness should be given, with the RLQ being most suggestive of appendicitis. Rebound tenderness, guarding and referred pain to the RLQ from palpation of the left lower abdomen (Rovsing's sign) can be indicative of secondary peritonitis. In these cases, children are most likely to be unwilling to jump or cough, suggesting peritoneal irritation if positive. Bowel sounds will usually become hypoactive as appendicitis progresses[27]. Rectal examination has become less routine in children as it does not add any further diagnostic value[14,29,30]. Other less specific findings in patients with an inflamed retro-caecal appendix include the obturator sign (pain with internal rotation of the right hip) and the psoas sign (pain with extension of the hip). A normal urinalysis can help rule out new onset diabetes and decrease the chance of urinary tract infection or nephrolithiasis. However, it is worth bearing in mind that many children with appendicitis will have sterile urine or pyuria with a few WBCs in the urine due to ureteral or bladder irritation[14,27].

According to Bundy *et al.* (2007), fever was found to be the most useful sign associated with appendicitis in children who present with abdominal pain and suspected appendicitis (positive likelihood ratio 3.4; 95% CI 2.4 to 4.8). Likewise, absence of fever decreased the possibility of appendicitis (negative likelihood ratio 0.32; 95% CI 0.16 to 0.64)[31]. Contradictory to this is a study of 492 patients by Andersson *et al.* (1999), in which a temperature greater than 37.7°C had a sensitivity of 70% and a specificity of 65%[29]. Evidence about the importance of fever in association with appendicitis still remains controversial.

Laboratory markers

As the sensitivity of clinical examination alone ranges from 54-70% compared to 70-80% in adult patients, simply examining the child with abdominal pain is not sufficient to rule it out[14]. Other diagnostic means need to be added in order to pick up appendicitis more sensitively. WBC, C-reactive protein (CRP), the proportion of polymorphonuclear leukocytes (PMN), fever and other factors have been widely studied and shown to be useful in predicting risk of appendicitis[32]. However, WBC is not a specific marker and commonly elevated in patients with other inflammatory conditions, including differential diagnoses for appendicitis[14,33]. Taken alone, elevated WBC has been found to have limited predictive value for diagnosing appendicitis with a recent study reporting that 21% of negative appendectomies had an elevated neutrophil count at presentation[32,34]. Similarly, Gronroos *et al.* (2009) found that 20% of paediatric patients with pathologically proven appendicitis had a normal WBC count[34,35]. Yet, Wang *et al.* (2007) could illustrate that the presence of both white blood cell count and left shift together has a specificity of 94% and a positive likelihood ratio of 9.8 when appendicitis is suspected clinically[36].

Although no consensus about which WBC thresholds are optimal for maximising predictive value has been found, a large meta-analysis by Andersson *et al.* (2004)[37] found the sensitivity and specificity of a WBC >10,000 cells/mm^3 to be 83% and 67%, respectively.

In a similar fashion, CRP levels show an increase between 8-12 hours after the onset of inflammatory processes with a peak between 24 and 48 hours. Consequently, it is of little diagnostic benefit in the early phase of simple appendicitis[33]. It is important to note that the rise of WBC occurs earlier than that seen in CRP. Again, in a large meta-analysis by Andersson *et al.* (2004), taken alone, CRP >10mg/L was reported to exhibit a sensitivity and specificity of 81% and 59%, respectively[37]. Regardless, a normal CRP and WBC do not rule out appendicitis in children[14].

Novel markers

Recently, researchers have begun looking at other laboratory markers including procalcitonin, interleukin 6, amyloid A, rinoleukograms and others[32,33,38]. However, to date, the power of these studies is considered to be limited with regard to clinical practice[33].

Clinical scoring systems

Children presenting with vague symptoms of appendicitis often pose a vigorous challenge for clinicians as performing surgery might lead to higher negative appendectomy rates while placing the child on observation may increase the risk of appendiceal perforation. As the suspicion of appendicitis relies on the constellation of signs, symptoms and complementary findings, researchers have developed multiple scoring systems to help determine the probability of a patient having acute appendicitis and aid in further decision making[34,39]. Two of the best known scoring systems are the Alvarado Score (1986)[40] and the Paediatric Appendicitis Score (2002)[41].

Table 4. – Alvarado Score (MANTRELS)

Features	Score
Migration of pain	1
Anorexia	1
Nausea and Vomiting	1
Tenderness in the right lower quadrant	2
Rebound tenderness	1
Elevation of temperature >37.5°C	1
Leukocytosis	2
Shift to the left	1
Total	10

Adapted from Alvarado *et al.* (1986)[40]

Alvarado Score

The most commonly used score is the Alvarado score focusing on the typical signs of localised peritonitis and an abnormal leukocytosis (*table 4*). The higher the score, the higher the likelihood of appendicitis[34]. An Alvarado score of greater than 7 was recently shown to have a PPV of 65%[14,42]. However, critics argue that the score is only partially applicable to paediatric populations as it requires children to identify migration of pain, nausea and anorexia which is not easily identified in very young children[33]. Nevertheless, a recent systematic review found that a score <5 points was 94% to 99% sensitive in 'ruling out' appendicitis while the data did not find enough support to 'rule in' for surgery[34].

Paediatric Appendicitis Score

The Paediatric Appendicitis Score (*table 5*), also known as the Samuel Score, was developed based on a patient cohort in the UK[41]. Patients who had a score ≤2 could be discharged home without any further investigation reaching a sensitivity and specificity of 96% and 74%, respectively. Nonetheless, a score ≥6 in children younger than 10 years of age had a PPV of 45%[42]. Thus, the Paediatric Appendicitis Score, similar to the Alvarado score, exhibits similar flaws that if applied to the decision to operate, it would have led to a negative appendectomy rate of 12.9% in one study[34,43].

Concluding, clinical scores may be suitable as an objective instrument for selecting patients for further examination with imaging techniques or next-day re-evaluation[39]. However, they should not be used as the only method of determining the necessity for surgery as they do not have a sufficient positive predictive value and would otherwise increase the negative appendectomy rate[14,34,44].

Table 5. – Paediatric Appendicitis Score

Features	Score
Fever >38°C	1
Anorexia	1
Nausea/Vomiting	1
Cough/percussion/hopping tenderness	2
Right lower quadrant tenderness	2
Migration of pain	1
Leukocytosis >10,000 (10^9/L)	1
Polymorphonuclear-neutrophilia >7500 (10^9/L)	1
Total	10

Adapted from Samuel *et al.* (2002)[41]

Imaging of appendicitis in children

Ultrasound Sonography

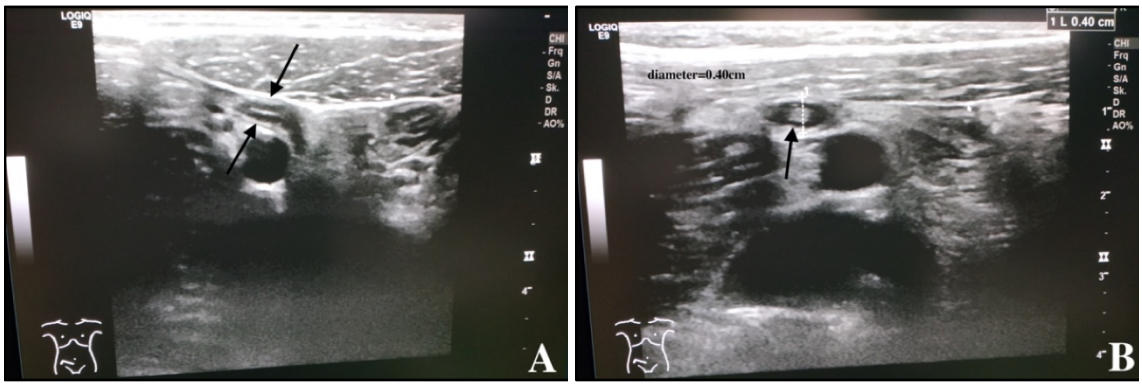
Patients who present with classical signs and symptoms of appendicitis, and are assessed by an experienced surgeon, generally do not require any radiological investigation. However, the problem arises in presentations that are non-specific or atypical[14]. Due to numerous advantages, USS has become a widespread primary diagnostic imaging modality in children presenting with suspected appendicitis[45]. In 1986, Puylaert *et al.* first described the technique of real-time graded compression ultrasound facilitating the

diagnosis of appendicitis[18,46]. This technique is operator dependent in which manual pressure is applied using a linear array of ultrasound transducers to compress and displace bowel loops in order to visualise the appendix[28,46].

In 2014, Coyne *et al.* was able to demonstrate that the diameter of the normal appendix (mean anteroposterior diameter 4.4 ± 0.9 mm, mean transverse diameter 5.1 ± 1.0 mm) does not change with age and is normally distributed in children (*image 1*)[47]. A fluid-filled, non-compressible tubular structure with a diameter of more than 6 mm (*image 1*) is therefore indicative of appendicitis[18,28,46]. Supplementary, Wiersma *et al.* (2005) found that this criterion is independent of age, height and weight[48]. Other features congruent with the diagnosis of appendicitis are the presence of an appendicolith, peri-caecal or peri-appendiceal fluid, increased peri-appendiceal echogenicity secondary to inflammation as well as hyperaemia around the appendiceal wall (*image 2*)[18,21,28].

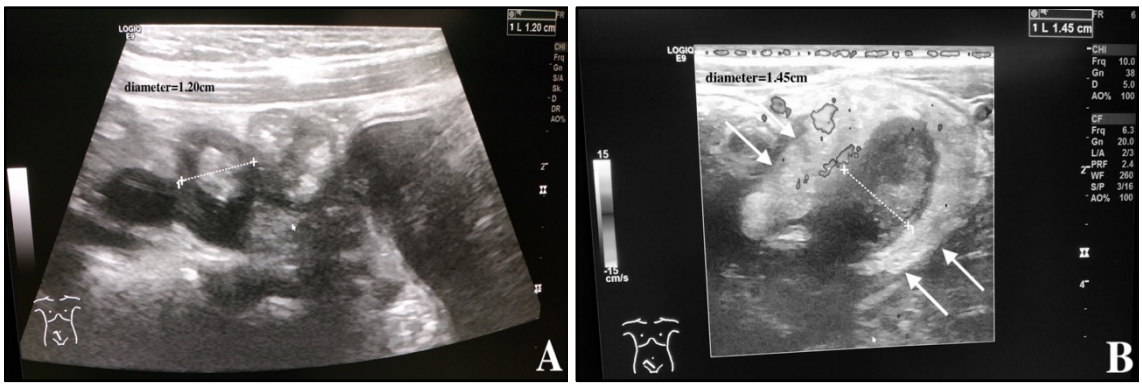
Test characteristics of USS performance for the diagnosis of appendicitis are usually expressed with the terms 'sensitivity' and 'specificity'. The current literature refers to a meta-analysis of studies published between 1986 and 2004 with a reported sensitivity and specificity of 88% and 94%, respectively[6]. More recent literature has found that the sensitivity of USS varies between 72.5-100% and the specificity from 88-98%[49-52]. However, more recent reports have illustrated a sensitivity of 91-99% and a specificity of 97-98%[53,54]. In general, the sensitivity of USS in children presenting with acute appendicitis has often been reported to be lower than that of the closely related specificity. In other words, a positive finding on USS is therefore highly specific and could 'rule in' appendicitis, however, the lower sensitivity cannot "rule it out"[14].

Image 1. – Ultrasound appearance of a normal appendix:



Patient with negative diagnosis. **(A)** Longitudinal abdominal ultrasound sonography of the right lower quadrant demonstrates a compressible, ovoid structure (arrows) without abnormal inflammatory structures, suggestive of a normal appendix. **(B)** Transverse section showing a non-inflamed appendix (arrow) with a diameter of 0.40 cm. No abnormal inflammatory signs are surrounding this appendix.

Image 2. – Ultrasound appearance of an inflamed appendix:



Patient with diagnosed appendicitis. **(A)** Transverse ultrasound sonography of the right lower quadrant demonstrates an enlarged round appendix with a diameter of 1.20cm, as well as abnormal inflammatory appearances (increased abdominal fat echo, fluid-filled centre). **(B)** Longitudinal section of the right lower quadrant demonstrates an enlarged, partially fluid-filled appendix with a diameter of 1.45 cm, surrounded by increased echogenic fat (arrows). These abnormal inflammatory changes together with the enlarged appendix are suggestive of acute appendicitis.

Anandalwar *et al.* (2015) listed the five most common constellations of USS findings[32]:

1. Incompletely or non-visualised appendix, no primary or secondary signs (28.4%)
2. Secondary signs of appendicitis present with/without a visualised appendix, primary signs of appendicitis (23.9%)
3. Incompletely or non-visualised appendix, fluid present in the RLQ or pelvis, no primary or secondary signs of appendicitis (22.8%)
4. Primary signs of appendicitis present without secondary signs, with/without fluid in the RLQ or pelvis (18.1%)
5. Normal appendix visualised without primary/secondary signs of appendicitis (6.7%)

From the above list, one can see that incomplete or non-visualisation of the appendix is the leading constellation in USS findings. This is particularly the case in adipose patients, representing one of the major weaknesses of USS examinations. Increased thickness of the adipose tissue both increases the distance between the ultrasound scanning transducer beam and limits the compressibility which further compromises visualisation of the appendix[9,55]. This has been confirmed in earlier studies identifying the body mass index (BMI) as a factor influencing ultrasound accuracy[56]. In the past and within the United States, children with inconclusive USS results therefore often underwent CT scan examination as non-visualisation did not rule out appendicitis[9]. Furthermore, Emil *et al.* (2001) associated inconclusive USS results with long emergency department stays as a potential increase in perforation rates based on the diagnostic delay that inconclusive results would bring[9,57]. Other limitations are excessive bowel gas, inadequate bladder filling, retro-caecal localisation of the appendix as well as its operator-dependant nature[18].

Despite the flaws that this imaging modality might bring, investigators noted an increased trend towards reliance on USS and decreased use of CT scans for paediatric populations within the United States[58]. More importantly, the American College of Emergency Physicians recommends that USS be considered as the initial imaging modality to diagnose suspected appendicitis in children[9,50]. Many authors support the position that when USS does not identify the appendix clearly, clinicians need to consider other modalities before diagnosing or excluding appendicitis. These can include but are not limited to repeat clinical assessment, laboratory testing, admission for repeat clinical exams or other imaging modalities such as CT or MRI[49].

Computer Tomography (CT)

The diagnostic use of radiography is based on the varied absorption of radiation by different tissues such as bone, air and others. Radiation that is not absorbed by the body passes onto a radiosensitive film. In CT scans, multiple beams of radiation from different directions are emitted producing a 3-dimensional representation of the relative tissue densities[49].

In a large meta-analysis by Doria *et al.* (2006), sensitivity for CT to differentiate appendicitis from non-appendicitis was reported at 94% which is 6% higher than that of USS as reported in the same study. Likewise, specificity for CT was 95%, 1% higher than

that reported for USS[6]. In the United States and potentially in many other countries, availability of USS support is limited at night, further increasing the use of CT for patient evaluation. Moreover, CT scans are often preferred because of their general availability, ability to detect appendicitis in obese patients and uniform use in contrast to paediatric USS requiring specialised training for its performance and interpretation[59]. However, CT introduces a significant risk of radiation-induced malignancy in children[9,60,61]. Because even a single CT implicates risks, its use must be limited.

In a large study by Bachur *et al.* (2012) assessing 1810 children, researchers were able to show that USS sensitivity to differentiate appendicitis from non-appendicitis increases from 81% in the first 12 hours after onset of abdominal pain to 96% after 48 hours of pain. In a similar fashion, the NPV increased with the duration of pain with a marked improvement after 36 hours[5]. Against the presumption that abdominal pain duration would increase the sensitivity of CT, this was not shown in the study. Similarly, the specificities of both USS and CT were not affected by duration of abdominal pain[5]. Despite the higher sensitivity of CT, it is worth noting that the sensitivity of USS is reasonably high in children and that the exposure to radiation is a concern that needs to be balanced[62]. Whenever a clinician opts for a CT, he or she needs to consider a trade-off between the future risk of cancer with the use of CT and the risk of missing positive cases with USS[6].

The improved performance of USS over time indicates that this imaging modality should be considered more as a first-line option in daily practice rather than to perform a CT scan in a child presenting with inconclusive USS results. For patients presenting with mild RLQ pain for less than a day, this could mean to abstain from multiple imaging and instead consider monitoring and repeated clinical examinations to avoid overreliance on CT[5].

Why the concern of ionising radiation in paediatric imaging?

In recent times, increasing concerns within the American healthcare system over exposure to ionising radiation from medical diagnostics have led to efforts trying to reappraise the use of CT in children[60]. Possible increase in future cancer risk has been estimated with risk projection models that were mainly derived from studies of atomic bomb survivors

in Japan[60,61,63]. It is believed that children have a higher number of dividing cells; radiation will therefore increase the risk of genetic alteration and facilitate cancerous transformation if cellular reproduction is modified[18,64]. These risk projection models revealed three significant factors with regards to exposure and risk of malignancy. First, radiation induced malignancy (RIM) is considered to occur at the same age as spontaneous cancers of the same type seen in adults but with higher frequency.
However, in children it would take over 50 years to judge the impact of radiation exposure. Furthermore, because of the relatively large organ size and lack of shield tissue, adult type radiation doses result in higher effective organ doses in children. The effective dose is therefore larger in an infant than in an adult. Lastly, younger children have a higher risk of RIM compared to adults receiving the same dose of radiation as they have a longer life expectancy in which potential cancer-promoting effects of radiation can manifest whereby females were found to be more radiosensitive than males[61]. To give a rough estimate as to what extent of radiation children are exposed to, Brody *et al.* (2007) reported in a clinical report that the estimated medical radiation dose for a 5-year-old child undergoing abdominal CT scan is equivalent to 250 chest radiographs at once[65]. Concluding, the National Research Council has cautioned that there is no lower threshold of exposure of radiation that has been identified to be without risk[66,67].

Negative appendectomy

Removing the appendix surgically is one of the most commonly performed emergency operations in children[68]. Evaluating a child with presumed acute appendicitis means to balance the early surgical intervention, hoping to prevent potential perforation against a more restrictive course of action where the risk of unnecessary surgery is reduced in case the initial presumption proves wrong[33]. A false-positive diagnosis of appendicitis may lead to an unnecessary surgery, termed negative appendectomy (NA). Over the years, improvements in diagnostic modalities, especially USS and CT, have both helped to decrease the incidence of NA[5,69]. In a large retrospective study by Oyetunji *et al.* (2012), including over 250,000 children from the years 2000, 2003 and 2006, researchers determined an overall NA rate of 6,7%. Moreover, they could confirm a decreasing trend towards lowered NA rates (8.1% in 2000 to 5.2% in 2006)[69]. Not surprising, they asserted that the NA rate was highest among children younger than 5 years of age when compared to older children which may be again due to their atypical presentation and greater diagnostic uncertainty. More precisely, the top primary diagnosis accounting for NA were

RLQ abdominal pain (22.2%), other diseases of the appendix (16.2%), mesenteric lymphadenitis (12.5%), ovarian cyst (3.4%) and intussusception (3.0%)[69]. It comes as little surprise that female sex was associated with increased NA with gynaecological conditions possibly confusing the diagnosis.

Nonetheless, the same researchers found that an increased NA rate was linked to a prolonged length of hospital stay which may be due to initial misdiagnosis and subsequent delay in establishing the correct diagnosis[69]. More importantly, the spectrum of morbidity that NA brings includes but is not limited to early complications such as abscess formation, wound infection as well as later complications such as infertility and intra-abdominal adhesions, often presenting many years after the initial procedure has taken place[18]. Still, higher rates of NA have to be accepted in order to reduce the rate of appendiceal perforation as described above[33]. Major goals that need to be achieved are to timely and accurately diagnose a child with suspected appendicitis, to minimise missed appendicitis (false-negative), to avoid misdiagnosis (false-positive) leading to negative appendectomy and to properly identify the appendicitis before perforation[5].

Management of acute appendicitis in children

Operative treatment

Once appendicitis is confirmed, antibiotics should be administered to all patients as recommended by the American Paediatric Surgical Association to prevent intra-abdominal abscess formation, septicaemia or wound infection[27,70]. Surgical removal of the appendix should then follow. Accompanying consequences reported with the surgical procedure include post-operative abscess formation, wound infection and subsequent adhesive bowel obstruction. Yet, the risk of acquiring these complications has been shown to be higher in populations in whom the appendix was found to be perforated[70].

Non-operative treatment

With over 130 years since first describing appendicitis and subsequent surgical removal of the appendix by R. H. Fitz, appendectomy has become the mainstay of treatment of acute appendicitis[13]. However, just recently, the widely recognised principle of surgery is questioned with growing literature pointing towards an antibiotic-only approach as effective treatment for acute appendicitis[71]. In fact, environments without real access to

surgical care (e.g. submarines or merchant navy) have treated appendicitis in a non-operative fashion for many years. Well-documented evidence exists proving high success rates[71]. Yet, evidence for non-operative treatment suggests it may only be effective for uncomplicated non-perforated appendicitis. Further evidence remains to be seen[71,72].

In an ideal world, medical expertise would allow to select those children who could be predicted to respond to non-operative treatment and proceed without surgery, while preserving surgery for those remaining. Currently, with all the medical achievements we are unable to make this prediction. Despite the availability of clinical scoring systems, none so far can reveal those children likely to respond to non-operative treatment, those who need surgery and those in whom appendiceal perforation is inevitable[71].

1.3 Aims of the study

As outlined above, two potential situations in children presenting with suspected acute appendicitis need to be avoided: Any delay in diagnosis and subsequent appendiceal perforation as well as unnecessary appendectomy. It is therefore of upmost importance that children presenting with suspected acute appendicitis receive a timely and, more crucially, an accurate diagnosis translating into lower missed appendicitis (false-negative) rates, fewer misdiagnoses (false-positive) leading to negative appendectomy as well as proper identification of the appendix that is close to perforation.
In face of considerable efforts to establish and validate clinical scoring systems to guide the management of patients with suspected appendicitis, their performance has in the past shown to be inadequate for clinical management. Continuous improvements in technology and imaging modalities over the past 15 years have substantially improved the accuracy of imaging techniques to diagnose acute appendicitis. Unfortunately, children are often still overlooked and misdiagnosed, contributing to greater morbidity.

As abdominal pain duration has previously been associated with the severity of the disease, it was hypothesised that the performance of ultrasound imaging may be diminished in the early phase of acute appendicitis. Clinicians should therefore determine the optimal timing to detect acute appendicitis at the time of greatest diagnostic certainty. The aim of this study was to find clinical parameters that are associated with a diagnostically conclusive ultrasound on admission to help improve the accuracy of ultrasound performance according to the duration of abdominal pain. Following previously conducted research on prediction models of appendiceal perforation, we tried to further broaden the existing scientific evidence on this topic and searched for parameters that are associated with appendiceal perforation in a cohort of children with histologically proven appendicitis. Lastly, we investigated parameters that were collected upon admission to the emergency department that led clinicians to call for a prompt surgical intervention (<6 hours) over delayed appendectomy. The investigation was supposed help determine the association of each parameter with the necessity to perform surgery urgently in a child with acute appendicitis. The results of this study shall help optimise the validity of this imaging modality, decrease negative appendectomies performed in children and further help to reduce the appendiceal perforation rate and its complications.

2.0 Material and Methods

2.1 Study Design and Setting

After obtaining ethical approval from the University of Giessen ethics committee *(Aktenzeichen: 225/18)*, we performed a retrospective cohort study of all patients, aged 0-18 years, who underwent abdominal USS evaluation for abdominal pain and suspected appendicitis with subsequent appendectomy between January 2015 and June 2019 in a single tertiary medical centre. Owing to the retrospective nature of this research project, the necessity for an informed patient consent (from either parent or guardian) was waived as all data were collected as part of routine work-up. All collected data that were related to patient names were first pseudo-anonymised and later anonymised, based on the standards of good clinical practice.

2.2 Selection of Patients

At our tertiary institution, USS is the first-line imaging test for acute appendicitis and is typically attempted on all patients presenting with acute abdominal pain. All patients younger than 18 years of age who underwent right lower quadrant ultrasound during the study period with a final diagnosis of appendicitis based on the subsequent pathology report were included in the study population (*see table 6* below*).*

Table 6. – Inclusion Criteria

Inclusion Criteria	▪ Patients aged 0-18 years of age ▪ Primary symptom of acute abdominal pain ▪ Appendicitis as discharge diagnosis along with histological confirmation ▪ Admitted during sampled study period: 01.01.2015 – 01.06.2019

Table 7. – Exclusion Criteria

Exclusion Criteria	➢ Patients aged ≥18 years of age ➢ Pregnancy at the time of admission ➢ Duration of abdominal pain >120 hours ➢ Prior abdominal operation ➢ Patients with concurrent psychiatric conditions (Autism, Intellectual Disability) ➢ Prior acute radiological studies before admission to emergency department ➢ Concurrent primary gastro-intestinal conditions: ▪ Inflammatory bowel disease ▪ Acute cholecystitis ▪ Meckel diverticulitis ▪ Lymphadenitis ▪ Perforated gastric ulcer ▪ Diverticulitis ▪ Pancreatitis ▪ Psoas abscess ▪ Renal colic ▪ Pyelonephritis ▪ Endometriosis ▪ Ectopic pregnancy ▪ Ruptured ovarian cyst ▪ Ovarian (-cyst) torsion ▪ Acute testicular torsion ▪ Salpingitis ▪ Gastroenteritis ▪ Terminal ileitis ▪ Diabetic ketoacidosis ▪ Porphyria

We excluded patients who went directly to the operating theatre without any imaging performed, patients who were pregnant upon presentation, those who presented >120 hours after onset of abdominal pain or had prior acute radiological studies performed before presenting to ED. We also excluded patients with major psychiatric conditions that might interfere with an accurate clinical assessment. Major psychiatric conditions included but were not limited to intellectual disability or autism spectrum disorder. Patients with incomplete or insufficient medical records were excluded, as well as patients with any of the gastro-intestinal conditions as outlined in table 7 *(see above)*.

2.3 Institutional Protocol

Our paediatric emergency department sees about 25.000 children each year. The standard protocol at our institution after children present to ED is to first clinically evaluate the child. A detailed history is taken, a complete physical examination is performed, and blood is taken for a complete blood count. Radiologic studies such as USS scans are usually done in all children with acute abdominal pain but remains at the discretion of the treating physician. The specific sonography approach to the RLQ includes a graded compression technique described by Puylaert *et al.* (1986)[46].

Children in whom the diagnosis of appendicitis was uncertain after physical examination and USS scan either underwent a second ultrasound scan, were placed on observation or had diagnostic laparoscopy performed. CT scans are only being performed in very rare cases when other severe abdominal conditions are suspected or to visualise appendiceal perforation that might otherwise not be detected on USS. Throughout the study period, no appendicitis was confirmed by use of CT scans. During office hours, ultrasound studies are generally performed by an onsite paediatric radiology resident with at least 3 months of experience and supervised by a paediatric radiology consultant. In ambiguous cases, children are seen a second time by the paediatric radiology consultant or the head of the department. Out of hour scans are available and provided by an onsite paediatric radiology resident, supervised by an on-call consultant radiologist.

2.4 Data Collection and Processing

Demographic and Clinical Data

For the purpose of this study, gathered data included demographic information (age, gender, presence of concurrent infection, date and time of admission to ED). Presence at time of admission or occurrence of fever after onset of abdominal pain was assessed either at the time of admission or based on the caretaker's report. A fever was any temperature reported or recorded at or above 38.5°C. Thus, a raised body temperature was categorised as 'subfebrile' if reported or recorded at 37-38,4°C. The lack of raised body temperature in the patient's/caretaker's own accounts and at the time of admission was coded as 'absent'.

Serological blood tests included inflammatory markers (CRP, WBC) and were both taken on admission. Laboratories that processed blood samples had prior been approved, ensuring current quality assurance standards (Teil A der Rili-BÄK 2014 [Deutsches Ärzteblatt, Jg. 111, Heft 38 vom 19. September 2014, S. A 1583])[73]. In cases where no blood tests were taken at the time of USS examination, it was the blood test taken most approximate to the time of USS examination. In accordance with previous studies, the upper limit of the reference interval for leukocyte count was 10 x 10^9 /L and that for CRP was 10 mg/L[74].

Abdominal Pain Duration and Ultrasound Examination

Duration of abdominal pain was defined as the point of earliest remembrance of acute abdominal pain onset leading up to the first USS examination in ED. Owing to the retrospective nature of this study, it was impossible to define an absolute temporal set point for the onset of abdominal pain. Therefore, both patients/caretakers, where possible, were interviewed on admission and enquired about the most proximate onset of abdominal pain. Depending on the time 'last seen well' and earliest point of remembrance of abdominal pain onset, an estimated point in time was set and used to compute the difference between time of USS examination and abdominal pain onset (in hours). This variable was later coded into a categorical variable (see *'Outcome measures'* below) with the true onset of abdominal pain not deviating more than a few hours from the estimated onset in time.

Date, time and results of the USS scans, histopathological/surgical reports, time-interval from time of admission to appendectomy, type of abdominal access (laparoscopic/ open surgical appendectomy) as well as final discharge diagnosis were taken from the patient's electronic file. Missing variables after a thorough patient's record search were presumed absent (e.g. no temperature taken) and coded as 'not documented'. We defined suspected appendicitis in patients whose treating physician obtained blood tests, USS studies and whose management involved paediatric surgical consultation for the purpose of diagnosing appendicitis.

In order to determine the time interval leading up to operation, we used time of admission to hospital as a surrogate marker for when surgical evaluation was performed. This was later subtracted from time of operation and the difference calculated. Although this may not always have been accurate, we felt it was a better approximation than other potential starting points because of confounding factors that may have biased the data. The final diagnosis was made on the basis of the pathologist's histological report reflecting the inflammatory nature of the tissue sample sent after surgical appendectomy.

Ultrasound Examination and Definition

From the paediatric ED, the great majority of cases in which the diagnosis of acute appendicitis was not evident right from clinical and physical examination were referred to the paediatric radiology department. All ultrasound studies were obtained using a 2-8 MHz broad-spectrum linear probe transducer for bowel and iliac fossa images and broad-

spectrum convex probe transducer (2-9 MHz) for pelvis and solid organs using GE Healthcare LOGIQ E9 and S8 ultrasound machines. Ultrasound images as well as the paediatric/general radiologist's report were retrospectively sought, and findings abstracted into a collection sheet and later categorised according to three diagnostic categories for the suspected presence of appendicitis (positive, equivocal and normal). Details of the diagnostic criteria for each subgroup are based on previous studies and are outlined below (*see table 8)*[5,19,75-81].

Table 8. – Interpretative Categories for Ultrasound Findings

		RADIOLOGICAL FEATURES
ULTRASOUND CATEGORIES	**Normal**	➢ Appendix as blind-ending, tubular structure, axial diameter of <6 mm ➢ No free fluid in abdominal space ➢ Negative pain reaction when compressing on the appendix ➢ Appendix is compressible ➢ Normal echogenicity and vascularity of the appendix ➢ No appendiceal- or peri-appendiceal abscess ➢ No sign of appendiceal perforation
	Positive	➢ Appendiceal axial diameter of ≥6 mm ➢ Presence of thickened echogenic peri-appendiceal/ intra-abdominal fat ➢ Extraluminal fluid collection (e.g. Douglas pouch) ➢ Hyperaemic appendiceal wall using doppler-sonography ➢ Presence of an appendicolith ➢ Non-compressibility of the appendix ➢ Positive pain reaction when pressing on appendix ➢ Complex space-occupying lesion as in 'conglomerate tumour'
	Equivocal	➢ Unable to visualise the appendix but positive pain reaction when pressing on RLQ ➢ Able to visualise appendix but no radiological evidence to definitely categorise appendix as 'normal'

According to these abstraction rules, the radiologist's report was coded as *'normal'* when the appendix was visualised and no secondary signs of appendicitis were present (thickened echogenic mesenteric fat, enlarged intra-peritoneal lymph nodes, extraluminal fluid collection, appendicolith, positive pain reaction when compressing the appendix, conglomerate tumour); or no visualisation of the appendix plus no secondary signs of

appendicitis[5,48,75-77,80]. A radiologist's report was coded as '*positive'* for appendicitis if it explicitly stated "appendicitis" or "consistent with appendicitis" and included ultrasound findings both compatible with primary and secondary sonographic findings of appendicitis as outlined in the table above[5,19,53,75,76,78,80,81]. For the categorisation of a '*normal'* and a *'positive'* finding, the appendix had to be visualised in full length on ultrasound examination. A perforated appendix seen upon ultrasound exam also counted as a *'positive'* finding. Lastly, the radiologist's report was coded '*equivocal'* if the appendix could not be visualised and one or more secondary signs of appendicitis were present or if the final impression stated "unclear", "unsure" or "not conclusive" (termed non-diagnostic USS)[5].

2.5 Outcome Measures

The primary outcome was the presence or absence of appendicitis upon USS scanning. Final diagnosis of appendicitis was based both on the histological findings at pathological examination and the surgical report describing the macroscopic nature of the appendix at operation. The pathologists were blinded to the ultrasound scanning results.

USS examination results were categorised as *'true-positive'* if the ultrasound was conclusive for acute appendicitis along with a positive histopathological report, or *'true-negative'* when the ultrasound findings would rule out the presence of appendicitis with negative findings on histopathology and no subsequent development of appendicitis. As described above, it is not uncommon for patients with *'equivocal'* findings on USS to receive additional USS imaging throughout the course of the ED stay if clinical examination is not sufficient enough to diagnose appendicitis on its own. In the event that initial *'equivocal'* cases turned into *'positive'* cases on subsequent ultrasound imaging, it was always the USS examination with the greatest diagnostic value to rule in appendicitis. In these circumstances, *'positive'* cases were chosen over *'equivocal'* cases and used for the analysis. Similarly, for *'equivocal'* cases that remained *'equivocal'* despite further imaging, it was always the most initial ultrasound exam that was used for the analysis.

'False-positive' results were considered when the USS read appendicitis, but with a subsequent appendectomy and no evidence of appendicitis on the histopathology report (negative appendectomy). *'False-negative'* results were reports in which USS read negative for appendicitis, but the patient had appendicitis confirmed at subsequent surgery and histopathology. Diagnostically accurate USS results included both *'true-*

negative' and *'true-positive'* examinations (termed diagnostic USS). Inaccurate USS results were reports with *'false-positive'*, *'false-negative'* interpretations and those with *'equivocal'* readings (termed non-diagnostic USS) albeit equivocal USS results on their own are not synonymous with non-diagnostic USS. For the sake of completeness, *'equivocal'* cases were retrospectively reappraised by a paediatric radiology resident and/or consultant together with the principal investigator in order to resolve any inconsistencies. Secondary outcomes included the presence or absence of perforation which was determined by the paediatric surgeon's operative report and verified by the microscopic findings as described in the histopathological report. None of the included patients with appendicitis were managed non-operatively, all appendices were studied histologically. Other secondary outcomes included the duration of abdominal pain. In order to allow for subgroup analysis of the patient cohort, abdominal pain duration was grouped and categorised into 10 mutually exclusive groups. Table 9, below, illustrates the different categorical groups.

Table 9. – Duration of Abdominal Pain

0< x <12 hours	12≤ x <24 hours	24≤ x <36 hours	36≤ x <48 hours	48≤ x <60 hours
60≤ x <72 hours	72≤ x <84 hours	84≤ x < 96 hours	96≤ x <108 hours	108≤ x ≤120 hours

Additionally, sex, age, presence of fever and inflammatory markers (CRP, WBC) were later incorporated into the statistical analysis. As the manifestation of appendicitis symptoms is hypothesised to strongly correlate with the patient's age, we divided the study population and rearranged all patients into three mutually exclusive groups depending on their age at presentation to ED (*see table 10* below). The allocation was conducted for practical reasons taking the different steps in a child's development into account.

Table 10. – Age allocation

0< x <6 years of age (pre-school age)	6≤ x <12 years of age (common school age)	12≤ x <18 years of age (adolescent age)

Lastly, the time interval from admission to ED and time of surgery (termed *latency*; in hours) was assessed to compare secondary outcomes of short latency (<6 hours) and longer latency (≥6 hours).

2.6 Primary Data Analysis

For statistical analyses and graphics, SPSS® v26 (IBM, Armonk New York, USA), Microsoft Excel (Microsoft Excel; Microsoft, Redmond, WA, USA) and R 3.6.3 (R Core Team 2020) together with the packages openxlsx (Schauberger and Walker 2019), stringr (Wickham 2019), lubridate (Spinu, Grolemund, and Wickham 2018),Hmisc (Harrell Jr, Charles Dupont, and others. 2020), lattice (Sarkar 2018), car (Fox, Weisberg, and Price 2020), and effects (Fox *et al.* 2019) were used. Descriptive analyses were performed to characterise the study population. In order to provide reference values for groups of primary and secondary outcome (positive *vs.* equivocal appendicitis on ultrasound; perforated appendicitis *vs.* non-perforated appendicitis) as well as for all other subgroup-analyses performed, group means and standard deviations (mean ± SD) were calculated for all normally distributed data. Medians with quartiles 1 and 3 (Q_1; Q_3) and minimum – maximum were computed for non-normally distributed data of all variables. Normal distribution of all outcome results was assessed by means of visual inspection of quantile-quantile plots (Q-Q-plots).

Pearson's chi-squared analysis (χ^2) or Fisher's exact test were used to compare categorical variables, and independent sample t-tests were used to compare means of parametric continuous variables. In order to compare non-parametric continuous variables, the Mann-Whitney U test was used. Intergroup differences were assessed using One-Way analysis of variance (ANOVA) as well as Spearman's Rank Correlation coefficient. A p-value of ≤ 0.05 was considered statistically significant. Missing variables were not included in the final statistical analysis. Subgroup analyses were performed for cases with appendiceal perforation as well as for each of the three age-allocated groups *(table 9)*.

Logistic Regression Models

In order to test for factors associated with an increased likelihood of showing the outcome of interest, logistic regression models were used. Binary outcomes (response variable) of interest were: [1] USS exam result (positive *vs.* equivocal appendicitis) on admission to ED, [2] perforation status (perforated *vs.* non-perforated appendicitis) at surgery/pathology and [3] latency to operating theatre (OT) (<6 hours *vs.* ≥6 hours), respectively. Input (Co-) variables used in these analyses included age on admission, abdominal pain duration (continuous parameter), leukocyte count, CRP level as well as sex. The logistic regression model for showing associations for a perforated appendicitis

also included abdominal ultrasound result as additional input variable. In [3] we tested for factors allowing prediction of short latency to OT (<6 hours). Presence of fever and subfebrile temperature (≥37.0°C) were not taken into this regression model as the amount of missing data was too large to allow an adequate assertion about the effect of fever and subfebrile temperature on the short latency prediction model. For reference, the equations of the utilised multiple logistic regression models are given below:

[1] P(Result.USS = Positive APP | Age, Abdominal Pain Duration, Leukocytes, CRP, Sex)

$$= \frac{1}{1 + \exp\left(-\left(\beta 0 + \beta 1 \cdot \text{Age} + \beta 2 \cdot \text{Abd. Pain Dur.} + \beta 3 \cdot \text{Leukocyte} + \beta 4 \cdot \text{CRP} + \alpha_{1,Sex}\right)\right)}$$

[2] P(Perforation.Result= Perforated APP | Age, Abd. Pain Duration, Leukocytes, CRP, Sex, Result.USS)

$$= \frac{1}{1 + \exp(-(\beta 0 + \beta 1 \cdot \text{Age} + \beta 2 \cdot \text{Abd. Pain Dur.} + \beta 3 \cdot \text{Leukocyte} + \beta 4 \cdot \text{CRP} + \beta 5 \cdot \text{Result. USS} + \alpha_{1,Sex}))}$$

[3] P(Latency= <6 Hours | Age, Abd. Pain Duration, Leukocyte, CRP, Sex, Result.USS.)

$$= \frac{1}{1 + \exp(-(\beta 0 + \beta 1 \cdot \text{Age} + \beta 2 \cdot \text{Abd. Pain Dur.} + \beta 3 \cdot \text{Leukocyte} + \beta 4 \cdot \text{CRP} + \beta 5 \cdot \text{Result. USS} + \alpha_{1,Sex}))}$$

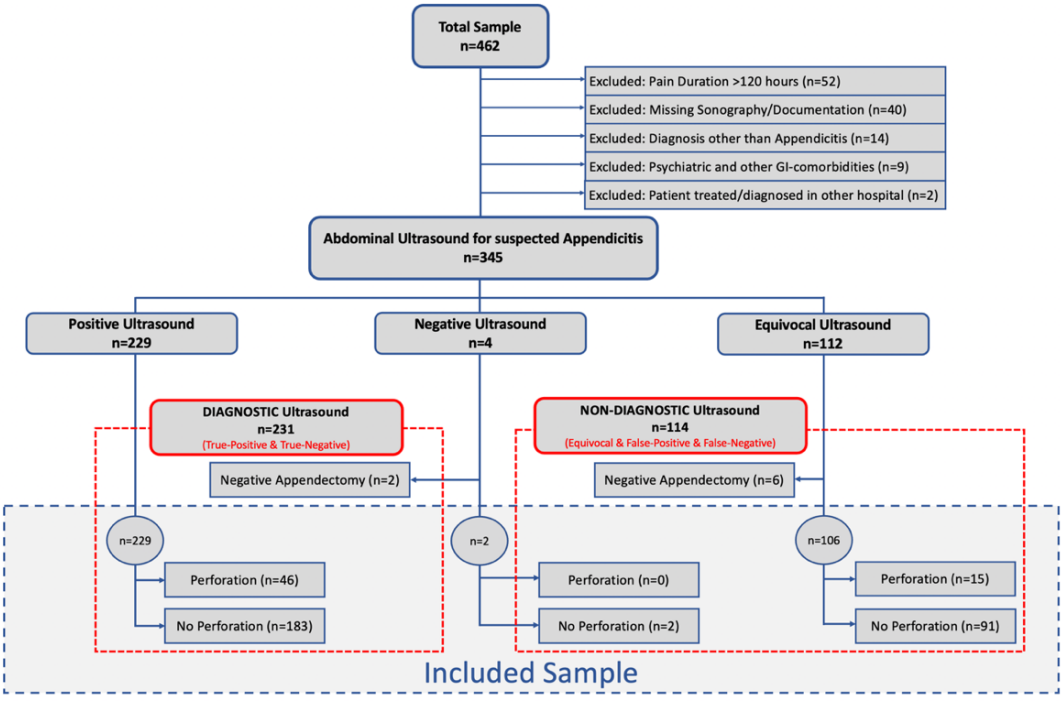
Statistical output was displayed using odds, odds ratio, odds confidence interval (2.5%; 97.5%) and p-value. To help visualise the magnitude of each individual parameter predicting the outcome of interest, each parameter's predicting effect was illustrated graphically. Each graph includes a vertical axis which is supplied with the estimated probability of the 'occurring event' (scaled non-linearly from 0-1) of either 'positive ultrasound' or 'perforated appendicitis'. The horizontal axis exhibits the range of raw data for each parameter, respectively. Also, every single case showing the outcome of interest is listed on this scale with a black "tick mark". A continuous blue line for continuous input variables, and a blue ring for categorical input variables indicate the estimated probability of achieving the outcome of interest in dependence of the specified co-variable. The surrounding blue shade (continuous input variables) and the pink error bars (categorical input variable) mark the 95% confidence interval.

3.0 Results

3.1 Characteristics of Study Subjects

Four hundred sixty-two paediatric patients were screened for presumed appendicitis between January 2015 and June 2019 (4.5 years) whose discharge diagnosis mentioned 'appendicitis' and whose primary treatment included surgical appendectomy. After retrospectively reviewing inclusion- and exclusion criteria, 52 patients (11.25%) were found to have had an abdominal pain duration of >120 hours and were subsequently excluded from further analysis (see *Figure 1 – Patient Flowchart*). 40 patients (8.66%) were excluded due to lack of and/or inconsistent documentation that prevented any further analyses. In 14 cases (3%), patients were retrospectively found to have had a diagnosis other than appendicitis and were subsequently excluded. Psychiatric and other concurrent gastrointestinal comorbidities were found in 9 patients (1.95%) and removed from further analysis. Diagnoses in this group included but were not limited to nonspecific peritonitis, salpingitis, ovarian pathology, severe constipation as well as autism. Lastly, 2 patients (0.43%) were diagnosed and/or treated in other hospitals prior to presenting to our emergency department. These cases were excluded as well since presenting symptoms and abdominal pain duration were deemed to be incomprehensive given the delay in diagnosis and appendiceal visualisation at the time of admission.

Figure 1. – Patient flowchart



Demographic and clinical characteristics are displayed in table 11. Eight patients (1.73%) were found to have had a histologically normal appendix on the pathologist's report. The remaining 337 cases (73%) were categorised as histologically proven appendicitis and were entered into the final study sample after meeting our pre-defined inclusion criteria.

Table 11. – Demographic and clinical characteristics

Criteria	**Population (n=337)**
Age (y)	
Mean ± Std. Deviation	11.51 ± 3.46
Minimum – Maximum	2.25 – 17.83
0< x <6 years, n (%)	23 (6.8)
6≤ x <12 years, n (%)	151 (44.8)
12≤ x <18 years, n (%)	163 (48.4)
Sex, male, n (%)	176 (52.2)
Duration of abdominal pain (h)	
0< x <12, n (%)	48 (14.3)
12≤ x <24, n (%)	103 (30.6)
24≤ x <36, n (%)	82 (24.3)
36≤ x <48, n (%)	28 (8.3)
48≤ x <60, n (%)	27 (8.0)
60≤ x <72, n (%)	19 (5.6)
72≤ x <84, n (%)	17 (5.0)
84≤ x <96, n (%)	6 (1.8)
96≤ x <108, n (%)	7 (2.1)
108≤ x ≤120, n (%)	0 (0)
Leukocyte count on admission (x 10^3/µL)	
Mean ± Std. Deviation	14.58 ± 5.21
CRP on admission (mg/l)	
Median (Q_1; Q_3),	20.0 (6.0; 59.3)
Fever, n	
≥38.5°C, n (%)	44 (13.1)
37.0-38.4°C, n (%)	49 (14.5)
No Fever, n (%)	184 (54.6)
No documentation, n (%)	60 (17.8)
Time from Admission to OT, h	
Median (Q_1; Q_3)	8.0 (5.0; 21.0)
Perforation, n (%)	61 (18.1)
Abdominal Access:	
Laparoscopic, n (%)	302 (89.6)
Open surgery, n (%)	35 (10.4)

C – Celsius
CRP – C-reactive Protein (Values classified as <3 were computed as CRP=2)
h – Hours
OT – Operating Theatre
Q_1, Q_3 – Quartile 1 and Quartile 3
y – Years

The following analyses are based on this sample group (n=337). The male-female ratio was 1.1:1, and mean age was 11.51 years. Based on our pre-defined age groups, 6.8% of included patients were below the age of 6 years at the time of admission (preschool age), with the remaining children almost equally distributed among the groups 'common school age' and 'adolescent age' (6≤ x <12 years: 44.8%; 12≤ x <18 years: 48.4%). The most frequent abdominal pain duration reported across all 10 subgroups was the '12≤ x <24-hours' category (30.6%), followed by the '24≤ x <36-hours' category (24.3%) and '<12-hours' category (14.3%). Mean leukocyte count on admission for the entire population was reported at 14.58 ± 5.21 x 10^3/µL. Median CRP level on admission was found to be 20.0 (6.0; 59.3) mg/l. Occasionally, CRP values were measured at levels lower than 3mg/L and indicated as '<3 mg/L' on the laboratory record. The categorical value '<3 mg/L' was converted into a continuous one by computing and coding it as 2 mg/L, aiming at the middle of CRP 1-3 mg/L. Among all patients, the vast majority did not report any signs or symptoms of raised temperature at the time of admission (54.6%), followed by 14.5% reporting slightly elevated to subfebrile temperature (37.0-38.4°C) and only 13.1% reporting or presenting with signs and symptoms of acute fever (≥38.5°C). No information was documented in 17.8%. All patients underwent surgical appendectomy with 35 (10.4%) performed open surgeries and 302 (89.6%) laparoscopic procedures. Overall, appendiceal perforation rate for our cohort was reported in 61 (18.1%) cases. Median duration from time of admission to time of surgery was 8 hours (Q_1: 5h; Q_3: 21h).

3.2 Comparison of Ultrasound Results on Admission

At the time of admission, 229 (68%) patients were identified as having a 'positive ultrasound' result (*see table 8 - Interpretative Categories for Ultrasound findings, section 2.4 Data Collection and Processing*), 106 cases (31.5%) were classified as 'equivocal ultrasound' and 2 cases (0.5%) were reported with an ultrasound exam that was 'normal' (with no follow-up ultrasound exam) despite the pathologist's report that later confirmed acute appendicitis. Since the focus of the present work is set on scenarios where the USS result (equivocal & positive USS) would provide therapeutic consequences for the treating clinician (opt for surgery or wait), we decided to perform further group comparisons between patients showing 'positive' and 'equivocal' ultrasound results only (n=335) and refrained from analysing negative USS results separately. Patients identified with a 'positive ultrasound' upon arrival to ED were statistically significantly younger (11.12 ± 3.26 years *vs*. 12.26 ± 3.72 years; p=0.005) *(table 13)*. Among all cases with a

positive ultrasound exam, more than half of them presented in the age group of '6≤ x <12 years' (n=115; 50.2%) whereas the majority of equivocal ultrasounds on admission presented in the age group of '12≤ x <18 years' (n=62; 58.5%). The proportion of positive ultrasounds stratified by age group (preschool age *vs.* common school age *vs.* adolescent age) was highest among the group of common school age (6≤ x <12 years) followed by pre-school age (in order of rising age group: 65.22% *vs.* 76.16% *vs.* 60.74%). For patients with equivocal ultrasound findings, the highest proportion of equivocal ultrasounds was therefore seen among children of adolescent age (12≤ x <18 years) followed by preschool age children (0< x <6 years) (in order of rising age group: 34.78% *vs.* 23.84% *vs.* 39.26%). These findings were found to be statistically significant (p=0.040).

Stratified by ultrasound exam result, more positive ultrasounds were found in overall males than in females (74.86% *vs.* 61.25%) whereas the opposite was true for patients with equivocal ultrasounds (males: 25.14% *vs.* females: 38.75%; p=0.010).

Table 12. – Distribution of ultrasound results stratified by abdominal pain duration

Duration	Positive Ultrasound	Equivocal Ultrasound	p-value
0< x <12h, n (%)	37 (16.2)	11 (10.4)	
12≤ x <24h, n (%)	77 (33.6)	26 (24.5)	
24≤ x <36h, n (%)	48 (21)	32 (30.2)	
36≤ x <48h, n (%)	16 (7)	12 (11.3)	
48≤ x <60h, n (%)	21 (9.2)	6 (5.7)	Group differences: 0.293¥
60≤ x <72h, n (%)	12 (5.2)	7 (6.6)	
72≤ x <84h, n (%)	12 (5.2)	5 (4.7)	
84≤ x <96h, n (%)	3 (1.3)	3 (2.8)	
96≤ x <108h, n (%)	3 (1.3)	4 (3.8)	
108≤ x ≤120h, n (%)	0 (0)	0 (0)	

h – Hours ¥ Pearson's chi-squared test

When analysing abdominal pain duration *(see table 12)*, there was a slight trend of appendicitis cases presenting early with an initially positive ultrasound result with more than 70% (114/151 cases) of all cases presenting within the first 24 hours (classified as 'positive ultrasound'). On closer inspection, almost half of all patients with a positive ultrasound exam on admission (114/229; 49.8%) presented within the first 24 hours after onset of abdominal pain, whereas it was only 34.9% (37/106) of patients with an equivocal ultrasound result presenting under 24 hours. The highest proportion of 'positive ultrasounds' was reported among patients presenting between 12 and 24 hours (33.6%),

closely followed by patients presenting between 24 and 36 hours after onset of abdominal pain (21%).

As abdominal pain duration increased, the proportions of positive ultrasound results stratified by pain duration group slightly decreased (24≤ x <36h: 21%; 36≤ x <48h: 7%) before taking a slight increase after 48 hours (48≤ x <60h: 9.2%) followed by a gradual reduction (60≤ x <72h: 5.2%; 72≤ x <84h: 5.2%; 84≤ x <96h: 1.3%; 96≤ x <108h: 1.3%; 108≤ x ≤120h: 0%). Similarly, with growing abdominal pain duration, the ratio of equivocal results increased at the beginning of the disease (0< x <12h: 10.4%; 12≤ x <24h: 24.5%; 24≤ x <36h: 30.2) followed by a subsequent decrease. However, differences between different groups of abdominal pain duration were not found to be statistically significant (p=0.293) (table 12).

Leukocyte count upon admission was significantly higher for patients presenting with a positive ultrasound result compared to patients with equivocal results (15.06 ± 5.22 x 10^3/µL *vs.* 13.66 ± 5.04 x 10^3/µL; p=0.022). However, this was not the case for the CRP level. Here, the median count for patients with a positive ultrasound was 19.0 mg/l compared with patients with equivocal ultrasound results of 22.50 mg/l, p=0.597. No statistically significant group difference was found for children with regard to subfebrile temperature and fever. Here, children of both groups (positive ultrasound *vs.* equivocal ultrasound) distributed similarly among subgroups (p=0.698).

Time from admission to surgery was measured in hours. Comparison of both groups yielded a statistically significant shorter duration for the group with positive ultrasound findings compared to patients with an equivocal ultrasound result (Median [Q_1; Q_3]; 7h [4h; 15h] *vs.* 17h [7h; 27.75h]; p<0.001). At surgery, open approach to appendectomy was not chosen more often over the laparoscopic approach when comparing both groups (positive ultrasound: 10.1% *vs.* equivocal ultrasound 9.4%, p= 0.420).

Table 13. – Comparison of demographic and clinical characteristics of paediatric patients with positive and equivocal ultrasound upon admission to the emergency department

Criteria	Positive Ultrasound (n=229)	Equivocal Ultrasound (n=106)	p-value
Age (y)			
Mean ± Std. Deviation	11.12 ± 3.26	12.26 ± 3.72	0.005‡
Minimum – Maximum	2.25 – 17.83	2.58 – 17.67	
0< x <6 years, n (%)	15 (65.22)	8 (34.78)	
6≤ x <12 years, n (%)	115 (76.16)	36 (23.84)	0.040§
12≤ x <18 years, n (%)	99 (60.74)	62 (39.26)	
Sex, male, n (%)	131 (74.86)	44 (25.14)	0.010§
Sex, female, n (%)	98 (61.25)	62 (38.75)	
Leukocyte count on admission			
Mean ± Std. Deviation (x 10^3/µL)	15.06 ± 5.22	13.66 ± 5.04	0.022‡
CRP on admission (mg/L)			
Median (Q_1; Q_3),	19 (5.25; 53.50)	22.50 (6.0; 64.75)	0.579†
Fever			
≥38.5°C, n (%)	30 (13.1)	13 (12.3)	
37.0-38.4°C, n (%)	31 (13.5)	18 (17)	0.698§
No Fever, n (%)	127 (55.5)	56 (52.8)	
No documentation, n (%)	41 (17.9)	19 (17.9)	
Time from Admission to OT (h)			
Median (Q_1; Q_3)	7.00 (4.0; 15.0)	17.0 (7.0; 27.7)	<0.001†
Perforation Status:			
Perforated, n (%)	46 (90.2)	15 (9.8)	0.224§
Abdominal Access:			
Laparoscopic, n (%)	204 (89.1)	96 (90.6)	0.420§
Open surgery, n (%)	25 (10.9)	10 (9.4)	

C – Celsius
CRP – C-reactive Protein
§ Fisher's Exact test
h – Hours
‡ Independent-t test
† Mann-Whitney U test
OT – Operating Theatre
¥ Pearson's chi-squared test
Q_1; Q_3 – Quartile 1 and Quartile 3
y – Years

3.3 Logistic Regression Model Predicting a Positive Ultrasound upon Admission

A multiple logistic regression model was fitted to the data to determine factors that are associated with the development of a positive ultrasound result in patients suspected of acute appendicitis (*refer to Methods 2.6 Primary Analysis)*. This model revealed that only age and leukocyte count on admission were statistically significantly associated with developing a positive ultrasound result on admission. For clarity, for an increase in age by one year, the odds of a positive ultrasound upon admission would decrease by 7.06% (OR [95% CI]: 0.92937 [0.86291-0.99899]; p=0.0493) *(see figure 2)*. Likewise, for every increase in leukocyte count by $10^3/\mu l$, the odds of a positive ultrasound on examination would increase by 5.6% (OR [95% CI]: 1.05551 [1.00558-1.10951]; p=0.0308) *(see figure 3)*. When abdominal pain duration is made a continuous variable, rising abdominal pain duration showed a decreasing tendency of exhibiting a positive ultrasound (OR [95% CI]: 0.99536 [0.98456-1.00639]; p=0.4044). However, abdominal pain duration was not significantly associated with a positive ultrasound on admission. Neither were CRP level (p=0.3483), nor female sex (p=0.0587) *(see table 14)*.

Table 14. – Multiple logistic regression of factors potentially associated with the development of a positive ultrasound result upon admission to the emergency department

Parameter	Odds	Odds Ratio	Odds-CI (95%)	p-value
Age	-0.073240	0.92937	0.86291; 0.9989	0.0493
Leukocyte	0.054029	1.05551	1.00558; 1.10951	0.0308
Abdominal Pain Duration	-0.004646	0.99536	0.98456; 1.00639	0.4044
Female Sex	-0.460313	0.63108	0.39050; 1.01613	0.0587
CRP	-0.002045	0.99795	0.99368; 1.00229	0.3483

CRP – C-reactive Protein *CI* – Confidence Interval

Figure 2. – Age as predictor of 'positive ultrasound'

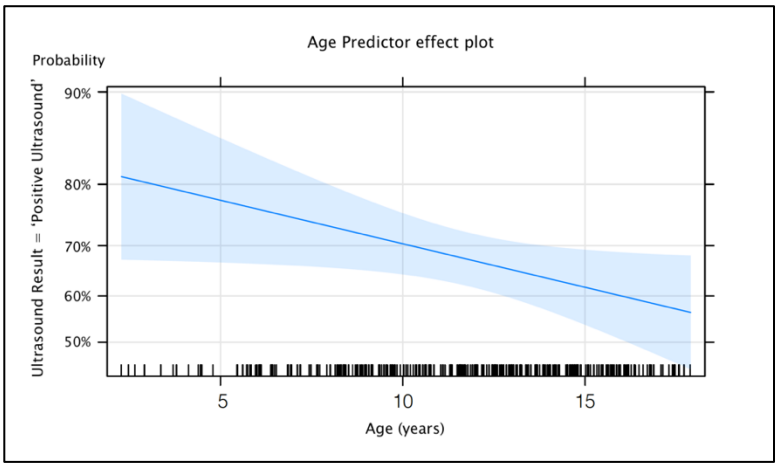


Figure 3. – Leukocyte count as predictor of 'positive ultrasound'

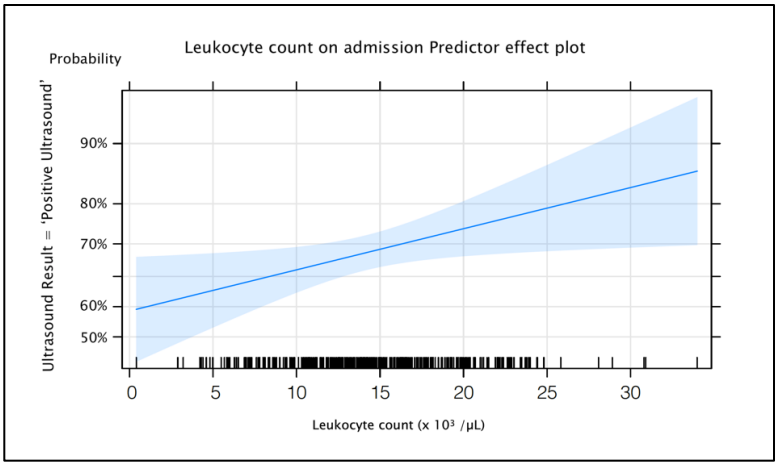


Figure 4. – Abdominal pain duration as predictor of 'positive ultrasound'

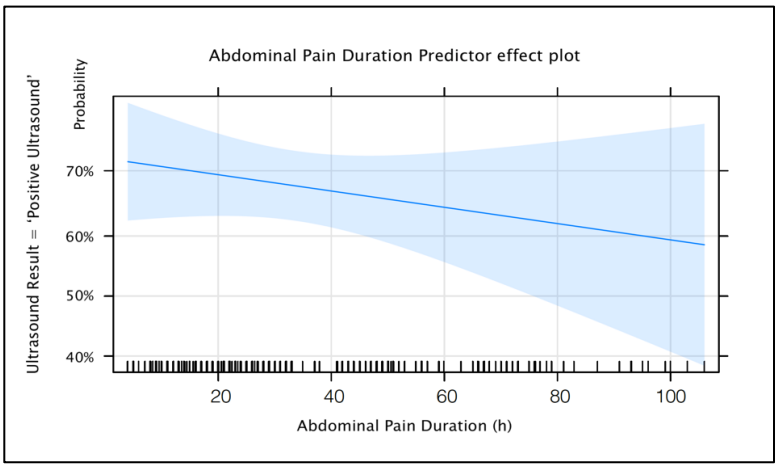
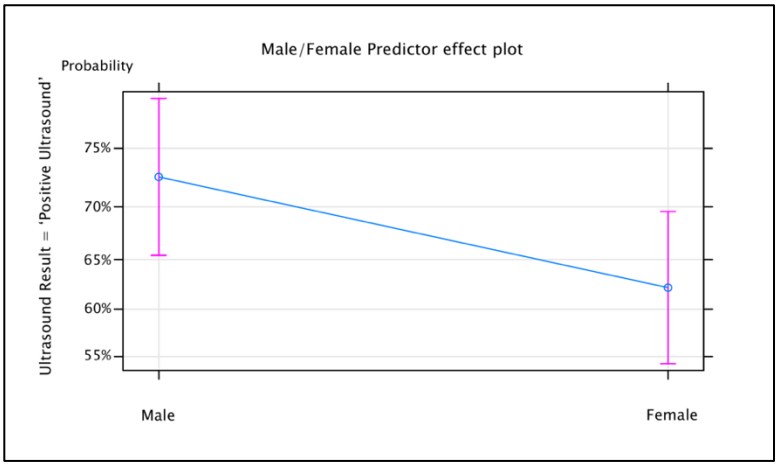


Figure 5. – Sex as a predictor of 'positive ultrasound'



3.4 Comparison of Perforated and Non-Perforated Paediatric Appendicitis

When comparing perforated with non-perforated appendicitis *(see table 16)*, children were statistically younger in the group of perforated compared to non-perforated appendicitis (10.14 ± 3.92 years *vs.* 11.81 ± 3.28 years; p=0.003). Stratified by age group, the proportion of perforation was highest among preschool-children (<6 years of age) and subsequently decreased with rising patient age (0< x <6 years: 39.1%; 6≤ x <12 years: 17.2%; 12≤ x <18 years: 16%) *(see table 16)*. The opposite was true for patients who did not have a perforation as detected by histology. Here, the proportion of non-perforation subsequently rose with rising patient age (0< x <6 years: 60.9%; 6≤ x <12 years: 82.8%; 12≤ x <18 years: 84%) *(see figure 6)*. These group differences were found to be statistically significant (p=0.024).

Figure 6. – Patient age distribution and proportion of patients with and without perforation

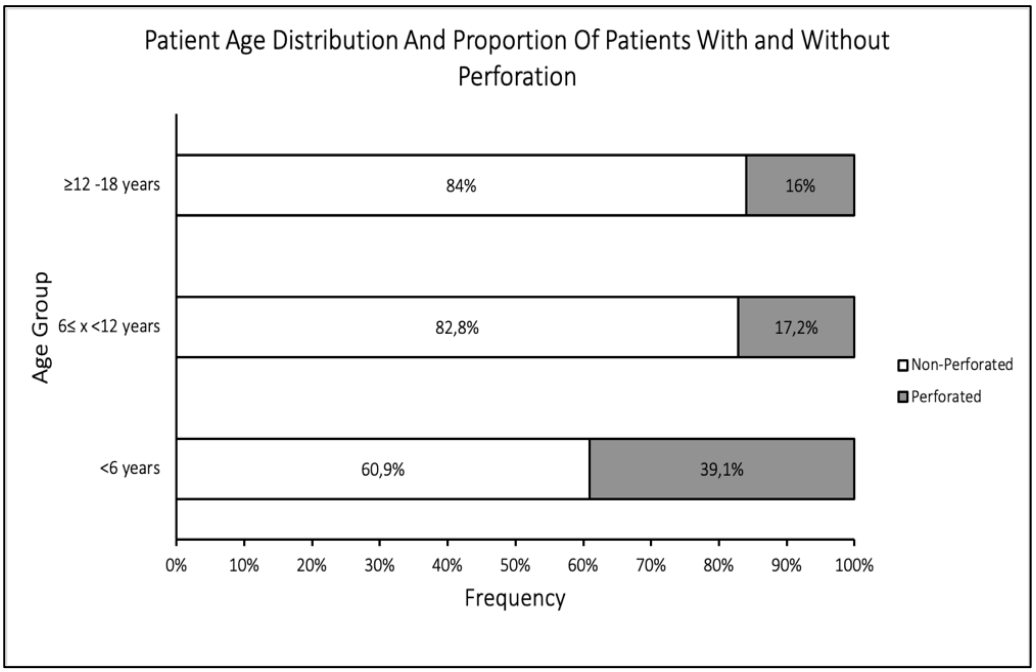
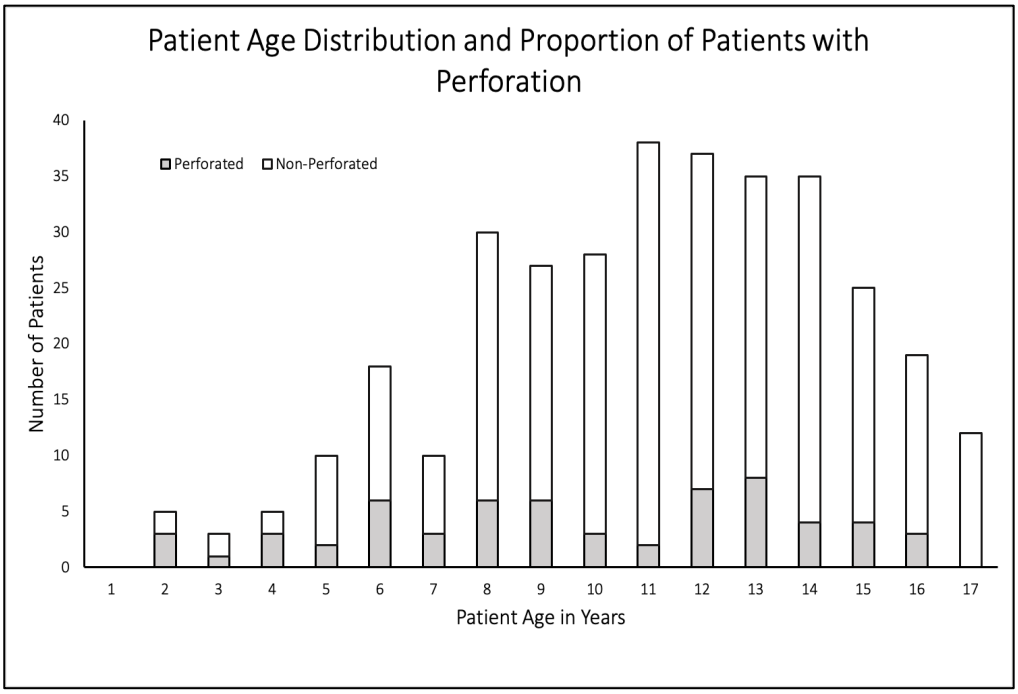


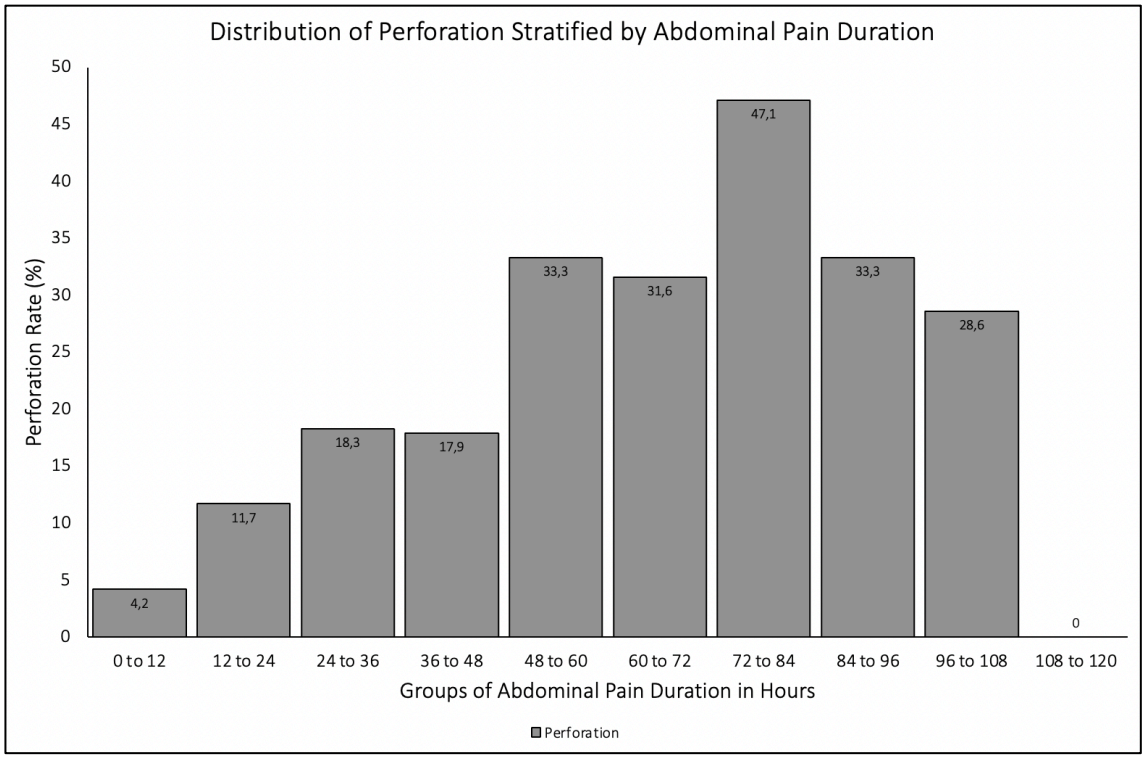
Figure 7 below depicts the perforation rate by age category (2-17 years) in our cohort of perforated appendices (n=61). As patient age rises, perforation rate decreases. The perforation rate for children aged 2 to 5 years was between 20-60% (n=9). After decreasing to 5-33% (n=26) for the ages 6 to 11, the appendiceal perforation rate further declined to 0-23% (n=26) for children aged 12 to 17 years. Regarding sex, there were slightly more females associated with appendiceal perforation. Among all males, 16.5% showed appendiceal perforation at surgery, whereas this was true for 19.9% of all females (p=0.497). The ratio of positive ultrasound results upon admission was higher among patients with later appendiceal perforation (75.4% *vs.* 66.3%; p=0.338). However, this finding was not found to be statistically significant.

Figure 7. – Patient age distribution and proportion of patients with perforation



With special focus on abdominal pain duration, there was a clear trend of appendiceal perforation later in the course of abdominal pain as abdominal pain duration increased *(see figure 8)*. The lowest proportion of perforation (4.2%) was seen right at the beginning of the abdominal pain course, in children presenting with <12 hours after onset of abdominal pain (n=2). Thereafter, the proportion gradually increased with the highest proportion for appendiceal perforation (47.1%) seen among children presenting between 72 and 84 hours after onset of abdominal pain. Conversely, the proportion of non-perforation was highest among children presenting in the earlier abdominal pain duration groups ($0< x <12h$: 95.8%; $12\leq x <24h$: 88.3%) and decreased gradually as abdominal pain duration increased. These group differences were found to be statistically significant ($p<0.001$).

Figure 8. – Distribution of perforation and non-perforation stratified by abdominal pain duration



Among children with no detected fever throughout the clinical course or on admission before presenting to ED *(see table 16)*, 9.2% matched the group of perforated appendicitis whereas 90.8% were attributed to the group of non-perforated appendicitis. Among children with subfebrile temperature (37.0-38.4°C), more than a third (34.7%) were found in the group of perforated appendicitis. The proportion increased with rising temperature where up to 38.6% of patients with perforated appendicitis upon surgery/histology had shown signs or symptoms of fever (38.5°C). These group differences tested statistically significant ($p<0.001$).

Patients who turned out to show appendiceal perforation showed higher inflammatory counts on admission, compared to patients without appendiceal perforation (Leukocytes: 16.66 ± 5.52 x $10^3/\mu L$ *vs.* 14.12 ± 5.04 x $10^3/\mu L$; $p=0.001$) *(see table 16)*. Likewise, CRP level on admission was equally higher (median [Q_1; Q_3]; 74.0 [27.0; 163.5] mg/L *vs.* 15.0 [4.0; 42.5] mg/L) in cases of perforated ($p<0.001$) compared to non-perforated cases.

Table 15. – Median (Q_1; Q_3) CRP level (mg/L) on admission for cases with and without appendiceal perforation, stratified by abdominal pain duration

Duration	Perforated Appendicitis	Non-Perforated Appendicitis	p-value
Total	74.0 (27.0; 163.5)	15.0 (4.0; 42.5)	<0.001†
0< x <12h	55.0 (15.0; 95.0)	5.0 (2.0; 10.0)	
12≤ x <24h	21.5 (11.0; 29.5)	13.0 (3.0; 35.0)	
24≤ x <36h	35.0 (16.0; 122.0)	25.0 (6.5; 58.0)	
36≤ x <48h	83.0 (48.5; 149.5)	14.0 (6.0; 75.0)	
48≤ x <60h	155.0 (120.0; 215.5)	36.0 (16.0; 63.0)	
60≤ x ≤72h	69.5 (54.0; 91.5)	57.0 (7.5; 122.5)	
72≤ x ≤84h	153.0 (59.0; 243.8)	11.0 (6.0; 35.5)	
84≤ x ≤96h	247.0 (180.0; 314.0)	3.5 (2.0; 35.8)	
96≤ x ≤108h	147.0 (95.0; 199.0)	2.0 (2.0; 13.0)	
108≤ x ≤120h	-	-	

CRP – C-reactive Protein
h -Hours
† Mann-Whitney U test
Q_1; Q_3– Quartiles 1, Quartiles 3

Further, patients with later confirmed appendiceal perforation were found not only to have higher median CRP levels on admission in virtually all abdominal pain categories *(see table 15)*, also median CRP level in this group was found to be on the rise as abdominal pain duration increased. The highest median CRP level detected in the group of perforated appendicitis was found at 247.0 [180.0; 314.0] mg/L while patients with non-perforated appendicitis never went beyond a median CRP level of 57.0 [7.5; 122.5] mg/L.

Comparison of both groups regarding time from admission to operation yielded a statistically significant difference. In children with later perforated appendicitis, median time to OT was shorter (median [Q_1; Q_3]; 6h [3.5h; 16.5h] *vs.* 9h [5h; 21h]) compared to non-perforated appendicitis (p=0.008). Following the start of surgery, the vast majority of patients in either group were operated using a laparoscopic access to the abdomen. However, when comparing each group against each other, open approach to appendectomy was performed more often in patients with perforated appendicitis (24.6% *vs.* 7.25%; p<0.001).

Table 16. – Comparison of demographic and clinical characteristics of paediatric patients with and without appendiceal perforation

Criteria	Perforated (n=61)	Non-Perforated (n=276)	p-value
Age (y)			
Mean ± Std. Deviation	10.14 ± 3.92	11.81 ± 3.28	0.003‡
Minimum – Maximum	2.42 – 16.75	2.25 – 16.75	
0< x <6 years, n (%)	9 (39.1)	14 (60.9)	
6≤ x <12 years, n (%)	26 (17.2)	125 (82.8)	0.024§
12≤ x <18 years, n (%)	26 (16)	137 (84)	
Sex, male (%)	29 (16.5)	147 (83.5)	0.479§
Sex, female (%)	32 (19.9)	129 (80.1)	
Positive Ultrasound, n (%)	46 (75.4)	183 (66.3)	
Equivocal Ultrasound, n (%)	15 (24.6)	91 (33)	0.338§
Negative Ultrasound, n (%)	0 (0)	2 (0.7)	
Leukocyte count on admission			
Mean ± Std. Deviation (x 10^3/μL)	16.66 ± 5.52	14.12 ± 5.04	0.001‡
CRP on admission (mg/L)			
Median (Q_1; Q_3),	74.0 (27.0; 163.5)	15.0 (4.0; 42.50)	<0.001†
Fever			
≥38.5°C, n (%)	17 (38.6)	27 (61.4)	
37.0-38.4°C, n (%)	17 (34.7)	32 (65.3)	<0.001§
No Fever, n (%)	17 (9.2)	167 (90.8)	
No documentation, n (%)	10 (16.7)	50 (83.3)	
Time from Admission to OT (h)			
Median (Q_1; Q_3)	6.0 (3.5; 16.5)	9.0 (5.0; 21.0)	0.008†
Abdominal Access:			
Laparoscopic, n (%)	46 (75.4)	256 (92.75)	<0.001§
Open surgery, n (%)	15 (24.6)	20 (7.25)	

C – Celsius
CRP – C-reactive Protein
§ Fisher's Exact test
h – Hours
‡ Independent-t test
† Mann-Whitney U test
OT – Operating Theatre
¥ Pearson's chi-squared test
Q_1, Q_3 – Quartile 1 and Quartile 3
y – Years

3.5 Logistic Regression Model Predicting Appendiceal Perforation

To test for factors associated with the outcome of appendiceal perforation when presenting to ED, another logistic regression model was created (*see table 17*). Co-variables to be tested included age, sex, abdominal pain duration, leukocyte and CRP level on admission as well as ultrasound result on admission.

Of all co-variables only CRP level (OR [95% CI]: 1.012 [1.00714-1.01813]; p<0.001) and abdominal pain duration (OR [95% CI]: 1.019 [1.00441-1.03392]; p<0.001) were found to be statistically significantly associated with predicting appendiceal perforation later in the course of the disease. Figure 9 illustrates the distribution of abdominal pain duration in both perforated and non-perforated groups. Both figures 10 and 11 illustrate the changing probability of appendiceal perforation as CRP and abdominal pain duration increase, respectively. However, age, leukocyte count on admission, female sex as well as USS result under ED conditions were not significantly associated with appendiceal perforation.

Table 17. – Multiple logistic regression of factors associated with appendiceal perforation

Parameter	Odds	Odds Ratio	Odds-CI (95%)	p-value
Age	-0.079876	0.9232	0.84046; 1.01268	0.091500
Leukocyte	0.050607	1.052	0.98924; 1.11980	0.107409
Abdominal Pain Duration	0.018946	1.019	1.00441; 1.03382	<0.001
Female Sex	0.297162	1.346	0.70273; 2.59083	0.369799
CRP	0.012319	1.012	1.00714; 1.01813	<0.001
Equivocal Ultrasound Result	-0.526504	0.5901	0.27646; 1.19816	0.156656

CRP – C-reactive Protein *CI* – Confidence Interval

Figure 9. – Distribution of abdominal pain duration among perforated and non-perforated cases

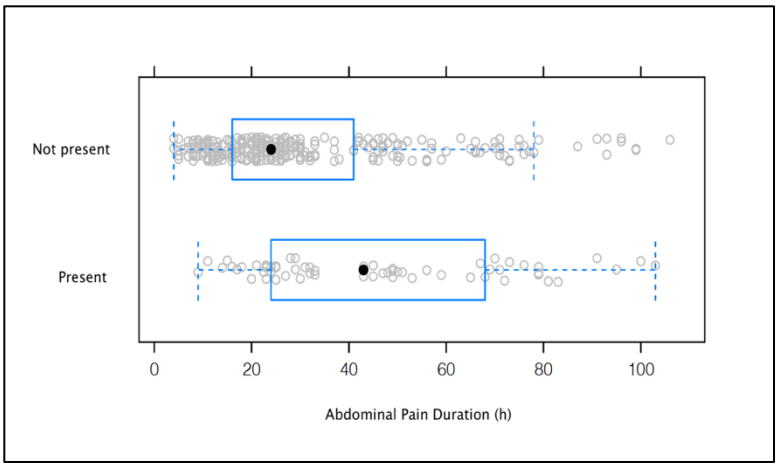


Figure 10. – Abdominal pain duration predicting appendiceal perforation

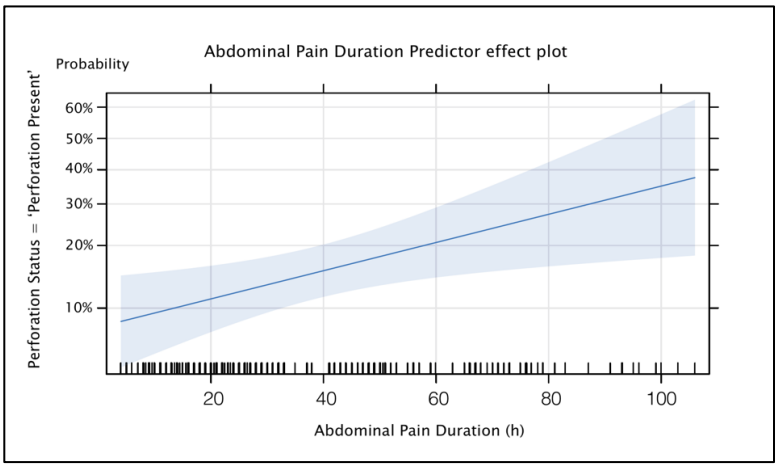
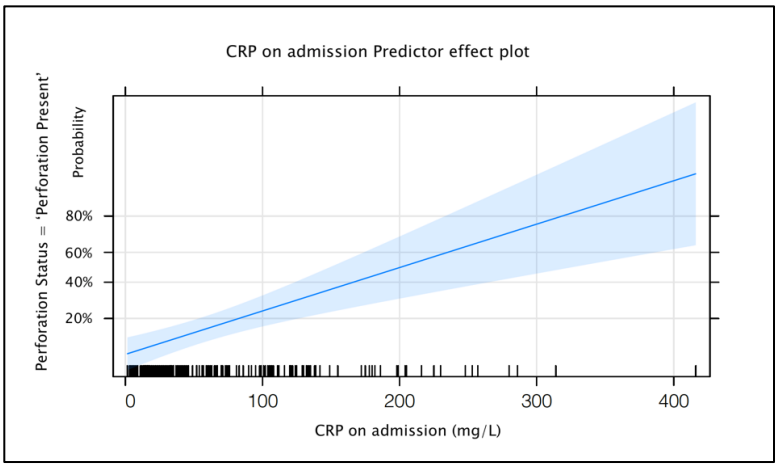


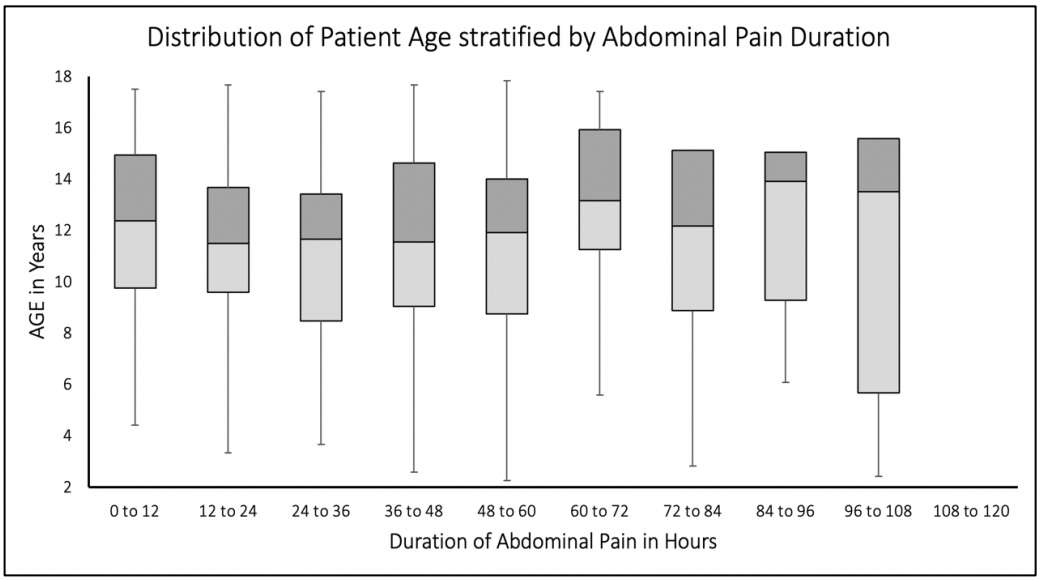
Figure 11. – CRP level predicting appendiceal perforation



3.6 Effect of Age on Abdominal Pain Duration

Figure 12 illustrates the distribution of patient age for each abdominal pain duration category given for the overall cohort (n=337). Among all abdominal pain duration groups, patient mean age varied between 11.04 and 12.05 years with an overall mean patient age of 11.51 ± 3.46 years. One-Way ANOVA for inter-group differences revealed no statistically significant differences (p=0.688). Neither did Spearman's Rank Correlation Coefficient (p=0.932).

Figure 12. – Distribution of patient age stratified by abdominal pain duration



3.7 Effect of Pain Duration on Ultrasound Results in Perforated Appendicitis

Comparing cases showing a positive ultrasound on admission with regards to their later perforation status yielded significant group differences when stratifying them according to abdominal pain duration *(table 18)*. The ratio of cases with later perforation found to be 'positive' on ultrasound sonography upon admission was observed to be gradually growing with increasing abdominal pain duration. No cases of later appendiceal perforation presented below 12 hours of abdominal pain duration. Taken together, these group differences turned out to be statistically significant (p<0.001).

Table 18. – Distribution of positive ultrasound results in perforated and non-perforated appendicitis stratified abdominal pain duration

Duration	Perforated	Non-Perforated	p-value
0< x <12h	0 (0)	37 (100)	
12≤ x <24h	11 (14.3)	66 (85.7)	
24≤ x <48h	11 (22.9)	37 (77.1)	
48≤ x <60h	2 (12.5)	14 (87.5)	
60≤ x <72h	6 (28.6)	15 (71.4)	Group differences: <0.001¥*
72≤ x <84h	5 (41.7)	7 (58.3)	
84≤ x <96h	8 (66.7)	4 (33.3)	
96≤ x <108h	1 (33.3)	2 (66.7)	
108≤ x ≤120h	2 (66.7)	1 (33.3)	

¥ Pearson's chi-squared test *Group differences of categorical pain durations between perforated and non-perforated cases of appendicitis

3.8 Factors Leading to a Short Latency to Operation

Average time from admission to OT (latency) was revealed as statistically significantly different when comparing perforated *vs.* non-perforated cases, as well as positive *vs.* equivocal ultrasounds on admission. Hence, latency was shorter in patients exhibiting a positive, diagnostic ultrasound on admission, as well as those patients showing later perforated appendicitis on histology. Latency was therefore shortest in patients exhibiting both features (positive ultrasound + appendiceal perforation). One question of significant interest was whether there are other factors that are associated with a short latency (<6 hours) after arrival to ED *(table 19)*.

Table 19. – Comparison of patients with short (<6h) and long (≥6) latency to operating theatre

Criteria	Latency <6h (n=106)	Latency ≥6h (n=229)	p-value
Age (y)			
Mean ± Std. Deviation	10.82 ± 3.51	11.79 ± 3.38	0.016‡
0< x <6 years, n (%)	13 (12.3)	10 (4.4)	
6≤ x <12 years, n (%)	49 (46.2)	102 (44.5)	0.018§
12≤ x <18 years, n (%)	44 (41.5)	117 (51.1)	
Sex, male (%)	55 (51.9)	120 (52.4)	0.512§
Sex, female (%)	51 (48.1)	109 (47.6)	
Ultrasound			
Positive, n (%)	84 (79.2)	145 (63.3)	0.002§
Equivocal, n (%)	22 (20.8)	84 (36.7)	
Duration of abdominal pain (h)			
0< x <12, n (%)	13 (12.3)	35 (15.3)	
12≤ x <24, n (%)	29 (27.4)	74 (32.3)	
24≤ x <36, n (%)	27 (25.5)	53 (23.1)	
36≤ x <48, n (%)	4 (3.8)	24 (10.5)	0.065¥
48≤ x <60, n (%)	13 (12.3)	14 (6.1)	
60≤ x <72, n (%)	6 (5.7)	13 (5.7)	
72≤ x <84, n (%)	11 (10.4)	6 (2.6)	
84≤ x <96, n (%)	1 (0.9)	5 (2.2)	
96≤ x <108, n (%)	2 (1.9)	5 (2.2)	
108≤ x ≤120, n (%)	0 (0)	0 (0)	
Fever			
≥38.5°C, n (%)	15 (14.2)	28 (12.2)	
37.0-38.4°C, n (%)	17 (16)	32 (14.0)	
No Fever, n (%)	57 (53.8)	126 (55)	0.831§
No documentation, n (%)	17 (16)	43 (18.8)	
Leukocyte (x 10^3/µL)			
Mean ± Std. Deviation	16.12 ± 5.02	13.92 ± 5.14	<0.001‡
CRP (mg/L)			
Median (Q_1; Q_3),	30.0 (8.5; 105.5)	15.0 (5.0; 50.5)	0.001†

C – Celsius
CRP – C-reactive Protein
§ Fisher's Exact test
h – Hours
‡ Independent-t test
† Mann-Whitney U test
OT – Operating Theatre
¥ Pearson's chi-squared test
Q_1; Q_3 – Quartile 1 and Quartile 3
y – Years

Inferential statistics comparing both groups of short and long latency showed that patients with short latency were often times younger (10.82 ± 3.51 *vs.* 11.79 ± 3.38 years, p=0.016) and more frequently had a positive, diagnostic USS exam when presenting to ED (79.2% *vs.* 63.3%; p=0.002). Also, both CRP level and leukocyte count on admission

were statistically significantly different in both groups, whereby patients of short latency exhibited higher inflammatory markers compared to those of longer latency (leukocyte p<0.001; CRP p=0.001). Sex, abdominal pain duration as well as temperature distributed similarly in both groups and did not show statistically significant group differences.

3.9 Logistic Regression Model Associated with Short Latency to Operation

In order to determine factors that would allow prediction of a short latency to OT, a logistic regression model was fitted. Input variables incorporated into this model were age, gender, CRP level on admission, abdominal pain duration, ultrasound result on admission and leukocyte count on admission (*see table 20*). Presence of fever and subfebrile temperature (≥37.0°C) were not taken into this regression model as the amount of missing data was too high to allow an adequate assertion about the effect of fever and subfebrile temperature on the short latency prediction model.

Of these co-variables, only leukocyte count and equivocal ultrasound result upon admission were statistically significantly associated with short latency (<6 hours) to OT. On closer inspection, increasing leukocytes were associated with increasing odds of reaching OT in 6 hours or less (OR [95% CI]: 1.068 [1.016-1.122]; p=0.010), while an equivocal ultrasound on admission decreased odds of prompt appendectomy (OR [95% CI]: 0.472 [0.268-0.833]; p=0.010). Besides, increasing patient age and male gender were both associated with decreasing the odds of early appendectomy, while both increasing CRP level on admission and increasing abdominal pain duration were both associated with prompt appendectomy. However, these findings were not found to be statistically significant.

Table 20. – Multiple logistic regression of factors associated with short latency (<6 hours) to OT

Factors	Odds	Odds Ratio	95% CI	p-value
Age	-0.041	0.960	0.892; 1.033	0.275
Male Gender	-0.131	0.877	0.536; 1.435	0.601
CRP Level upon Admission	0.003	1.003	0.999; 1.088	0.140
Abdominal Pain Duration	0.007	1.007	0.995; 1.019	0.250
Equivocal USS on Admission	-0.751	0.472	0.268; 0.833	0.010
Leukocyte Count upon Admission	0.065	1.068	1.016; 1.122	0.010

CI – Confidence Interval
CRP – C-reactive Protein
OT – Operating Theatre
USS – Ultrasound Sonography
The reference category is: ≥6 Hours

3.10 Sensitivity, Specificity and Positive Predictive Value

Formula **[4]** below describes the sensitivity of the ultrasound detecting true cases of appendicitis with positive test (conclusive USS) characteristics in all individuals in whom appendicitis was truly ruled in. The sensitivity was calculated to be 0.6795 *(table 21)*.

[4] Sensitivity

$$= \frac{\text{Number of True Positives}}{\text{Number of True Positives} + \text{Number of False Negatives}} = \frac{229}{229 + 108} = 0.6795$$

Formula **[5]** below describes the specificity of the ultrasound differentiating between all numbers of true negative cases and negative test characteristic and total number of individuals in whom appendicitis was truly ruled out. The specificity was calculated 1.0 *(table 21)*.

[5] Specificity

$$= \frac{\text{Number of True Negatives}}{\text{Number of True Negatives} + \text{Number of False Positives}} = \frac{2}{2 + 0} = 1$$

Formula **[6]** below describes the positive predictive value of the ultrasound in our cohort. This gives an estimate as to what chance patients with a positive ultrasound result will have the condition (appendicitis). Due to virtually no existing false positive cases, the positive predictive value was computed as 1.0 *(table 21)*.

[6] Positive Predictive Value (PPV)

$$= \frac{\text{Sensitivity x Prevalence}}{\text{Sensitivity x Prevalence} + (1 - \text{Specificity}) \text{ x } (1 - \text{Prevalence})}$$

$$= \frac{0.6795 \text{ x Prevalence}}{0.6795 \text{ x Prevalence} + (1 - 1) \text{ x } (1 - \text{Prevalence})} = \frac{0.6795 \text{ x Prevalence}}{0.6795 \text{ x Prevalence}} = 1$$

Table 21. – Two-by-Two contingency table

	Positive Histology	Negative Histology	Total
DIAGNOSTIC USS	229	2	231
NON-DIAGNOSTIC USS	108	6	114
Total	337	8	345

4.0 Discussion

It is the irony of recent medical advances that despite innovative improvements in medical imaging and therapeutics, the condition most commonly requiring immediate surgery in children still poses a diagnostic challenge. Over the years, rates of complicated appendicitis have fairly stayed unchanged, and many questions remain unanswered[3,10,12,19,25,26,82-84]. Especially in the young, history taking is often difficult and vague and remains a challenge in those who cannot speak for themselves. Delayed presentation and misdiagnosis of appendicitis translate into higher incidence rates of appendiceal perforation and significant comorbidity[16,85].

For children presenting with a classic 'textbook' description of acute appendicitis, the decision to proceed with operative care without advanced or additional imaging may be straightforward. Yet, children with prolonged symptom duration and findings suspicious of appendiceal perforation oftentimes require diagnostic imaging to rule out other pathologies. Major outcomes for paediatric cases of suspected appendicitis at the emergency department relate to timely diagnosis, minimising missed cases of appendicitis, avoiding misdiagnosis leading to negative appendectomy and proper identification of appendicitis before progressing to perforation. For each of these outcomes, diagnostic imaging serves a major role in cases of clinical uncertainty[5].

We know that ultrasound visualisation of the appendix depends on many factors including but not limited to abdominal tenderness, guarding, adipose tissue as well as the proficiency of the sonographer[86,87]. In uncertain cases, laboratory work as well as the clinical exam may drive the decision to take a child to the operating theatre or observe the child overnight for serial abdominal examinations and repeat laboratory testing. Additional imaging such as CT or MRI scans may be considered for cases with a high index of suspicion for appendicitis where ultrasound examinations show equivocal results or remain inconclusive overall[82]. In spite of recent advances in CT imaging, better resolution and quicker performance times, it has become a great challenge to replace CT with ultrasound over increasing concerns radiation poses to children whose tissues are still growing and their accumulative lifetime radiation exposure[18,82]. Consequently,

efforts have been made to minimise CT utilisation in favour of increased reliance on ultrasonography[6,9,50,88,89].

It appears comprehensible that more advanced cases of appendicitis should be more visible by ultrasonography, yet little research has been conducted regarding the test characteristics of diagnostic and non-diagnostic USS results according to the duration of abdominal pain[5,90].

In the present study we retrospectively investigated USS performance in a cohort consisting of histologically proven cases of appendicitis only and tried to improve the diagnostic validity of USS by correlating positive, diagnostic USS results with the time-dependent development of radiological features and clinical parameters in acute appendicitis. By doing so, we attempted to reveal and reduce the limitations of USS and guide the clinician in his decision-making process when encountering radiologically equivocal cases of acute appendicitis. The findings obtained in this study shall help to decrease negative appendectomies performed in children and further help to reduce the appendiceal perforation rate and its complications.

4.1 Factors associated with a positive ultrasound result on admission

The setup to this research question were which factors and clinical parameters are associated with the development of a positive, diagnostic USS result on admission in children presenting with symptoms of acute appendicitis. In our cohort of 335 children, positive USS results were detected more frequently with children presenting in the earlier phases of the condition after onset of abdominal pain with a decreasing trend of developing a positive, diagnostic USS result thereafter. Yet, we could not find a statistically significant positive association of abdominal pain duration with the development of a positive, diagnostic USS on admission suggesting that serial abdominal USS exams would not necessarily bring more clarity if used as a sole and only diagnostic modality to detect appendicitis.

The sensitivity of USS in our study to correctly diagnose appendicitis upon admission was 68%. While the setup of our study population may limit the prospect on the real-world scenario as we intentionally refrained from including non-diagnostic USS cases of non-appendicitis cases, we decided to choose this retrospective approach of positive,

diagnostic USS cases to further define and characterise this imaging modality for appendicitis-specific cases only. Moreover, the diagnostic capacity of ultrasound goes along with a probability of error due to false-positive results. Yet, this factor is negligible in our study design as the study focus is set on positive, diagnostically conclusive USS cases with histological proof. This cohort was intentionally kept free from other comorbidities mimicking the presence of appendicitis in order to make a definitive statement about the diagnostic significance of ultrasound in paediatric appendicitis. All positive, diagnostic ultrasound results therefore confirmed the presence of appendicitis. Vast studies have in the past described the sensitivity of ultrasound in paediatric appendicitis with reported values ranging from 72.5-100% but failed to exclude patients with conditions other than appendicitis that gave rise to abdominal pain[5,6,49,50,52,53,91]. Their use of abdominal sonography was therefore not always for confirmatory purposes. Comparing the sensitivity obtained in our cohort to previously published studies therefore needs to be treated with caution. Yet, our reported sensitivity of 68% is close to that of previous reported studies[49].

These aforementioned studies have shown that the sensitivity of ultrasound to differentiate between cases of appendicitis and non-appendicitis on admission increases linearly as abdominal pain duration increases. Specifically, Bachur *et al.* (2012) showed in a cohort of 832 children with suspected appendicitis that ultrasonographic sensitivity increased from 81% in the first 12 hours of pain to 96% after 48 hours of pain (test for trend: OR [95% CI]: 1.39 [1.14-1.71]). Similarly, the negative predictive value increased as abdominal pain duration increased with a marked improvement after reaching the threshold of 36 hours after onset of abdominal pain[5]. Yet, the used cohort only showed a prevalence of proven appendicitis of 38% which makes transferring and comparing their results with ours difficult as our cohort showed a prevalence of 100%. Likewise, a retrospective analysis by Partain *et al.* (2016) including 271 children with non-diagnostic USS, symptom duration of 3 or more days was associated with 2.3 times greater odds of finding secondary signs of appendicitis compared to patients of just 24-hour symptom duration[92,93]. Malia *et al.* (2019) showed in their retrospective study of 490 children, whose appendix could not be visualised, that the odds of appendicitis were 66% lower for patients presenting with more than 3 days of abdominal pain (OR [95% CI]: 0.34 [0.003-0.395])[82,91]. However, in this group of non-diagnostic ultrasounds (n=490) only 6.7% of children were diagnosed with appendicitis suggesting that other pathologies might have

masked the acute symptoms of appendicitis. Regardless, these studies clearly highlight a temporal relationship of ultrasound sensitivity and abdominal pain duration.

Yet, despite lacking statistical significance on group comparison in our cohort, we could demonstrate that over 70% of all positive, diagnostic ultrasound results presented in under 36 hours (overall, 50% of positive ultrasound results presenting in under 24 hours). Moreover, the positive detection rate on ultrasound decreased in intervals of 12 hours which gives enough reason to assume that the likelihood of getting hold of a meaningful (diagnosis defining) ultrasound result in patients with later proven appendicitis is higher if performing the ultrasound early on in the disease process. Our findings showed that the differentiation between diagnostic and equivocal ultrasound results becomes gradually more complex after 60 hours of abdominal pain duration, with better results expected below 24 hours after onset of abdominal pain. In fact, other causes of abdominal pain were detected far less frequently in our cohort of young patients as all patients had later proven appendicitis, thereby reducing the influence of other comorbidities on the clinical presentation. Sonographies performed beyond this threshold could make it more difficult to reach a clinically and diagnostically sound conclusion. We therefore argue for a speedy performance of abdominal ultrasound, if possible, within 60 hours after onset of abdominal pain, in patients presenting with acute symptoms of appendicitis.

In particular, given that the largest proportion of positive, diagnostic ultrasounds were detected in school-age children (n=115), we were able to depict what we assume to be a 'classic' course of appendicitis in this age group with the greatest overall absolute percentage increase of positive, diagnostic ultrasounds (+59.1%) between 12-36 hours after onset of abdominal pain. On the basis of the aforementioned studies which confirmed an increasing diagnostic value of ultrasound with rising abdominal pain duration, we would therefore advocate serial abdominal ultrasounds in this age group between 12 and 36 hours so that chances of diagnostically conclusive ultrasound findings can be maximised in children presenting with an initial equivocal ultrasound upon admission to ED. Yet, our retrospective study design was not able to reproduce this assumption objectively and therefore needs to be interpreted with caution.

Another factor that is associated with a positive, diagnostic ultrasound result was age. We could demonstrate statistically significant differences between groups of positive and

equivocal ultrasound results with regards to patient age (11.12 ± 3.26 *vs.* 12.26 ± 3.72 years; p=0.005*)*. Children with a positive, diagnostic ultrasound exam were on average more than a year younger compared to those with equivocal ultrasound results. Our finding is in accordance with a prospective study by Malia *et al.* (2019) who used a cohort of 1252 children to assess the diagnostic accuracy of laboratory and ultrasound findings in patients with a visualised appendix. Overall, their subgroup of visualised appendices showed a younger age on admission compared to non-visualised cases of appendicitis (mean age: 10.2 ± 4.1 *vs.* 11.7 ± 4.4 years; p=0.05)[82]. However, their cohort included both confirmed (n=225; 18%) and unconfirmed cases of appendicitis.

Of the two children in our cohort with negative ultrasounds on admission, they presented at 30 and 33 hours after pain onset, respectively. None of them exhibited later perforation. Even further, in our cohort with only confirmed cases of appendicitis and positive ultrasound (n=229), we were able to identify age-specific group differences. Among pre-school children (<6 years of age) and common-school age children (≥6-12 years of age), positive, diagnostic ultrasound detection rates were 65.22% and 76.16%, respectively, with a decrease in detection in adolescents (≥12-18 years of age) at 60.74%. These findings were found to be statistically significant (p=0.040) suggesting that reaching a diagnostic ultrasound result may be easier in children of common-school age. This finding is partially conflicted by the fact that our logistic regression modelling revealed decreasing odds of finding a positive ultrasound as patient age rose (OR [95% CI]: 0.93 [0.86-0.99]; p=0.049), reflecting on the different pre-existing conditions of appendicitis in pre-school children. Studies show an overall higher incidence of complicated appendicitis in pre-school children due to the fact that contagious co-infections such as measles, chicken pox, and other viral or bacterial diseases are more prevalent and exacerbate a rapid development of appendicitis. This later leads to subsequent premature perforation as they do not exhibit the same immune resistance compared to older peers[2,90,94]. Higher perforation rates on surgery and a more complicated development can therefore be assumed to translate into greater chances of positive ultrasound detection upon admission to ED in this age-group.

In cases where uncertainty remains, secondary predictors can play an important role in helping the emergency physician and paediatric surgeon provide a fully informed evaluation of the presenting case. We could show in our cohort that children with a

positive, diagnostic ultrasound exam on admission had significantly higher levels of leukocyte counts compared to children with an equivocal ultrasound result (15.06 ± 5.22 x 10^3/µL *vs.* 13.66 ± 5.04 x 10^3/µL; p=0.022), yet CRP levels did not show significant group differences on admission. These results may be explained by the pathophysiological response that takes place during the inflammatory process[85]. In most cases, appendicitis is secondary to bacterial infection with an invading pathogen activating the innate immune system, thereby stimulating the bone marrow to produce and release leukocytes via a cascade of various cytokines and inflammatory mediators[85,95]. Blood serum levels of leukocytes and neutrophils increase significantly following bacterial infection, later followed by neutrophils exiting from peripheral vasculature and migrating to the site of infection. In later stages of appendicitis, neutrophil consumption often exceeds its production, resulting in decreased peripheral leukocyte counts and neutrophil percentages[96]. Hence, white cell and neutrophil counts have been associated with high sensitivity in early but not later stages of appendicitis[85]. Lastly, CRP is an acute phase protein that is generated in the liver. With an average doubling time during an infection of approximately 8 hours, CRP requires almost 24 hours to exceed its cut-off value of 5mg/L from an average baseline level of 0.8mg/L[97]. As CRP will continue to rise in the presence of inflammation, it is an excellent marker for delayed and severe inflammation[85,98].

In our study, logistic regression revealed that a higher leukocyte count was independently associated with greater odds of showing a positive USS result (OR [95% CI]: 1.055 [1.005-1.109]; p=0.0308) whereas CRP was not associated with developing a positive ultrasound result on admission (OR [95% CI]: 0.997 [0.9937-1.0023]; p=0.348). This finding is in line with current studies whereby WBC had better diagnostic performance in differentiating appendicitis from non-appendicitis[85,94,99]. Chiang *et al.* (2020) found that leukocyte and neutrophil counts demonstrated high sensitivity for early appendicitis, whereas CRP performed better in the later stages of the condition. Taken all three inflammatory markers together, their data demonstrated a sensitivity of acute appendicitis exceeding 99%[85]. Other studies have assessed the influence of inflammatory markers on the predictive value of ultrasound sonography in children with suspected appendicitis[100]. Anandalwar *et al.* (2014) used a retrospective single centre study of 845 children and demonstrated the negative predictive value (NPV) of a non-diagnostic USS ranging from 41.9% to 95.8% depending on the presence or absence of leukocytosis (WBC >9 x

$10^3/\mu L$)[32]. Similar findings were reported in another retrospective study of 1383 patients, where a normal leukocyte count (WBC <7.5 x $10^3/\mu L$) increased the NPV of an equivocal USS result from 86.3% to 98.9%[83], as well as in the study by Malia *et al.* (2018) where the presence of leukocytosis (WBC >10 x $10^3/\mu L$) was associated with 4.4 times greater odds of appendicitis in patients with a non-diagnostic USS result[82,100].

Going into more detail, children within our cohort of younger age (<8 years of age) showed significant group differences in inflammatory counts (median CRP: 42.0 [16.0; 105.0] mg/L *vs.* 16.0 [5.0; 49.0] mg/L; p<0.001 and mean leukocyte count: 16.68 ± 5.73 x $10^3/\mu L$ *vs.* 14.24 ± 5.02 x $10^3/\mu L$; p=0.002) compared to older children (≥8 years of age). Again, this finding reflects known differences in the disease process between younger and older children. These findings must however be taken in the context of their limitations as inflammatory markers are significantly higher in young children[94,99]. Overlap of symptoms from co-infections, bacterial or viral in origin, can accelerate inflammatory cell production and expedite the processes needed to produce primary and secondary radiological findings that support a diagnostically conclusive result.

Laboratory markers such as CRP-, WBC- and neutrophil counts significantly guide clinicians in their decision process and help raise the suspicion for a clinically diagnosed appendicitis. However, a raised WBC on its own is insufficient as a diagnostic modality and only alters the probability of a diagnosis to a modest degree[33]. Likewise, the CRP level on its own is quite limited for the diagnosis of acute appendicitis in general but may serve as a strong predictor for appendiceal perforation when presenting later in the course of the condition[33,35,74].

Thus, young age and increased WBC are associated with the development of a positive, diagnostic ultrasound result over an equivocal result in children presenting with symptoms of acute appendicitis. Despite abdominal pain duration not showing a clear association with the development of a positive ultrasound result on logistic regression testing, and only showing a decreasing tendency of picking up a positive ultrasound with rising abdominal pain duration, we recommend performing an abdominal ultrasound as early as possible with our results showing a tendency towards catching a positive ultrasound result in children presenting with symptoms of acute appendicitis below 60 hours. This specifically applies to school-age children who present with equivocal USS

results in whom we assume that serial abdominal ultrasound testing between 12-36 hours after pain onset will bring more diagnostic clarity. Our data suggests that the differentiation will become more difficult thereafter.

4.2 Factors associated with appendiceal perforation

In the past decades much research has been conducted on clinical factors that are associated with appendiceal perforation in children to help reduce the perforation rate and detect cases of appendicitis before progressing to perforation[3,5,10,12,22,23,101]. Our study investigated common clinical parameters (age, sex, abdominal pain duration, ultrasound result, WBC and CRP levels) to assess their association of appendiceal perforation confirmed upon surgery.

In accordance with the literature, we could show that children with later perforation had statistically significantly longer abdominal pain duration with more than half of all perforated cases (52.5%) presenting after 36 hours after onset of abdominal pain. Moreover, the ratio of perforated cases increased gradually as abdominal pain duration increased, presumably reducing tissue perfusion and causing subsequent tissue damage/necrosis that would later go along with rupturing the adjacent organ wall[23]. Our logistic regression testing revealed that rising abdominal pain duration on admission was independently associated with later appendiceal perforation (OR [95% CI]: 1.019 [1.004-1.034]; p<0.001). In 2012, Bachur *et al.* demonstrated the risk for perforated appendicitis increasing significantly with rising abdominal pain duration, with rates of 3% (0-12h), 7.6% (≥12-36h), 13% (≥36-48h) and 23.3% (≥48-72h) from shortest to the longest pain duration categories (test for trend: OR [95% CI]; 1.65 [1.50-1.82])[5]. Our study was able to reproduce similar results [4.16% (0-12h), 11.65% (≥12-36h), 18.29% (≥36-48h), 32.6% (≥48-72h) and 40% (≥72h)]. Likewise, Mallick *et al.* (2008) showed that appendiceal perforation in children 5 years of age or less occurred in 26.6% of patients with symptoms for <48 hours and in 73.4% of patients with symptoms for ≥48 hours[3]. Children of the same age (≤5 years old, n=23) reported in our study showed a similar tendency (perforation occurring in 28.57% of children with <48 hours of abdominal pain duration [4/14] and in 55.56% of children with ≥48 hours of abdominal pain duration [5/9]) and did confirm abdominal pain duration to be a primary factor associated with perforation. Overall, 9 (39.1%) children below the age of 5 showed perforation at the time of surgery.

Besides abdominal pain duration, patient age may also be a factor associated with appendiceal perforation. Despite the common assumption that appendicitis behaves differently in younger compared to older patients, the exact age cut-off is still not clear[12]. We could show that the perforation rate was highest (39.1%) among pre-school children (<6 years) and decreased with rising patient age. Overall, mean patient age was significantly younger in children with later perforation compared to those without perforation (10.14 ± 3.92 *vs.* 11.81 ± 3.28 years; p=0.003) which once again confirms the findings of previous studies demonstrating the effects of age exerting an influence on appendiceal perforation[12,23]. Our findings are within the range of those reported in other studies[5,102]. Narsule *et al.* (2011)[12] used a paediatric cohort (aged 3-18 years) with 197 confirmed cases of appendicitis. With a perforation rate of 31.5% (62/197 cases), their cohort showed a much higher perforation rate compared to ours, yet their children with subsequent perforation were considerably younger than ours (median age 8 years). Studies in the past have linked high incidence rates of complicated appendicitis with pre-school children, even reaching 75% or more in children younger than 5 years[10,23]. This might as well be due to the prolonged work-up needed to differentiate acute appendicitis from other differential diagnoses in a not yet verbally communicating child and the anatomic immaturity as young children oftentimes lack an adequate omental barrier, contributing to more rapid progression of the ongoing inflammation and later perforation[103,104]. Particularly, Alloo *et al.* (2004) highlighted the complexity of symptom presentation when they assessed a cohort of 27 children below the age of 3 years who were all found to have appendiceal perforation at the time of surgical intervention (average symptom duration: 72 hours)[104].

In light of these differences, one can assume that specific constellations of concurrent co-infections, inability to communicate and anatomical differences that are found specifically in this young age group contribute to the increased perforation rate[3]. Gurien and her colleagues (2016)[102] reported in their cohort of 484 children a perforation rate of 22.5% (109/484 cases) with an almost identical age distribution compared to ours (10.5 ± 3.6 years), however, their cohort included a rate of comorbidities of 20%. Our perforation rate of 18.1% (61/337 cases) was lower, as our cohort included more children of older age (11.51 ± 3.46 years) but also had fewer comorbidities detected on admission (4.2%; 14/337 cases) which translated into fewer counted perforations.

In view of our great number of common-school age children (≥6-12 years, n=151) and adolescents (≥12-18 years, n=163), we could illustrate a 'classic development' of appendiceal perforation. Compared to non-perforated cases, both common-school age children and adolescents with later perforation presented most often on day 2 (36≤ x < 48h) after onset of abdominal pain, while pre-school children (n=23) presented most frequently two days after pain onset. Given the statistical significance of inter-group differences (p<0.001), we were able to characterise the progression to perforation in each of these age groups. This will help to estimate the perforation risk of future children with similar demographic characteristics who present to ED with acute appendicitis.

Nevertheless, despite significant group differences between perforated and non-perforated cases in our cohort, logistic regression modelling could not confirm age to be independently associated with appendiceal perforation (OR [95% CI]: 0.9232 [0.840-1.013]; p=0.0915). Yet, with a decreasing odds ratio of 0.9232, we could at least identify a trend of appendiceal perforation moving towards younger age.

In accordance with the above-mentioned study, our cohort did not show any positive association with either male or female sex for appendiceal perforation, (OR [95% CI]: 1.346 [0.7027-2.591]; p=0.369)[102]. Still, we could show clear group differences between cases of perforation *vs.* non-perforation regarding inflammatory markers (WBC, CRP). Among patients with later perforation, mean WBC count on admission was higher compared to those without later perforation (16.66 ± 5.52 x 10^3/µL *vs.* 14.12 ± 5.04 x 10^3/µL; p=0.001) which is in line with other studies reflecting on the increased inflammatory count of patients with appendiceal perforation *vs.* non-perforation[12,23,102]. Despite Gurien et al. [102] showing a clear positive association of WBC count in predicting appendiceal perforation (OR [95% CI]: 1.08 [1.04-1.12]; p<0.001), we could not confirm these findings using logistic regression. In our cohort, WBC count was only marginally associated with appendiceal perforation (OR [95% CI]: 1.052 [0.989-1.120]; p=0.107). In comparison, Gurien's cohort (n=484) included patients with concomitant comorbidities (20%) and showed higher leukocyte counts on admission in the first place, suggesting that their children already showed a higher degree of morbidity upon admission to ED (Gurien et al. [102]: 15.4 ± 6.8 x 10^3/µL *vs.* our study: 14.58 ± 5.21 x 10^3/µL). Our cohort only included 4.2% of comorbidities so that the overall contribution from concurrent infections etc. on the elevated leukocyte count can be presumed to be

less pronounced and therefore may have decreased the predictor effect in our logistic regression modelling. As outlined above, appendiceal perforation occurs at the later phases of appendicitis when, presumably, neutrophil consumption exceeds its production which translates into decreased peripheral leukocyte counts[95,96]. In children presenting later in the course of the condition, serum WBC count may therefore be limited in predicting appendiceal perforation in combination with an equivocal ultrasound result on admission.

However, the CRP level has proven to be more reliable in patients with more advanced stages of appendicitis and/or later presentation[85]. Among children with later perforated appendicitis, median [Q_1; Q_3] CRP level on admission in our cohort was 74.0 [27.0; 163.5] mg/L compared to those without later perforation, 15.0 [4.0; 42.5] mg/L. Furthermore, in those with later confirmed perforation, their CRP level on admission was not only found to be elevated but higher than in those who presented at the same time without later appendiceal perforation. Virtually, the highest median CRP level measured in cases without perforation was detected at 57 [7.5; 122.5] mg/L (60-72 hours after onset of abdominal pain). For comparison, patients with later appendiceal perforation already achieved this level within the first 48 hours after onset of abdominal pain. As appendiceal perforation is "on its way", median CRP levels will steadily rise with continuous abdominal pain duration, making it an ideal parameter to triage children in ED. As anticipated, our logistic regression analysis showed a positive association of CRP and appendiceal perforation on admission (OR [95% CI]: 1.012 [1.007-1.018]; $p<0.001$). The clear emphasis of CRP over WBC associated with appendiceal perforation once again illustrates the underlying temporal relationship of inflammatory markers in the development of appendiceal perforation[85,99,100]. CRP contributes little diagnostic utility in the early phases of simple appendicitis but gains momentum when differentiating cases of perforated appendicitis from non-appendicitis, not alone due to the fact that perforated cases often present later in the course of the condition anyway[33].

Of major interest to us was whether the initial ultrasound on admission would allow any conclusions about a possible association with later appendiceal perforation. Among all perforated cases, 75% (46/61 cases) showed an initial positive, diagnostic ultrasound on admission, yet comparison between perforated and non-perforated cases did not result in a higher frequency of positive USS detection (pos. USS: [Perforated] 75.4% *vs.* [Non-perforated] 66.3%; p=0.338). As expected, the odds ratio of 0.5901 was only able to

depict an association between the ultrasound on admission and later appendiceal perforation (OR [95% CI]: 0.5901 [0.276-1.198]; p=0.157). However, this finding did not achieve statistical significance. Conclusive studies about the predictive value of a positive, diagnostic ultrasound on later appendiceal perforation are scarce. Current literature has shown that an appendicolith found on CT or USS scan was associated with higher risk of perforation[103,105,106]. But, to what extent a positive, diagnostic ultrasound is able to predict later perforation remains inconclusive. Given the lack of clear differences and statistical significance on our logistic regression model, we assume that the initial ultrasound result is not associated with the outcome of later perforation. Repeat clinical examinations and serial laboratory tests will then strengthen the diagnostic certainty in cases presenting with an initial equivocal ultrasound result.

Concluding, our findings demonstrate that in the context of recent symptom duration, an equivocal ultrasound with normal or near-normal leukocyte and CRP level should reassure the clinician that appendiceal perforation is less likely at the time. However, especially in pre-school children, high inflammatory counts on admission, CRP more than WBC, should at least increase suspicion for appendiceal perforation and encourage clinicians to actively seek diagnostic confirmation. The CRP level on admission has turned out to be a useful clinical parameter to hold on to in children presenting with suspected appendicitis whose abdominal wall could be on the brink of appendiceal perforation. While a non-diagnostic USS scan may not allow robust conclusions to be drawn regarding the perforating appendix, and while USS performance is dependent on the trained examiner, USS performed early in the course of appendicitis, rather than later, yields more diagnostic benefit in terms of its therapeutic consequence as the scan results are more likely to be less ambiguous and therefore more diagnostically conclusive.
Our study is one more piece of evidence that can be added to the growing body of literature describing the temporal relationship of appendiceal perforation in children which could help to realise the before mentioned major outcome of ED care to properly identify appendicitis before progressing to perforation. For future purposes, the results of the present study could be used to educate parents and caretakers as well as primary care providers about the risk of diagnostic delay in paediatric appendicitis. This way, we can further decrease both the negative appendectomy and the appendiceal perforation rate in children.

4.3 Factors associated with a short latency to operation

For decades, the optimal timing to perform appendectomy has been a central research topic in general surgery with articles found in the literature dating back to the 1980s[102,107]. Yet, despite published articles now stretching over almost 40 years, there is still no consensus with respect to this issue. A recent survey of nearly 500 U.S. paediatric surgeons from 2012 showed that a vast majority (96%) did not consider non-perforated appendicitis to be a surgical emergency while 92% believed that delaying overnight appendectomy would not go along with a clinically significant increased risk for perforation. More strikingly, more than two thirds of respondents reported no existing departmental guideline on when to delay overnight appendectomy[108]. Over the years, emphasis has been placed on the potential detrimental impact of prolonged working hours for residents[109,110]. Several studies have illustrated the negative effects of sleep deprivation in residents on clinical and surgical performance rates, calling the need for surgical procedures at night into question while suggesting that appendectomies performed at night be limited to absolute emergencies only[111-113]. A study in 2008 suggested that a reduction of operations performed out-of-hours may have a positive impact on both medical professionals and patients[114].

For this reason, several studies conducted over the past 16 years have evaluated differences in outcomes between immediate and delayed surgical interventions and proposed delayed appendectomy for acute appendicitis, particularly for patients admitted to the emergency department at night-time[101,103,114-120]. The evidence behind these studies illustrates a heterogenous landscape for the medical community with studies documenting a higher incidence of complications with prolonged latency to operation theatre[103,115], whereas others found no differences in adverse events when operating the child the next day or thereafter[101,114,116-120]. Many studies used mixed sample groups of children and adults alike making it difficult to compare outcomes among each other[22,26,102,111,114,118,121-126]. It comes as no surprise that previous studies have therefore reported conflicting results in finding an association between early appendectomy and risk of complicated disease due to prolonged latency to OT. The studies used were oftentimes single-institutional, were limited by inadequate sample sizes, differences in patient populations, transfer patterns to OT and peri-operative management[12,103,118,122,125-127]. Beyond that, other studies using administrative and registry data have tried to improve their study power by using subjective and imprecise ICD-9-CM based diagnostic codes for defining their study

outcomes and their inability to adjust for hospital-related effects on observed perforation rates[26,84,121,124].

One frequently encountered problem is the fact that time to OT only reflects a marginal proportion of overall duration in the development of appendiceal perforation. It is often the case that abdominal pain duration until presenting to ED is left out the analysis and will therefore not be considered as a contribution to appendiceal perforation in studies assessing intra-hospital delay and perforation risk. A meta-analysis of 11 studies (8858 patients), published in 2014 and led by the UK National Surgical Research Collaborative, revealed that a delay of 12 to 24 hours after presentation to the ED was not associated with an increase in the risk of complex appendicitis (OR 0.97, p=0.750)[128]. Yet, only 21.3% of included patients were below the age of 16 years, making this meta-analysis hardly representative for a paediatric cohort in general. Moreover, the meta-analysis claimed that even a delay beyond 48 hours did not show a significant increase in perforation rate, while illustrating a significant increase in wound infection rates (adjusted OR 2.24; p=0.039). For comparison, among all perforated cases of our paediatric patients only cohort, 56% of perforations were detected at 48 hours or below after onset of abdominal pain. To give another example, the perforation rate between 28-32h (18.5%) and 32-36h (36.4%) increased by 17.9% in just 4 hours. Delaying appendectomy until the next morning (e.g. *10:00 pm to 06:00 am*) already accounts for 8 hours which would go along with a needless increase in perforations the next morning if surgical procedures were not to be performed at night-time at all. A latency to OT of 24 hours or more represented a major fraction of overall abdominal pain duration in our cohort of children who already showed perforation at day 2 (48 hours). This observation had neither been reflected on in the meta-analysis nor in many other previously conducted studies[128].

Of those studies using paediatric only cohorts, many were not able to bring further clarity into this medical obscurity. Brender *et al.* (1985)[107], Surana *et al.* (1993)[129], Yardeni *et al.* (2004)[119], Taylor *et al.* (2005)[127] and Bush *et al.* (2011)[121] all found relatively greater rates of longer delay in those children showing perforation, yet none detected significantly increased perforation rates impacted by intra-hospital delay. None used pre-operative radiographic imaging studies to differentiate between groups compared. It is therefore unknown whether differences in group characteristics may have confounded the analysis[103,130]. One of the studies published in more recent times (2017) aimed at putting

an end to this scientific inconsistency. Serres *et al.* conducted a retrospective multicentre cohort study at 23 U.S children's hospitals between 2013 and 2014, based on the U.S Paediatric National Surgical Quality Improvement Program (2429 children aged <18 years who underwent appendectomy within 24 hours of presentation) and found that a delay of appendectomy within 24 hours after presenting to the emergency department was not associated with an increased risk of complicated appendicitis or adverse outcomes[84]. However, this study failed to exclude patients who may have already had a perforated appendicitis prior to arrival in ED. By including a cohort of patients which may already have reached the primary outcome of appendiceal perforation, detecting a statistically significant difference in perforation rates between early and late appendectomy is a critical endeavour and calls for further clarification[130].

This problem was overcome by the most recent study on this issue in 2019 when Meltzer et al. [130] performed a multicentre retrospective cohort study of children (<18 years of age) who had CT-scan-proven non-perforated appendicitis at the time of presentation to ED. After adjusting for potential confounders, every hourly increase in delay from ED to surgery was independently associated with a 2% increase in the odds of perforation (adjusted OR [95% CI]: 1.02 [1.0-1.04]; p=0.03). To put this into general terms, a 12-hour delay until appendectomy increased the odds of perforation by approximately 25%. This finding was particularly affected by both patient age and WBC count, two commonly cited risk factors for appendiceal perforation[10,103,130]. The critical fact of excluding children with prior perforated appendicitis before arrival to ED allows for clear conclusions to be drawn from the perforation rates. The results of this study were confirmed in a previous study by Bonadio et al. [103] in 2015 who also included children of CT-confirmed uncomplicated appendicitis only and was able to show that 96.4% of perforated cases occurred in patients with a delay of >9 hours from ED presentation to appendectomy.

The rate of perforation shown in this study was approximately six-fold greater in those with >9 hours of delay until appendectomy (25%) compared to those of ≤9 hours of delay (4.6%) (p<0.005). Unlike many other studies, Bonadio *et al.* reported that 94% of included patients exhibited an abdominal pain duration of 48 hours or less before presenting to ED, making every additional hour of delay spent waiting for appendectomy so critical. Although often cited as preventative, the study found that antibiotic therapy

before appendectomy was not able to halt progression to perforation as close to 25% of uncomplicated appendicitis cases progressed to appendiceal perforation despite prior antibiotic usage[119,127,131]. Taken together, the increased morbidity along with perforation on grounds of surgical delay to appendectomy serves as the primary justification for prompt surgical intervention in all cases of acute appendicitis regardless of time of admission[26]. Using our cohort, we set ourselves the goal of investigating clinical parameters that were associated with an early (<6 hours) or a late appendectomy (≥6 hours). In addition, we wanted to assess whether there were any age-related differences in terms of timing of appendectomy.

Descriptive analyses showed significant group differences between appendectomies performed before and after 6 hours after ED admission. For example, in our cohort children of short latency were statistically younger (10.82 ± 3.51 *vs.* 11.79 ± 3.38 years; p=0.016). Subgroup analyses showed that especially among pre-school children (<6 years of age), more children received an appendectomy within 6 hours after arrival (56.5% *vs.* 43.5%) while the vast majority of older children were all operated after 6 hours after arrival in ED (≥6-12 years: 32.5% *vs.* 67.5%; ≥12-18 years: 27.3% *vs.* 72.7%; p=0.018). Furthermore, children of short latency tested more often positive for appendicitis on ultrasound exam (79.2% *vs.* 63.3%; p=0.002) with older children exhibiting the largest group differences (≥12-18 years: 77.3% *vs.* 55.6%, p=0.009; ≥6-12 years: 85.7% *vs.* 71.57%; p=0.041). Children of pre-school age illustrated a reverse relationship with showing fewer positively detected ultrasound results in the group of early appendectomies. However, these results were based on a small sample size (n=23) and were not statistically significant (<6 years: 61.5% *vs.* 70%, p=0.510). Further, all children of early appendectomy taken together had significantly higher inflammatory markers on admission (WBC: 16.12 ± 5.02 x 10^3/µL *vs.* 13.92 ± 5.15 x 10^3/µL; p<0.001 and CRP: 30.0 [8.5; 105.5] mg/L *vs.* 15 [5.0; 50.6] mg/L; p=0.001). However, no significant group differences were found when assessing sex, abdominal pain duration or fever.

To assess any age-related differences in timing of appendectomy we stratified the data according to patient age. Of particular interest were inflammatory markers in pre-school children with short latency to operation theatre. Here, children showed strikingly higher CRP levels (107.0 [22.5; 198.5] mg/L *vs.* 64.0 [36.25; 97.5] mg/L; p=0.664) but no differences in terms of elevated leukocyte count (16.67 ± 6.97 x 10^3/µL *vs.* 16.8 ± 3.78 x

10^3/µL; p=0.958). No significant group differences were detected with regards to initial ultrasound result, sex, fever or abdominal pain duration.

This was different in older children (≥6-12 years of age) with short latency to OT who showed significantly higher inflammatory counts when compared among each other ([WBC: 17.28 ± 4.44 x 10^3/µL *vs.* 14.33 ± 5.47 x 10^3/µL; p=0.001] and [CRP: 26.0 (11.75; 94.5) mg/L *vs.* 17.0 (6.0; 49.75) mg/L; p=0.059]). Furthermore, they exhibited a significantly higher detection rate of positive ultrasound results on admission (85.7% *vs.* 71.57%; p=0.041) while other clinical parameters (sex, temperature and abdominal pain duration) showed no significant group differences when compared among each other. Similar results were obtained in children of adolescent age (≥12-18 years of age) where higher CRP levels (26.5 [5.25; 79.0] mg/L *vs.* 11.0 [3.0; 44.0] mg/L; p=0.050), leukocyte counts (14.7 ± 4.7 x 10^3/µL *vs.* 13.32 ± 4.85 x 10^3/µL; p=0.107) as well as their positive ultrasound (77.3% *vs.* 55.6%; p=0.009) lead clinicians to a more rapid surgery.

As anticipated, logistic regression modelling revealed that among the previously assessed clinical parameters, only leukocyte count and a positive, diagnostic ultrasound result on admission were significantly associated with a short latency to OT (<6 hours). Here, increasing leukocytes were associated with a more rapid referral to surgery (<6 hours) (OR [95% CI]: 1.068 [1.016-1.122]; p=0.010) while CRP level on admission was not showing any tendency of sending a child to early surgery (OR [95% CI]: 1.003 [0.999-1.088]; p=0.140). Again, CRP level taken on admission was associated with later appendiceal perforation but not sufficient enough to allow any statement about the urgency of appendectomy and presumably is not a robust clinical marker in a physician's decision-making process to call for prompt surgery if taken on admission. One can only estimate that the significance of CRP as a parameter associated with prompt surgery will change and allow a better prognosis further down en route to surgery. This would however require multiple blood takings which could further delay surgery. In our logistic regression model, patient age was not significantly associated with an earlier appendectomy despite showing a tendency towards prompt surgery with decreasing patient age. Meltzer et al. [130] did not exclude children with pre-existing illnesses or comorbidities and, with the necessity to perform CT-scans on admission, potentially included a cohort of children that was more "ill" compared to other studies of the same subject. In contrast, we used a cohort with only 4.2% of patients showing concurrent

infections or comorbidities so that the effect of young age on perforation can be assumed to be stronger. Yet, in spite of 'almost' significant results and as young age has been associated with earlier and higher rates of perforation, we argue for time to operation theatre in younger children be kept to a minimum.

Among the initial ultrasound results on admission, equivocal results were significantly associated with delaying appendectomy (OR [95% CI]: 0.472 [0.268-0.833]; p=0.010). The delay shown here can very well be explained with increased efforts in diagnostics and time needed to differentiate acute appendicitis from other differential diagnoses in a child presenting with acute abdominal pain. Our study has illustrated to some extent the 'classic' formation of appendicitis in school-age children, with the greatest overall percentage increase of positive ultrasounds within this age-group (+59.1%) between 12-36 hours after onset of abdominal pain. For this reason, we would argue for serial abdominal ultrasounds within this timeframe so that chances of an early appendectomy can be maximised in children presenting with an initially equivocal ultrasound upon admission to ED. However, our data was not able to reproduce this assumption objectively. Likewise, increasing age and male sex were both associated with delay to appendectomy, however, these findings were neither statistically significant.

Strikingly, abdominal pain duration did not appear to be linked to a faster admission to surgery, neither within the group as a whole, nor among all three age categories and only constituted minimal weight in terms of a clinician's decision-making process (OR [95% CI]: 1.007 [0.995-1.019]; p=0.250). Current scientific evidence shows that younger children in comparison to older children are far more affected by pre-hospital delay due to their inability to communicate to parents, atypical presentation and anatomic immaturity[3]. However, abdominal pain duration in a not-yet verbally communicating child may be very variable and accounts of parents and caretakers may not be reliable enough for the assessing clinician. This finding may as well be attributed to our retrospective study design and more conclusive results could be expected with a prospective sampling of patients where repetitive ultrasound sonographies are carried out to detect an appendix at the threshold of perforation as well as abdominal pain duration be recorded on a continuous scale rather than a categorical one. However, such a prospective longitudinal study is ethically controversial as a timely diagnosis is critical to

avoid complications. Still, we have reached a point of necessity, where robust criteria are needed as to when and if at all to delay appendectomy.

To sum up, strong parameters associated with early appendectomy were an increased leukocyte count and a positive (diagnostic) ultrasound on admission. Especially in younger children in whom ultrasound examination might not be conclusive enough, overall elevated white cell counts appeared to help expedite the process to the operation theatre and guide a clinician in his decision-making process. Moreover, higher leukocyte counts were helpful in both school-age children and adolescents. CRP levels were not clearly helpful in the decision-making process for early surgery. On top, our data suggests that abdominal pain duration is not necessarily associated with a quick route to operation.

4.4 Conclusion

Using our data, we could show that by performing abdominal ultrasound sonography upon admission the presence of appendicitis in true cases of appendicitis was detected in close to 70% which represents a result that is largely free from comorbidities which could potentially have confounded the detection rate. Moreover, this result displays a good threshold and justifies the routine use of sonography in the paediatric emergency department. In the context of contemporary clinical practice, our data has shown that the odds of detecting diagnostically conclusive findings on ultrasound sonography for cases of acute appendicitis upon admission decrease with rising patient age. Alongside the obligatory clinical and more specific abdominal examination, positive, diagnostic ultrasounds occurred statistically significantly more often in children of younger age. Moreover, while CRP levels on admission were not, higher leukocyte counts were associated with a positive ultrasound on admission reflecting the status of acute appendicitis at its early stage. Despite lacking statistical significance on logistic regression modelling, we could show a tendency of decreasing odds of a positive ultrasound finding with rising abdominal pain duration. Given our data, we recommend performing abdominal ultrasound sonography as early as possible with a decreasing trend of positive, diagnostic USS results after 60 hours after onset of abdominal pain. Especially younger children (<6 years of age) deserve a low threshold of suspicion on the part of parents and medical professionals alike when presenting with acute abdominal pain in order to allow for diagnostic cases to be recognised in time. Further, in children of school-age, serial abdominal ultrasounds together with repetitive clinical examinations

performed between 12 and 36 hours might increase chances of detecting a diagnostic ultrasound result. The results of this study could be used to educate parents and caretakers as well as primary care providers about the risk of diagnostic delay in acute paediatric appendicitis. This way, we can further decrease both morbidity rate as well as appendiceal perforation rates in children.

Regarding possible links between clinical parameters and later outcomes of appendicitis, our data has shown that rising abdominal pain duration is significantly associated with an increasing appendiceal perforation rate with more than half of our cases of perforation detected after 36 hours of abdominal pain. Perforation rate was found to be the highest among children of pre-school age and decreased subsequently as children were older. However, decreasing patient age was almost statistically significantly associated with an increased appendiceal perforation rate. Among CRP and leukocytes, CRP level on admission has proven to be an ideal parameter to triage children in ED with showing clear associations of later appendiceal perforation. Practically speaking, cases of elevated CRP levels on admission are highly suspicious of advanced stages of acute appendicitis and should be given further attention. However, the outcome of a positive ultrasound exam on admission was not associated with an increased risk of perforation and was not necessarily helpful to forecast perforation. Repeat clinical examinations and serial laboratory tests will then strengthen the diagnostic certainty in cases presenting with an initially equivocal ultrasound result.

Investigating the relationship between clinical parameters and time to surgery, our study found that positive ultrasound result and elevated white cell count on admission were significantly associated with a prompt pathway to appendectomy (<6 hours from admission to appendectomy). Our data suggests that especially in younger children in whom clinical and ultrasound examinations might not bring conclusive results, elevated leukocyte counts on admission can help to expedite the process to appendectomy and presumably reduce the morbidity that has been linked to increased intra-hospital delay. Differences in abdominal pain duration were not statistically significant in either group despite a positive relationship between pain duration and perforation rate and therefore constituted only marginal weight in the decision-making process leading up to prompt appendectomy. Given the most recent evidence on intra-hospital delay in paediatric appendicitis, we argue for early appendectomy in all cases of acute appendicitis

regardless of time of admission. Future studies are needed to detect the exact timing of an appendix at the brink of perforation as well as the optimal timing of appendectomy in cases of acute appendicitis. Ideally, these studies should contain multi-institutional data with standardised definitions, a prospective sampling of paediatric patients and the use of abdominal pain duration collected on a continuous rather than a categorical scale. This way, all doubts about modality and interpretation can be reduced to a minimum allowing for universal comparison among different patient groups. Additionally, these studies could potentially provide statistically significant results for the predictor effect of abdominal pain duration on the likelihood of a diagnostically conclusive USS on admission.

Within a tertiary care setting, abdominal ultrasound sonography may have proven to be a crucial contributor to correctly diagnosing paediatric appendicitis. However, its relative high rate of equivocal results, especially at advanced stages of appendicitis, in turn confirms the diagnostic relevance and emphasis of additional laboratory and clinical examinations.

Limitations:

There are several limitations to our study. (1) First of all, this study was performed at a single tertiary paediatric hospital and therefore may not be reproducible in other settings. Our hospital is the main referral centre for paediatric emergencies, both surgical and non-surgical, in the area. However, it is not excluded that very few children were discharged home due to lack of consistent clinical symptoms and may have subsequently been diagnosed with appendicitis at another hospital (although children are frequently referred to our institution for surgical management with nearly no existing alternative hospital in the immediate vicinity). (2) Furthermore, this study is limited by its retrospective nature. As with all retrospective studies, results are dependent on accurate coding. Historical data supplied by caregivers could not be verified for accuracy and can be expected to vary among children and caregivers as their perception of a child's physical well-being is limited.

(3) Similarly, the analyses are limited by the use of large time intervals for abdominal pain duration rather than time used as a continuous variable. As abdominal pain onset is not completely predictable nor can be precisely recalled, many clinical notes used broad terms to describe the onset of abdominal pain. For example, 'yesterday', 'last night' and 'this morning' were seen frequently within the medical records. This was later compared to the time of admission to the emergency department to determine the length of abdominal pain duration. This was done in accordance with previously conducted studies as this method is practical for most patients who cannot recollect the exact onset of their abdominal pain. Clinicians were not specifically trained to document the exact onset of pain and we were therefore reliant on the clinical notes and intrinsic variability in clinician documentation. This was however adequate to create groups of patients with intervals of 12 hours between each other. We used time of admission to hospital as a surrogate marker for when surgical evaluation and appendiceal ultrasound were performed. Although this may not always have been accurate, we felt it was a better approximation than other potential starting points because of confounding factors that may have biased the data. This may have lessened the power of this study.

(4) Even with a cohort of over 330 proven cases of appendicitis, the sample size of children less than 6 years of age was relatively small and therefore estimates of diagnostic performance of ultrasound as well as inflammatory markers will undoubtedly have less

precision than the other two age groups. Other factors that may influence patient outcomes were not specifically addressed in this analysis but are relevant to the medical decision-making process in cases of suspected appendicitis. (5) Our cohort included only patients who had histologically proven appendicitis. Comparing our outcomes to those of previously conducted studies on the predictive value of ultrasound upon admission to differentiate appendicitis from non-appendicitis is therefore not possible. For example, our negative appendectomy rate of 1.7% (8 out of 462 screened cases) is the artificial result of our retrospective study design where the search strategy excluded patients with comorbidities and aimed at patients with histologically proven appendicitis only.

(6) Another limitation to this study is the differentiation between diagnostic and non-diagnostic USS results. Given our retrospective approach to patient sampling, we excluded USS cases without histological proof of appendicitis. Therefore, false-positive, and true-negative cases were not accounted for in our population and analyses. We also set the focus of the present work on scenarios where the USS result would provide therapeutic consequences for the treating clinician. As such, we only compared true-positive (diagnostic) USS results with equivocal USS results, albeit equivocal USS results are not synonymous with non-diagnostic USS. However, future prospective sampling would be needed to make up for the remaining non-diagnostic (false-negative & false-positive cases) and diagnostic (truly negative cases) subgroups to gain clarity in its entirety.

(7) Ultrasounds which represented an urgent surgical indication (independent of suspected perforation) also constituted a ‘positive’ ultrasound result which is different from point-of-care diagnostics in a general hospital. Thus, the percentage of appendices seen on ultrasound was high (ca. 68%) in our cohort which may well represent the ‘best-case scenario’. As many children are seen at non-paediatric facilities, institutions without dedicated paediatric sonography may not have the same expertise when evaluating these sonographic predictors. Our results may therefore differ from other studies conducted at more rural or less specialised hospitals. Also, patient-related factors such as body mass index which can affect the test characteristics of abdominal ultrasound sonography, were not assessed.

A growing body of evidence has demonstrated some success in the non-operative management of patients with image-proven uncomplicated appendicitis to the benefit of selected patients who may safely avoid the risks of surgery[26,132]. These patients are mainly treated using antibiotics. Yet, current literature has shown that the non-operative management of acute uncomplicated appendicitis does not statistically increase appendiceal perforation while complication-free treatment success rates have been shown to be higher with surgical management[133]. Our study was conducted at a time when non-operative management of acute paediatric appendicitis was not a standard even in selected patients. Data about antibiotic usage is therefore lacking in this study. However, it is not unreasonable to believe that the diagnostic accuracy of ultrasound will be of crucial importance for future studies evaluating the non-operative treatment of uncomplicated appendicitis.

Lastly, this study was conducted in a single centre which sees over 25.000 children each year. Our sonographers have gone through extensive ultrasound training beforehand and therefore have significant exclusive paediatric experience in the evaluation of acute abdominal pain. We know that the performance and reading of paediatric abdominal ultrasounds are dependent on the examiner, yet, despite supervision by consultants and control sonographies, many ultrasound results remained non-diagnostic. The fact that appendicitis detection with ultrasound sonography is technically limited translates into an increasing significance of the clinical examination. Since ultrasound cannot replace the clinical examination, the results of this study will therefore only add further to the diagnostic certainty. Even if serial ultrasounds remain equivocal, and therefore non-diagnostic, the clinical symptoms of appendicitis should be guiding the treating physician in his decision-making and prevent a ‘false‘ reliance on technical diagnostics.

Our study has illustrated the ‘classic formation‘ of appendicitis in a cohort of common-school age children for which serial abdominal ultrasound exams may bring more diagnostic certainty if the initial ultrasound is ‘equivocal‘. However, our data was not able to reproduce this assumption objectively. The results obtained here will provide assurance in the evaluation of disease progression and provide a meaningful diagnosis in paediatric appendicitis. This will be of particular use in cases where medical professionals are facing medicolegal implications for the accusation of not having acted in a timely manner when children suffered additional harm due to appendiceal perforation.

Ultrasound sonography in paediatric appendicitis has its limitations and the clinical examination is of upmost importance in managing a child. Together with the ultrasound and other clinical parameters, the diagnostic certainty can only be increased while harmful consequences of negative appendectomies or false-negative cases are limited to the best of current abilities.

Summary

Background: Ultrasound sonography is the leading diagnostic modality in suspected cases of paediatric appendicitis. However, its diagnostic validity is limited due to the variability in ultrasound performance throughout the disease process.

Objective: In order to validate the routine use of abdominal sonography, the influence of patient-related data and the time-dependent nature of ultrasound results throughout the disease process will be investigated and analysed.

Material and Methods: Multiple logistic regression models of a single-centre retrospective cohort of children aged 2 to 18 years with histologically proven appendicitis who underwent ultrasonography were fitted.

Results: 337 children were analysed (11.5 ± 3.46 years); Ultrasound sensitivity on admission was 68%. Perforation occurred in 61 cases (18.1%). The odds of detecting a positive ultrasound increased with decreasing patient age (OR [95% CI]: 0.929 [0.863-0.999]; p=0.049). Despite improving ultrasound sensitivity with rising leukocyte counts (OR [95% CI]: 1.056 [1.006-1.110]; p=0.031), increasing pain duration was not significantly associated with a positive ultrasound result on admission. Odds of perforation rose with increasing pain duration (OR [95% CI]: 1.019 [1.004-1.034]; p<0.001) and increasing CRP levels on admission (OR [95% CI]: 1.012 [1.007-1.018]; p<0.001) but showed a decreasing tendency with older patient age. Yet, a conclusive ultrasound result was not significantly associated with higher rates of perforation.

Conclusion: Use of abdominal sonography in suspected cases of appendicitis has proven to be an important modality to reach a diagnosis in a timely manner. Increasing leukocyte counts and young age were significantly associated with a later confirmatory ultrasound result. Moreover, a significant relationship between young age, rising CRP and increasing pain duration was found for later stages of appendicitis with perforation. Despite no association of initial ultrasound result and appendiceal perforation, the confirmed time-dependent nature of appendiceal perforation and reduced operation latency in cases of initial positive ultrasound result highlights the ultrasound's potential to help determine the course of this disease.

Hintergrund: Die Ultraschall-Sonografie gilt in der Appendizitis Diagnostik bei Heranwachsenden als führende Modalität. Aufgrund ihrer Befundvariabilität im Krankheitsverlauf ist ihre diagnostische Zuverlässigkeit jedoch begrenzt.

Ziel der Arbeit: Zur Validierung des Routineeinsatzes dieser Methode wird der Einfluss patientenbezogener Daten und die Zeitabhängigkeit der Ultraschall-Ergebnisse im Krankheitsverlauf untersucht.

Methoden: Eine mono-zentrische Kohorte von Kindern (Alter 2-18 Jahre) mit histologisch gesicherter Appendizitis und Ultraschall-diagnostik bei Aufnahme wurde mittels multinominal logistischer Regressionsmodellen retrospektiv analysiert.

Ergebnisse: Die Abdomen-Sonografie an 337 Kindern (11.5 ± 3.46 Jahre) ergab bei Aufnahme eine Sensitivität von 68%. Eine Appendix-Perforation ergab sich für 61 Fälle (18.1%). Die Odds einen positiven Ultraschall bei Aufnahme zu detektieren, nahm mit steigendem Alter ab (OR [95% CI]: 0.929 [0.863-0.999]; p=0.049). Obwohl die Aussagekraft des Ultraschalls mit im Verlauf steigender Leukozyten-Zahl (OR [95% CI]: 1.056 [1.006-1.110]; p=0.031) zunahm, so korrelierte die steigende Krankheitsdauer nicht signifikant mit einer erhöhten Ultraschall-Detektionsrate. Die Odds einer Appendix-Perforation stieg mit steigender Schmerzdauer (OR [95% CI]: 1.019 [1.004-1.034]; p<0.001) und steigendem CRP bei Aufnahme (OR [95% CI]: 1.012 [1.007-1.018]; p<0.001), und zeigte eine abnehmende Tendenz mit steigendem Kindesalter. Hingegen war ein positiver Ultraschallbefund bei Aufnahme nicht signifikant mit einer späteren Perforation assoziiert.

Schlussfolgerung: Die Ultraschalluntersuchung bei Appendizitis Verdacht hat sich als wichtiger Beitrag zur Diagnosesicherung und zeitnahen Therapieentscheidung bestätigen lassen. Dabei erweisen sich junges Alter und erhöhte Leukozyten-Zahlen signifikant mit einem später bestätigten Ultraschallbefund assoziiert. Ein signifikanter Zusammenhang von jungem Alter, steigendem CRP, sowie andauernden Schmerzdauer findet sich ebenfalls für spätere Stadien der Appendizitis mit Perforation. Auch wenn sich kein Zusammenhang zwischen Ultraschallergebnis und Perforationsrate ergibt, lässt die gesicherte zeitabhängige Perforationsrate und geringere Operationslatenz bei positivem Ultraschallbefund auf dessen verlaufsbestimmendes Potenzial schließen.

Glossary

ANC	-	Absolute neutrophil count
BMI	-	Body mass index
C	-	Celsius
CRP	-	C-reactive protein
CT	-	Computer tomography
ED	-	Emergency department
IQR	-	Inter quartile range
MRI	-	Magnetic resonance imaging
NA	-	Negative Appendectomy
NPV	-	Negative predictive value
OR	-	Odds ratio
OT	-	Operating theatre
PMN	-	Polymorphonuclear leukocyte
PPV	-	Positive predictive value
Q_1; Q_3	-	Quartile 1; Quartile 2
RIM	-	Radiation induced malignancy
RLQ	-	Right lower quadrant
SD	-	Standard Deviation
USS	-	Ultrasound sonography
Vs	-	versus
WBC	-	White blood cell count

Key Words:
Paediatric appendicitis, abdominal pain duration, sonography, accuracy, validity, appendectomy

Table of Figures

Bibliography

118. Abou-Nukta F, Bakhos C, Arroyo K, et al. Effects of delaying appendectomy for acute appendicitis for 12 to 24 hours. *Arch Surg* 2006;141(5):504-6; discussioin 06-7.

14. Acheson J, Banerjee J. Management of suspected appendicitis in children. *Arch Dis Child Educ Pract Ed* 2010;95(1):9-13.

1. Addiss DG, Shaffer N, Fowler BS, et al. The epidemiology of appendicitis and appendectomy in the United States. *Am J Epidemiol* 1990;132(5):910-25.

104. Alloo J, Gerstle T, Shilyansky J, et al. Appendicitis in children less than 3 years of age: a 28-year review. *Pediatr Surg Int* 2004;19(12):777-9.

120. Almström M, Svensson JF, Patkova B, et al. In-hospital Surgical Delay Does Not Increase the Risk for Perforated Appendicitis in Children: A Single-center Retrospective Cohort Study. *Ann Surg* 2017;265(3):616-21.

40. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 1986;15(5):557-64.

32. Anandalwar SP, Callahan MJ, Bachur RG, et al. Use of White Blood Cell Count and Polymorphonuclear Leukocyte Differential to Improve the Predictive Value of Ultrasound for Suspected Appendicitis in Children. *J Am Coll Surg* 2015;220(6):1010-7.

29. Andersson RE, Hugander AP, Ghazi SH, et al. Diagnostic value of disease history, clinical presentation, and inflammatory parameters of appendicitis. *World J Surg* 1999;23(2):133-40.

37. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. *Br J Surg* 2004;91(1):28-37.

4. Ang A, Chong NK, Daneman A. Pediatric appendicitis in "real-time": the value of sonography in diagnosis and treatment. *Pediatr Emerg Care* 2001;17(5):334-40.

73. Ärztekammer B. Neufassung der „Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen – Rili-BÄK“. Deutsches Ärzteblatt, 2014: A 1583.

5. Bachur RG, Dayan PS, Bajaj L, et al. The effect of abdominal pain duration on the accuracy of diagnostic imaging for pediatric appendicitis. *Ann Emerg Med* 2012;60(5):582-90.e3.

100. Bachur RG, Dayan PS, Dudley NC, et al. The Influence of Age on the Diagnostic Performance of White Blood Cell Count and Absolute Neutrophil Count in Suspected Pediatric Appendicitis. *Acad Emerg Med* 2016;23(11):1235-42.

58. Bachur RG, Hennelly K, Callahan MJ, et al. Advanced radiologic imaging for pediatric appendicitis, 2005-2009: trends and outcomes. *J Pediatr* 2012;160(6):1034-8.

51. Baldisserotto M, Marchiori E. Accuracy of noncompressive sonography of children with appendicitis according to the potential positions of the appendix. *AJR Am J Roentgenol* 2000;175(5):1387-92.

34. Bates MF, Khander A, Steigman SA, et al. Use of white blood cell count and negative appendectomy rate. *Pediatrics* 2014;133(1):e39-44.

2. Becker T, Kharbanda A, Bachur R. Atypical clinical features of pediatric appendicitis. *Acad Emerg Med* 2007;14(2):124-9.

72. Bhangu A, Soreide K, Di Saverio S, et al. Acute appendicitis: modern understanding of pathogenesis, diagnosis, and management. *Lancet* 2015;386(10000):1278-87.

128. Bhangu A. Safety of short, in-hospital delays before surgery for acute appendicitis: multicentre cohort study, systematic review, and meta-analysis. *Ann Surg* 2014;259(5):894-903.

43. Bhatt M, Joseph L, Ducharme FM, et al. Prospective validation of the pediatric appendicitis score in a Canadian pediatric emergency department. *Acad Emerg Med* 2009;16(7):591-6.

25. Bickell NA, Aufses AH, Jr., Rojas M, et al. How time affects the risk of rupture in appendicitis. *J Am Coll Surg* 2006;202(3):401-6.

19. Birnbaum BA, Wilson SR. Appendicitis at the millennium. *Radiology* 2000;215(2):337-48.

24. Blumfield E, Nayak G, Srinivasan R, et al. Ultrasound for differentiation between perforated and nonperforated appendicitis in pediatric patients. *AJR Am J Roentgenol* 2013;200(5):957-62.

103. Bonadio W, Brazg J, Telt N, et al. Impact of In-Hospital Timing to Appendectomy on Perforation Rates in Children with Appendicitis. *J Emerg Med* 2015;49(5):597-604.

116. Boomer LA, Cooper JN, Deans KJ, et al. Does delay in appendectomy affect surgical site infection in children with appendicitis? *J Pediatr Surg* 2014;49(6):1026-9; discussion 29.

107. Brender JD, Marcuse EK, Koepsell TD, et al. Childhood appendicitis: factors associated with perforation. *Pediatrics* 1985;76(2):301-6.

18. Brennan GD. Pediatric appendicitis: pathophysiology and appropriate use of diagnostic imaging. *Cjem* 2006;8(6):425-32.

60. Brenner D, Elliston C, Hall E, et al. Estimated risks of radiation-induced fatal cancer from pediatric CT. *AJR Am J Roentgenol* 2001;176(2):289-96.

64. Brenner DJ. Estimating cancer risks from pediatric CT: going from the qualitative to the quantitative. *Pediatr Radiol* 2002;32(4):228-1; discussion 42-4.

65. Brody AS, Frush DP, Huda W, et al. Radiation risk to children from computed tomography. *Pediatrics* 2007;120(3):677-82.

31. Bundy DG, Byerley JS, Liles EA, et al. Does this child have appendicitis? *Jama* 2007;298(4):438-51.

59. Burr A, Renaud EJ, Manno M, et al. Glowing in the dark: time of day as a determinant of radiographic imaging in the evaluation of abdominal pain in children. *J Pediatr Surg* 2011;46(1):188-91

121. Busch M, Gutzwiller FS, Aellig S, et al. In-hospital delay increases the risk of perforation in adults with appendicitis. *World J Surg* 2011;35(7):1626-33.

20. Carr NJ. The pathology of acute appendicitis. *Ann Diagn Pathol* 2000;4(1):46-58.

85. Chiang JJY, Angus MI, Nah SA, et al. Time course response of inflammatory markers in pediatric appendicitis. *Pediatr Surg Int* 2020;36(4):493-500.

114. Clyde C, Bax T, Merg A, et al. Timing of intervention does not affect outcome in acute appendicitis in a large community practice. *Am J Surg* 2008;195(5):590-2; discussion 92-3.

83. Cohen B, Bowling J, Midulla P, et al. The non-diagnostic ultrasound in appendicitis: is a non-visualized appendix the same as a negative study? *J Pediatr Surg* 2015;50(6):923-7.

47. Coyne SM, Zhang B, Trout AT. Does appendiceal diameter change with age? A sonographic study. *AJR Am J Roentgenol* 2014;203(5):1120-6.

39. de Castro SMM, Geerdink TH, Macco S, et al. Mandatory imaging in the work-up of children suspected of having appendicitis reduces the rate of unnecessary surgeries. *J Pediatr Surg* 2018.

126. Ditillo MF, Dziura JD, Rabinovici R. Is it safe to delay appendectomy in adults with acute appendicitis? *Ann Surg* 2006;244(5):656-60.

67. Dorfman AL, Fazel R, Einstein AJ, et al. Use of medical imaging procedures with ionizing radiation in children: a population-based study. *Arch Pediatr Adolesc Med* 2011;165(5):458-64.

6. Doria AS, Moineddin R, Kellenberger CJ, et al. US or CT for Diagnosis of Appendicitis in Children and Adults? A Meta-Analysis. *Radiology* 2006;241(1):83-94.

62. Doria AS. Optimizing the role of imaging in appendicitis. *Pediatr Radiol* 2009;39 Suppl 2:S144-8.

22. Drake FT, Mottey NE, Castelli AA, et al. Time-of-day and appendicitis: Impact on management and outcomes. *Surgery* 2017;161(2):405-14.

108. Dunlop JC, Meltzer JA, Silver EJ, et al. Is nonperforated pediatric appendicitis still considered a surgical emergency? A survey of pediatric surgeons. *Acad Pediatr* 2012;12(6):567-71.

125. Eko FN, Ryb GE, Drager L, et al. Ideal timing of surgery for acute uncomplicated appendicitis. *N Am J Med Sci* 2013;5(1):22-7.

57. Emil S, Mikhail P, Laberge JM, et al. Clinical versus sonographic evaluation of acute appendicitis in children: a comparison of patient characteristics and outcomes. *J Pediatr Surg* 2001;36(5):780-3.

123. Fair BA, Kubasiak JC, Janssen I, et al. The impact of operative timing on outcomes of appendicitis: a National Surgical Quality Improvement Project analysis. *Am J Surg* 2015;209(3):498-502.

13. Fitz RH. *Perforating inflammation of the vermiform appendix, with special reference to its early diagnosis and treatment*. Philadelphia: Dornan; 1886.

45. Garcia Pena BM, Mandl KD, Kraus SJ, et al. Ultrasonography and limited computed tomography in the diagnosis and management of appendicitis in children. *Jama* 1999;282(11):1041-6.

111. Giraudo G, Baracchi F, Pellegrino L, et al. Prompt or delayed appendectomy? Influence of timing of surgery for acute appendicitis. *Surg Today* 2013;43(4):392-6.

15. Graham JM, Pokorny WJ, Harberg FJ. Acute appendicitis in preschool age children. *Am J Surg* 1980;139(2):247-50.

74. Gronroos JM. Do normal leucocyte count and C-reactive protein value exclude acute appendicitis in children? *Acta Paediatr* 2001;90(6):649-51.

35. Gronroos P, Huhtinen H, Gronroos JM. Normal leukocyte count and C-reactive protein value do not effectively exclude acute appendicitis in children *Dis Colon Rectum*. United States; 2009 p1028-9; author reply 29.

17. Grosfeld JL, Weinberger M, Clatworthy HW, Jr. Acute appendicitis in the first two years of life. *J Pediatr Surg* 1973;8(2):285-93.

102. Gurien LA, Wyrick DL, Smith SD, et al. Optimal timing of appendectomy in the pediatric population. *J Surg Res* 2016;202(1):126-31.

86. Hahn HB, Hoepner FU, Kalle T, et al. Sonography of acute appendicitis in children: 7 years experience. *Pediatr Radiol* 1998;28(3):147-51.

61. Hall EJ. Lessons we have learned from our children: cancer risks from diagnostic radiology. *Pediatr Radiol* 2002;32(10):700-6.

68. Hall MJ, DeFrances CJ, Williams SN, et al. National Hospital Discharge Survey: 2007 summary. *Natl Health Stat Report* 2010(29):1-20, 24.

71. Hall NJ, Eaton S. Non-operative management of appendicitis in children. *Arch Dis Child* 2018;103(5):498-502.

80. Hofmann V, Deeg K-H, Hoyer PF. Ultraschalldiagnostik in Pädiatrie und Kinderchirurgie. *Dtsch Arztebl International* 2005;102(50):3503-.

97. Honda T, Uehara T, Matsumoto G, et al. Neutrophil left shift and white blood cell count as markers of bacterial infection. *Clin Chim Acta* 2016;457:46-53.

56. Hormann M, Scharitzer M, Stadler A, et al. Ultrasound of the appendix in children: is the child too obese? *Eur Radiol* 2003;13(6):1428-31.

50. Howell JM, Eddy OL, Lukens TW, et al. Clinical policy: Critical issues in the evaluation and management of emergency department patients with suspected appendicitis. *Ann Emerg Med* 2010;55(1):71-116.

124. Ingraham AM, Cohen ME, Bilimoria KY, et al. Effect of delay to operation on outcomes in adults with acute appendicitis. *Arch Surg* 2010;145(9):886-92.

79. Jeffrey RB, Jr., Laing FC, Townsend RR. Acute appendicitis: sonographic criteria based on 250 cases. *Radiology* 1988;167(2):327-9.

55. Josephson T, Styrud J, Eriksson S. Ultrasonography in acute appendicitis. Body mass index as selection factor for US examination. *Acta Radiol* 2000;41(5):486-8.

92. Kashtan M, Graham D, Anandalwar S, et al. Influence of symptom duration and WBC profile on the negative predictive value of a nondiagnostic ultrasound in children with suspected appendicitis. *J Pediatr Surg* 2020.

117. Kearney D, Cahill RA, O'Brien E, et al. Influence of delays on perforation risk in adults with acute appendicitis. *Dis Colon Rectum* 2008;51(12):1823-7.

78. Kessler N, Cyteval C, Gallix B, et al. Appendicitis: evaluation of sensitivity, specificity, and predictive values of US, Doppler US, and laboratory findings. *Radiology* 2004;230(2):472-8.

30. Kremer K, Kraemer M, Fuchs KH, et al. [The diagnostic value of rectal examination of patients with acute appendicitis]. *Langenbecks Arch Chir Suppl Kongressbd* 1998;115:1120-2.

88. Krishnamoorthi R, Ramarajan N, Wang NE, et al. Effectiveness of a staged US and CT protocol for the diagnosis of pediatric appendicitis: reducing radiation exposure in the age of ALARA. *Radiology* 2011;259(1):231-9.

77. Lee MW, Kim YJ, Jeon HJ, et al. Sonography of acute right lower quadrant pain: importance of increased intraabdominal fat echo. *AJR Am J Roentgenol* 2009;192(1):174-9.

131. Lee SL, Islam S, Cassidy LD, et al. Antibiotics and appendicitis in the pediatric population: an American Pediatric Surgical Association Outcomes and Clinical Trials Committee systematic review. *J Pediatr Surg* 2010;45(11):2181-5.

11. Lee SL, Stark R, Yaghoubian A, et al. Does age affect the outcomes and management of pediatric appendicitis? *J Pediatr Surg* 2011;46(12):2342-5.

44. Lintula H, Kokki H, Kettunen R, et al. Appendicitis score for children with suspected appendicitis. A randomized clinical trial. *Langenbecks Arch Surg* 2009;394(6):999-1004.

27. Lipsett SC, Bachur RG. Current Approach to the Diagnosis and Emergency Department Management of Appendicitis in Children. *Pediatr Emerg Care* 2017;33(3):198-203.

82. Malia L, Sturm JJ, Smith SR, et al. Diagnostic accuracy of laboratory and ultrasound findings in patients with a non-visualized appendix. *Am J Emerg Med* 2019;37(5):879-83.

91. Malia L, Sturm JJ, Smith SR, et al. Predictors for Acute Appendicitis in Children. *Pediatr Emerg Care* 2019.

3. Mallick MS. Appendicitis in pre-school children: a continuing clinical challenge. A retrospective study. *Int J Surg* 2008;6(5):371-3.

7. Mallin M, Craven P, Ockerse P, et al. Diagnosis of appendicitis by bedside ultrasound in the ED. *Am J Emerg Med* 2015;33(3):430-2.

115. Mandeville K, Monuteaux M, Pottker T, et al. Effects of Timing to Diagnosis and Appendectomy in Pediatric Appendicitis. *Pediatr Emerg Care* 2015;31(11):753-8.

130. Meltzer JA, Kunkov S, Chao JH, et al. Association of Delay in Appendectomy With Perforation in Children With Appendicitis. *Pediatr Emerg Care* 2019;35(1):45-49.

132. Minneci PC, Hade EM, Lawrence AE, et al. Multi-institutional trial of non-operative management and surgery for uncomplicated appendicitis in children: Design and rationale. *Contemp Clin Trials* 2019;83:10-17.

49. Mittal MK, Dayan PS, Macias CG, et al. Performance of ultrasound in the diagnosis of appendicitis in children in a multicenter cohort. *Acad Emerg Med* 2013;20(7):697-702.

105. Miyauchi H, Okata Y, Hatakeyama T, et al. Analysis of predictive factors of perforated appendicitis in children. *Pediatr Int* 2020.

38. Mostbeck G, Adam EJ, Nielsen MB, et al. How to diagnose acute appendicitis: ultrasound first. *Insights Imaging* 2016;7(2):255-63.

10. Nance ML, Adamson WT, Hedrick HL. Appendicitis in the young child: a continuing diagnostic challenge. *Pediatr Emerg Care* 2000;16(3):160-2.

12. Narsule CK, Kahle EJ, Kim DS, et al. Effect of delay in presentation on rate of perforation in children with appendicitis. *Am J Emerg Med* 2011;29(8):890-3.

66. National Research C. *Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2*. Washington, DC: The National Academies Press; 2006.

23. Nelson DS, Bateman B, Bolte RG. Appendiceal perforation in children diagnosed in a pediatric emergency department. *Pediatr Emerg Care* 2000;16(4):233-7.

106. Obinwa O, Peirce C, Cassidy M, et al. A model predicting perforation and complications in paediatric appendicectomy. *Int J Colorectal Dis* 2015;30(4):559-65.

69. Oyetunji TA, Ong'uti SK, Bolorunduro OB, et al. Pediatric negative appendectomy rate: trend, predictors, and differentials. *J Surg Res* 2012;173(1):16-20.

26. Papandria D, Goldstein SD, Rhee D, et al. Risk of perforation increases with delay in recognition and surgery for acute appendicitis. *J Surg Res* 2013;184(2):723-9.

95. Partain KN, Patel A, Travers C, et al. Secondary signs may improve the diagnostic accuracy of equivocal ultrasounds for suspected appendicitis in children. *J Pediatr Surg* 2016;51(10):1655-60.

94. Partain KN, Patel AU, Travers C, et al. Association of Duration of Symptoms and Secondary Signs in Ultrasound for Pediatric Appendicitis. *Am Surg* 2016;82(9):e266-8.

63. Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet* 2012;380(9840):499-505.

133. Podda M, Gerardi C, Cillara N, et al. Antibiotic Treatment and Appendectomy for Uncomplicated Acute Appendicitis in Adults and Children: A Systematic Review and Meta-analysis. *Ann Surg* 2019;270(6):1028-40.

98. Póvoa P. C-reactive protein: a valuable marker of sepsis. *Intensive Care Med* 2002;28(3):235-43.

93. Prada-Arias M, Gómez-Veiras J, Vázquez JL, et al. Appendicitis or non-specific abdominal pain in pre-school children: When to request abdominal ultrasound? *J Paediatr Child Health* 2020;56(3):367-71.

21. Puig S, Staudenherz A, Felder-Puig R, et al. Imaging of appendicitis in children and adolescents: useful or useless? A comparison of imaging techniques and a critical review of the current literature. *Semin Roentgenol* 2008;43(1):22-8.

87. Puylaert JB, Rutgers PH, Lalisang RI, et al. A prospective study of ultrasonography in the diagnosis of appendicitis. *N Engl J Med* 1987;317(11):666-9.

46. Puylaert JB. Acute appendicitis: US evaluation using graded compression. *Radiology* 1986;158(2):355-60.

89. Ramarajan N, Krishnamoorthi R, Barth R, et al. An interdisciplinary initiative to reduce radiation exposure: evaluation of appendicitis in a pediatric emergency department with clinical assessment supported by a staged ultrasound and computed tomography pathway. *Acad Emerg Med* 2009;16(11):1258-65.

75. Reddan T, Corness J, Mengersen K, et al. Ultrasound of paediatric appendicitis and its secondary sonographic signs: providing a more meaningful finding. *J Med Radiat Sci* 2016;63(1):59-66.

96. Rosales C, Demaurex N, Lowell CA, et al. Neutrophils: Their Role in Innate and Adaptive Immunity. *J Immunol Res* 2016;2016:1469780.

81. Ryan C. Sonographic Evaluation of Acute Appendicitis. *Journal of Diagnostic Medical Sonography* 2012;28(6):320-23.

99. Sack U, Biereder B, Elouahidi T, et al. Diagnostic value of blood inflammatory markers for detection of acute appendicitis in children. *BMC Surg* 2006;6:15.

90. Sakellaris G, Tilemis S, Charissis G. Acute appendicitis in preschool-age children. *Eur J Pediatr* 2005;164(2):80-3.

41. Samuel M. Pediatric appendicitis score. *J Pediatr Surg* 2002;37(6):877-81.

42. Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring systems using a prospective pediatric cohort. *Ann Emerg Med* 2007;49(6):778-84, 84.e1.

9. Schuh S, Man C, Cheng A, et al. Predictors of non-diagnostic ultrasound scanning in children with suspected appendicitis. *J Pediatr* 2011;158(1):112-8.

84. Serres SK, Cameron DB, Glass CC, et al. Time to Appendectomy and Risk of Complicated Appendicitis and Adverse Outcomes in Children. *JAMA Pediatr* 2017;171(8):740-46.

33. Shogilev DJ, Duus N, Odom SR, et al. Diagnosing appendicitis: evidence-based review of the diagnostic approach in 2014. *West J Emerg Med* 2014;15(7):859-71.

52. Sivit CJ, Applegate KE, Stallion A, et al. Imaging evaluation of suspected appendicitis in a pediatric population: effectiveness of sonography versus CT. *AJR Am J Roentgenol* 2000;175(4):977-80.

18. Sivit CJ, Applegate KE. Imaging of acute appendicitis in children. *Semin Ultrasound CT MR* 2003;24(2):74-82.

8. Smink DS, Finkelstein JA, Garcia Pena BM, et al. Diagnosis of acute appendicitis in children using a clinical practice guideline. *J Pediatr Surg* 2004;39(3):458-63; discussion 58-63.

110. Steinbrook R. The debate over residents' work hours. *N Engl J Med* 2002;347(16):1296-302.

101. Stevenson MD, Dayan PS, Dudley NC, et al. Time From Emergency Department Evaluation to Operation and Appendiceal Perforation. *Pediatrics* 2017;139(6).

70. Stringer MD. Acute appendicitis. *J Paediatr Child Health* 2017;53(11):1071-76.

16. Surana R, Quinn F, Puri P. Appendicitis in preschool children. *Pediatric Surgery International* 1995;10(2):68-70.

129. Surana R, Quinn F, Puri P. Is it necessary to perform appendicectomy in the middle of the night in children? *Bmj* 1993;306(6886):1168.

109. Taffinder NJ, McManus IC, Gul Y, et al. Effect of sleep deprivation on surgeons' dexterity on laparoscopy simulator *Lancet*. England; 1998 p1191.

127. Taylor M, Emil S, Nguyen N, et al. Emergent vs urgent appendectomy in children: a study of outcomes. *J Pediatr Surg* 2005;40(12):1912-5.

122. Teixeira PG, Sivrikoz E, Inaba K, et al. Appendectomy timing: waiting until the next morning increases the risk of surgical site infections. *Ann Surg* 2012;256(3):538-43.

54. Toorenvliet B, Vellekoop A, Bakker R, et al. Clinical differentiation between acute appendicitis and acute mesenteric lymphadenitis in children. *Eur J Pediatr Surg* 2011;21(2):120-3.

76. van Breda Vriesman AC, Puylaert JB. Mimics of appendicitis: alternative nonsurgical diagnoses with sonography and CT. *AJR Am J Roentgenol* 2006;186(4):1103-12.

36. Wang LT, Prentiss KA, Simon JZ, et al. The use of white blood cell count and left shift in the diagnosis of appendicitis in children. *Pediatr Emerg Care* 2007;23(2):69-76.

112. Weinger MB, Ancoli-Israel S. Sleep deprivation and clinical performance. *Jama* 2002;287(8):955-7.

113. Wesnes KA, Walker MB, Walker LG, et al. Cognitive performance and mood after a weekend on call in a surgical unit. *Br J Surg* 1997;84(4):493-5.

48. Wiersma F, Sramek A, Holscher HC. US features of the normal appendix and surrounding area in children. *Radiology* 2005;235(3):1018-22.

53. Wiersma F, Toorenvliet BR, Bloem JL, et al. US examination of the appendix in children with suspected appendicitis: the additional value of secondary signs. *Eur Radiol* 2009;19(2):455-61.

119. Yardeni D, Hirschl RB, Drongowski RA, et al. Delayed versus immediate surgery in acute appendicitis: do we need to operate during the night? *J Pediatr Surg* 2004;39(3):464-9; discussion 64-9.

Appendix

Gesundheit Nordhessen

Patient ID	Date of Birth	Gender	Date of Admission	Time of Admission	Coincidence of Infection
1.					
2.					
3.					
4.					
5.					
6.					
7.					
8.					
9.					
10.					
11.					
12.					
13.					
14.					
15.					
16.					
17.					
18.					
19.					
20.					
21.					
22.					
23.					
24.					
25.					

Data Collection Sheet: Patient Demographics

Gesundheit
Nordhessen

Patient ID	RLQ	Umbilical	LLQ	McBurney	Lanz	Blumberg	Rovsing	Douglas	Baldwin	'Erschütterungsschmerz'
1.										
2.										
3.										
4.										
5.										
6.										
7.										
8.										
9.										
10.										
11.										
12.										
13.										
14.										
15.										
16.										
17.										
18.										
19.										
20.										
21.										
22.										
23.										
24.										
25.										
26.										

Data Collection Sheet: Symptoms of Pain

Gesundheit
Nordhessen

Patient ID	Temperature 1 = <37.0°C 2 = 37-38.4°C 3 = ≥38.5° 4= No doc.	Nausea 0 = Missing 1 = Present 2 = No doc.	Vomiting 0 = Missing 1 = Present 2 = No doc.	Bowel Movements 1 = Diarrhoea 2 = Obstipation 3 = Missing B.M 4 = Normal 5 = No doc.	WCC upon admission -66 = < 3 2 = No doc.	WCC upon USS -66 = < 3 2 = No doc.	CRP upon admission -66 = < 3 2 = No doc.	CRP upon USS -66 = < 3 2 = No doc.
1.								
2.								
3.								
4.								
5.								
6.								
7.								
8.								
9.								
10.								
11.								
12.								
13.								
14.								
15.								
16.								
17.								
18.								
19.								
20.								
21.								
22.								
23.								

Data Collection Sheet: Clinical Data

Gesundheit
Nordhessen

Patient ID	USS Date / Time	Abdo-Pain Duration	Radiological Report 1 = Normal APP 2 = Positive APP 3 = Equivocal	Histopath. Report 1 = Confirm 0 = Rule Out	Perforation 1 = Present 0 = Missing	Time to Appendectomy	Abdominal Access 1 = Laparascopic. 2 = Open Abdom.
1.							
2.							
3.							
4.							
5.							
6.							
7.							
8.							
9.							
10.							
11.							
12.							
13.							
14.							
15.							
16.							
17.							
18.							
19.							
20.							
21.							
22.							
23.							
24.							

Data Collection Sheet: Ultrasound and Histo-/Pathology Report

Gesundheit
Nordhessen

Patient ID	Duration >120 h	No USS available	No Appendicitis or Diagnosis other than Appendicitis	Psychiatric comorbidity	Prior Admission/Diagnosis	Other GI comorbidity of Exclusion
1.						
2.						
3.						
4.						
5.						
6.						
7.						
8.						
9.						
10.						
11.						
12.						
13.						
14.						
15.						
16.						
17.						
18.						
19.						
20.						
21.						
22.						
23.						
24.						
25.						

Data Collection Sheet: Exclusions

List of Publications

None declared.

Statutory Declaration

Erklärung zur Dissertation

„Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nichtveröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt an der Entstehung der vorliegenden Arbeit beteiligt waren.
Mit der Überprüfung meiner Arbeit durch eine Plagiatserkennungssoftware bzw. ein internetbasiertes Softwareprogramm erkläre ich mich einverstanden.“

Hann. Münden, den 12.10.2021

Ort, Datum	Unterschrift

Acknowledgements

Throughout the drafting, editing and writing of this dissertation I have received a great deal of support and assistance from people that deserve to be mentioned and acknowledged.

I would first like to thank my supervisor, Professor Dr. med. F. K. Tegtmeyer, whose expertise was invaluable in formulating the research question at hand as well as the methodology leading up to this project. Your insightful feedback and comments have given me the guidance needed to finalise and complete my dissertation. Your role as a paediatric consultant with decades of experience has shown me how current this topic still is and the magnitude of scientific gaps that yet need to be filled. Thank you for the encouragement and the opportunities you gave me to broaden my knowledge on this topic. Over and above, I would like to thank you for inspiring me to hold on to a career in academia as well as in clinical medicine.

Moreover, I would like to thank Dr. G. Eichner (Institute of Mathematics, University of Giessen). The meetings and phone conversations we had over the months were vital for me to think outside the box, comprehend the often times complex statistical analyses and served as an objective critique for our methodology. Your sound understanding of these matters were essential and an integral part I could not have completed this work without.

I would also like to thank D. Jaeger (Department of Paediatric Radiology, Klinikum Kassel GNH) whose patient afternoon meetings have given me great support and helped reduce the uncertainty with paediatric ultrasound results as well as technical advice. You contributed thoroughly with helpful comments and provided me with the tools I needed to successfully access my data. Additionally, I would like to express my gratitude to Dr. med. Peter Illing, Carl-Martin Junge and Christine Huslage for helping to provide the research data at hand.

Lastly, I would like to thank my parents for their wise counsel and sympathetic ear. They provided stimulating discussions as well as happy distractions to rest my mind outside of my research.