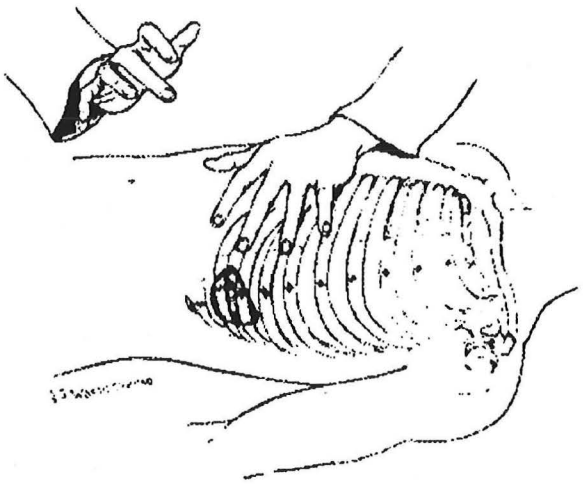
The Abdominal Examination: An Evidence-Based Medicine Approach



February 16, 1995

James M. Wagner, M.D.

Internal Medicine Grand Rounds
University of Texas Southwestern Medical Center
Dallas, Texas

Table of Contents

I. Introduction
- A. Evidence-Based Medicine
- B. The Internist and the Abdomen
- C. Objectives for this Grand Rounds

II. The Evaluation of Diagnostic Tests
- A. A Model of Clinical Judgement
- B. Hypothesis Testing and Bayes Theorem
  1. SnNout = PID
  2. SpPin = NIH
  3. LR+ = Sn/1-Sp

II. Assessing Abdominal Pain
- A. Probable and/or Utilitarian Diagnoses
  1. Admitted Patients
  2. Emergency Department Patients
  3. Ambulatory Patients
- B. Non-Specific Abdominal Pain
  1. Wasson's Rule
  2. Abdominal Wall Tenderness Test
- C. Appendicitis
  1. History
     a. Migration
     b. Pain Before Vomiting
  2. Examination
     a. Rebound Tenderness Test
     b. Psoas/Obturator Signs
     c. Rovsing's Sign
  3. Women of Childbearing age
  4. The Elderly
- C. Cholecystitis

III. Assessing Spleen Size
- A. Nixon Technique
- B. Castell Technique
- C. Barkun Technique

IV. Assessing Liver Size
- A. Percussion/Palpation
- B. Scratch Test

V. Limitations of the Data

VI. Summary

VII. Conclusion

The Abdominal Examination: an Evidence-Based Medicine Approach

In the mid-1960's, medicine was neatly divided into two distinct physician populations: the artists and the scientists. The artists stood at the patient's bedside, and the scientists stood at the laboratory bench. The artists' trade was built on experience, empathy, inference, and compassion. The scientists trade was built on experimentation, data analysis, and scientific scrutiny. Clinical examination techniques were championed by the sage clinician with the most experience. Indeed, several of the leading clinical examination textbooks today serve as monuments to the clinical experience of their single authors[1,2]. The true scientist would ask, "What is the evidence for that recommendation?" and would be dissatisfied when the answer began, "Well, in my experience...".

Although this anecdotal experience lends a healthy perspective, it suffers from recall and referral bias, confounding, and the fickleness of human memory. In his classic 1967 text *Clinical Judgement*[3], Feinstein advocated applying the scientific methods previously reserved for the laboratory to the bedside examination of the patient. Recently, several authors have again called for a more scientifically rigorous evaluation and application of the clinical examination[4,5].

As a response to these calls, scientific evaluation of the clinical examination has been accumulating in the literature for several decades. The application of the principles of Clinical Epidemiology[6] and Clinical Decision Analysis[7] has enriched this database with fresh perspectives. The traditional experience-based-approach to the clinical examination should be replaced by an evidence-based approach.

Evidence-based medicine[8] has become a reality, in part because access to the rich resource of the medical literature has become simple. In addition, the scientific methodology of the laboratory can now be used in a practical way at the bedside[9]. Application of these resources toward answering the daily questions of clinical medicine is the goal of evidence-based medicine.

The Abdomen and the Internist

In practice, the internist often serves as a primary care physician, regardless of subspecialty. In addition, the internist is a consultant for many other specialties. S/he is therefore confronted with the full spectrum of medicine. The training of the internist, however, is often compartamentalized by the triage nurse in the emergency department, "Pain below the xiphoid goes to surgery". The internist is therefore a reluctant participant in the assessment of abdominal complaints, particularly abdominal pain.

There is evidence that even our surgical colleagues could improve upon the clinical evaluation of the abdomen. Many surgeons believe that any increase in diagnostic accuracy is taken at the expense of a higher perforation rate, and hence higher peri- and post-operative morbidity and mortality[10]. This belief has been questioned, however. The highest diagnostic accuracy is obtained when physicians are required to obtain more details from the clinical examination for research purposes[11]. In addition, in large-area variation studies, it is clear that some hospitals can attain the high accuracy and low perforation rate deemed impossible by others[12,11].

The literature documents the need for improved clinical skills[13,14,15]. Frequently observed errors in the abdominal examination include the assessment of abdominal tenderness and organomegaly[16]. Several authors have called for more attention to be spent teaching and observing these skills to trainees[3,13].

The objective of this grand rounds is to review the scientific evidence for the abdominal examination. The review will be presented in the context of the tools used in the laboratory for the practical evaluation of diagnostic testing. Bayes Theorem will allow the scientific calculation of the precision and accuracy of various techniques[17,9]. The data to be presented has often been re-calculated from the original data. This was done to present the data in a uniform and most clinically useful manner.

These methods will be described and demonstrated. The evidence for each recommendation will be cited, along with an estimate of the limitations of the data.

Review of Clinical Judgement

Prowess in the clinical examination is the hallmark of the astute clinician. This clinical examination, consisting of the history and physical examination, allows the clinician to formulate a working diagnosis, differential diagnosis, and judiciously use the technology at his/her disposal. The clinical examination alone produces an accurate diagnosis in 90% of all encounters[18], with the history being the most important of the two components. The physical examination produces a diagnosis in a smaller number of encounters, but also enhances the confidence in diagnoses made from the history in a larger proportion[18,19].

When tending to a patient, the astute physician gathers clinical facts, then assimilates them into a working differential diagnosis. This working list is formulated by considering Occam's Razor and the theories of probability and utility, and hypothesis testing[20]. **Occam's Razor** is a dictum that states the best explanation of a constellation of findings is a single diagnosis. This heuristic is helpful, but not always correct.

The **theory of probability** states that the physician must be aware of the most likely diagnoses of a given symptom complex. Ignorance of any particular diagnosis will amount to omission of it in the differential, and perhaps unnecessary testing and risk to the patient. The **theory of utility** states that the physician must also be aware of those diagnoses, regardless how uncommon, must not be missed, as either treatment is particularly effective or expensive, or missing the diagnosis particularly catastrophic. **Hypothesis testing** occurs as the physician tests the validity of any postulated diagnosis with additional questions to the patient, additional physical examination techniques.

Once the probability of a diagnosis for the appropriate clinical setting is determined, the physician conducts the clinical examination. This probability of disease prior to the examination is termed the **pre-test probability** in the rubric of scientific diagnostic testing. The clinician then uses the findings of the clinical examination to modify this probability into a **post-test probability**.

The clinical examination is actually a series of diagnostic tests. Each test is an experiment in hypothesis testing. Each experiment adds or detracts from the probability of any particular diagnosis. This grand rounds will quantitate how much each diagnostic test adds or detracts from any particular diagnosis.

The Evaluation of a Diagnostic Test

Consider the senses of the physician the basic instruments of his/her trade. With these instruments the astute clinician feels, percusses, sees, smells, and hears the clues to the diagnosis and resolution of the patient's suffering. Using these instruments, the clinical examination is a sequence of diagnostic tests. They are the effectors of the Psoas and Obturator signs, Rebound tenderness, Rovsing's sign. They percuss Traube's Space and palpate the edges of the solid organs in the abdomen.

Diagnostic tests are measured in terms of sensitivity, specificity, positive and negative predictive values, and likelihood ratios in various clinical situations. Bayes Theorem is at the root of these measurements. A detailed discussion of the history and derivation of the practical applications of this theorem is beyond the scope of this grand rounds. A review of the practical application of Bayes Theorem will be presented.

The value of Bayes Theorem is often lost in its mathematical equations. Its value lies in the ability to measure the accuracy and precision of diagnostic tests. When aware of and armed with these parameters, a physician can determine with confidence when further diagnostic testing is needed. This, in turn, allows for the judicious use of the more expensive technologies at the physician's disposal.

A helpful mnemonic for calculating sensitivity and specificity has been suggested by Sackett[6] and is depicted in Figures 1 and 2.

Figure 1

SnNout = PID

A highly sensitive test rules OUT a disease when negative (SnNout). Sensitivity is the frequency that the test is Positive in Disease (PID).

Figure 2

SpPin = NIH

A highly specific test rules IN a disease when positive (SpPin). Specificity is the frequency that the test is negative in health (NIH).

Table 1

		Disease +	Disease -	
Test	+	a	b	a+b
	-	c	d	c+d
		a+c	b+d	

Sensitivity = a/a+c
Specificity = d/b+d
Prevalence of dz = a+c/b+d
Positive Predictive Value (PPV) = a/a+b
Negative Predictive Value (NPV) = d/c+d
Accuracy = a+d/a+b+c+d
Likelihood Ratio of a Positive Test (LR+)= Sensitivity/1-Specificity
Likelihood Ratio of a Negative Test (LR-) = Specificity/1-Sensitivity

At the root of Bayes Theorem is the 2X2 table. As depicted in Figures 1 and 2, sensitivity and specificity can be calculated from this table and defined in a practical and easily remembered manner. Sensitivity is used to rule OUT a disease (SnNout), because it is the frequency that a test is positive in disease (PID). For example, if we know that a leukocytosis has a high specificity for appendicitis, we might be less likely to commit operating room resources to a patient with a normal leukocyte count. The highly **Sen**sitive WBC count, being **N**egative, rules **out** the disease.

Specificity is use to rule IN a disease (SpPin), because it is the frequency that a test is negative in health (NIH). For example, if Rovsing's sign is highly **Spe**cific for appendicitis, a **P**ositive sign rules **in** the disease.

Sensitivity and specificity are not influenced by the prevalence of disease. This makes these parameters consistent throughout different populations. The performance of any diagnostic test, however, is clearly influenced by the prevalence of disease. For these reasons, the positive and negative predictive values (PPV and NPV) may be seen as more useful parameters, but they change from population to population. for this reason, PPV and NPV are less useful.

Likelihood ratios (LR) can be calculated from sensitivity and specificity. LRs can be used in any population to modify the probability of disease, as demonstrated in the nomogram in Figure 3.

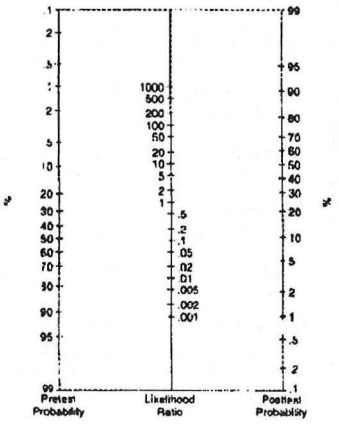


Figure 3

Evaluation of Abdominal Pain

The clinician uses the theories of probability and utility instinctively. Hypothesis testing occurs in the history and physical. In our model of clinical judgement, the most probable and utilitarian diagnoses must be quantified. The probability (prevalence) of any particular disease varies with the population studied. Several studies have been performed on patients admitted for abdominal pain[21,22,23,24,25], patients evaluated in emergency department for abdominal pain[26,27,28], and patients evaluated in the walk-in clinics for abdominal pain[29,30,31]. The probability of disease in these studies are summarized in Tables 3-5.

The studies of admitted patients select for sicker patients. The criteria for admission were not clearly elucidated in any of the studies, so the characteristics of a "sicker" patient cannot be defined. Nonetheless, several points can be derived from Table 3. First, the diagnosis that carries the largest proportion is Non-Specific Abdominal Pain (NSAP). Most of the in-patient studies required a pathological specimen to confirm the diagnosis, so many diagnoses fall into this category (e.g. gastroenteritis, Pelvic Inflammatory Disease [PID], or Irritable Bowel Syndrome).

The second most probable diagnosis was clearly appendicitis, making this an important diagnosis from both a probabilistic and utilitarian standpoint. This diagnosis will be discussed in more detail in another section. The remaining probable diagnoses are lead by intestinal obstruction and cholecystitis, followed by perforated peptic ulcer, pancreatitis, diverticulitis, and PID.

The utility of these diagnoses is apparent. Laparotomy and removal or repair of the diseased organ or tissue is the standard of care for most of these diagnoses (excluding pancreatitis, most diverticulitis and gynecological disorders).

Table 3
Probability (%) of Top 8 Diagnoses in Admitted Populations

Author	NSAP	APP	GYN	OBST	CHOL	PERF	PANC	DIVT
Staniland	17	17	-	8	17	17	8	17
deDombal	51	26	-	3.6	7.6	3.1	2.9	2.0
Jess	32	20	9	5	6	2	5	3
Thomson	63	19	-	4	3	4	-	-
Irvin	35	17	1.1	15	5	3	2	4
Simmens	39	20	1	11	11	4	1.7	12

NSAP is non-specific abdominal pain; APP is appendicitis; OBST is obstruction; GYN is a gynecological source of pain; PERF is perforated peptic ulcer; PANC is pancreatitis; DIVT is diverticulitis.

Knowing the probability of disease in the hospitalized patient is less useful when the physician is first encountering the patient with abdominal pain. Table 4 lists the probabilities of diseases presenting as abdominal pain to the emergency department. Most studies define the abdominal

pain as occurring for less than one week. NSAP has become more prevalent when compared to the in-patient group, and the highly utilitarian diagnoses requiring emergent surgical intervention are becoming diluted out by the less urgent, but still highly utilitarian diagnoses of PID, urinary tract infections (UTI), urolithiasis, and peptic ulcer disease.

Table 4
Probability (%) of Top 8 Diagnoses in the Emergency Department

Author	NSAP	GYN	GE	UROL	APP	OBST	CHOL	PUD
Brewer	41	9.5	7.3	11.2	4	2.5	2.5	2
Wilson	47*			3.2	16	2.6	6	2.3

NSAP is non-specific abdominal pain; GYN indicates a gynecological source of pain, most commonly PID, dysmenorrhea, ovarian cyst; GE is gastroenteritis; UROL indicates a urological source of pain, most commonly a UTI or stone; APP is appendicitis; OBST is obstruction.
* Wilson's study required a pathological specimen for diagnosis, hence NSAP usually included gynecological sources of pain and GE.

Patients presenting for evaluation of abdominal pain to the walk-in clinics represent a less ill subset. As illustrated in Table 5, NSAP, as defined in the previous studies, has increased in prevalence. Several highly utilitarian and emergent diagnoses have dropped off the list, being replaced by diagnoses that are usually self-limited (Irritable Bowel Syndrome [IBS}, Gastroenteritis [GE], Hepatitis).

Table 5
Probability (%) of Top 8 Diagnoses in the Ambulatory Clinic

Author	NSAP	IBS	GYN	GE	PUD	HEP	UROL	APP
Wasson	79*				5	3.3	1.6	1
Orient	33	7	13	7	8	2	13	2
Britt	16	6.3	-	11.4	7.4	-	4	1.6

NSAP is non-specific abdominal pain; IBS indicates Irritable Bowel Syndrome; GYN indicates a gynecological source of pain, most commonly PID, dysmenorrhea, ovarian cyst; GE is gastroenteritis; PUD indicated radiologically proven ulcer; HEP is hepatitis; UROL indicates a urological source of pain, most commonly a UTI or stone; APP is appendicitis.
* Wasson's study required a pathological specimen for diagnosis, hence NSAP usually included gynecological sources of pain and GE.

The elderly is another population to be considered. As shown in Table 6, the probability of various diseases has changed again. Appendicitis is much less prevalent, NSAP continues to be prevalent (but less so), and cholecystitis, obstruction, and diverticulitis are now leading the list. Cancer has also become much more prevalent.

Table 6

Probability (%) of Top 7 Diagnoses in the Elderly Population

Author	NSAP	CHOL	OBST	DIVT	APP	UROL	CANC
Fenyo	18.6	40.8	7.3	3.4	6.7	-	-
Fenyo	9.6	26	10.7	7.0	3.5	-	13.2
Irvin	22.5	8.9	28	8.5	4.2	3.2	5.5
Bugliosi	35	12	12	6	4	6	-

Across these populations, NSAP is the most prevalent (or probable) diagnosis. Appendicitis is both prevalent and utilitarian, since surgery can markedly reduce morbidity and mortality. For these reasons, these two diagnoses will be the focus of the discussion.

Non-Specific Abdominal Pain (NSAP)

Once the appropriate probability of any particular diagnosis is realized, the clinician proceeds with testing that diagnosis with the history and clinical examination

Given the apparent high probability of this diagnosis, a more detailed description is warranted. Several studies have documented the favorable prognosis of this category of patients. Lukens reported a 0% mortality of 403 patients diagnosed in the emergency department on 2 week follow-up[32]. Wasson reported that of patients with a diagnosis of NSAP, 4.6% required admission and 1.2% required surgery within one week of initial evaluation[29].

Jess reported that of 230 patients diagnosed with NSAP, only 18 (8%) were found to have a diagnosis causing abdominal pain requiring intervention over a 5 year period. These diagnoses included gynecological disorders in eight, appendicitis in five, diverticulitis in two, duodenal ulcer in two, and one each of Hodgkin's Disease and Acute Intermittent Porphyria. Mortality of one patient may have been related to the original presentation with NSAP, as she latter presented again with NSAP, was found to have leukemic infiltration of the skin On autopsy, she was found to have leukemic infiltration of the stomach[23].

Wasson's Rule[29] is the best published decision rule for identifying patients with a low likelihood for serious disease, and hence less likely to benefit from further diagnostic procedures. Wasson's Rule has a positive predictive value of 95% for correctly identifying patients with low-risk abdominal pain. Although the rule is somewhat cumbersome, it illustrates some important points in the clinical evaluation of patients with abdominal pain. This rule is outlined in Figure 4.

Figure 4

Wasson's Rule for NSAP[29]

If the history or physical reveals:	**Assign a score of:**
1. Pain has lasted at least six months or there has been at least 10 previous occurrences of pain	-3
2. Epigastric pain and tenderness	-2
3. Age 60 or over	+4
4. Pain is constant or unrelieved by any food or medication	+4
5. Pain has affected sleep	+3
6. Abdominal examination reveals any of the following: mass, rigidity, rebound, distention, absent bowel sounds, abnormal liver	+3
7. Stool contains occult blood	+3
8. Vomiting has occurred since pain began	+2
9. Weight loss (10 lbs. or more)	+1
10. History or cancer, diverticular dz, pancreatitis, gallstones or inflammatory bowel dz.	+6
Subtotal:	
Subtract:	-5
Total:	

Rule: total score less than zero means a low likelihood of serious disease.

Wasson's Rule has a sensitivity of 44% and a specificity of 89%. This makes the likelihood ratio 4, one of the stronger modifiers of pre-test probability. If diagnoses of peptic ulcer disease are excluded, the specificity rises to greater than 97%. The clinical application of this information is as follows. Given the high specificity, if the rule indicates that NSAP is likely, NSAP is more likely (SpPin), and serious disease less likely. Given the mediocre sensitivity, however, ruling out NSAP is less powerful (SnNout). In other words, a positive Wasson's Rule is helpful, but a negative Rule is not.

The abdominal wall may be under-recognized as a source of pain in NSAP. Several authors indicate this may be the source in the majority of patients within this category[33,34,35,36,37]. The **Abdominal-Wall Tenderness Test (AWT)** was first described by Carnett[38] in 1926 and has subsequently been studied and found to be useful. Thomson[36] describes the maneuver as follows: "the examiner sat at the bedside with the patient lying flat and relaxed; he palpated the abdomen until the tender area was found and then, with the examining hand remaining in place, asked the patient to cross arms and sit forward; with the patient midway between sitting and recumbency and anterior-abdominal-wall muscles tensed, the examiner palpated the previously tender spot. If the tenderness was reduced the test was recorded as negative; if made worse, as positive".

Thomson went on to validate the AWT test prospectively in 120 patients and found it to be highly sensitive and specific (96% and 100% respectively). The test incorrectly classified one patient as having abdominal wall pain when they actually had appendicitis, but the inflamed appendix was found to be adhered to the anterior abdominal wall at surgery.

The power of this test has not been fully evaluated, nor confirmed in subsequent studies. If the AWT test is as powerful as this study indicates, recognition of the abdominal wall as a source of pain could result in effectively focusing therapy with a successful outcome for the patient[33]. Some authors also speculate that fewer resources are needed to evaluate these patients, and hence the economic impact of routinely evaluating patients with abdominal pain with the AWT test would be beneficial[39,34].

Ruling out serious disease with tools like Wasson's Rule and the AWT test will provide confidence when not pursuing further diagnostic testing. Benefits include not subjecting the patient to potentially harmful tests and procedures. It will also encourage more judicious use of limited medical resources. This technique will not, however, delineate the best approach to the patient once serious disease is suspected. This requires ruling in the specific disease that is causing the serious illness. Two of the more common and best studied diseases will be reviewed.

Appendicitis

Appendicitis was first described by Fitz in 1886[40]. Appendicitis is a diagnosis of importance from the standpoint of both the theory of probability and utility. As demonstrated above, this diagnosis carries a significant prevalence in most populations studied. In addition, unrecognized appendicitis carries a grave prognosis and treatment is profoundly beneficial. In his review of the history of appendicitis, Berry reports Fitz's original mortality rate of 26% has fallen to 0.8%. This triumph can be attributed to increased clinical acumen, improved surgical techniques (Fitz's operative mortality was 40%), and the availability of antibiotics.

The key points of the clinical examination for appendicitis are outlined in the following tables. Table 7 summarizes the studies that have included assessment of key features of the history[21,26,41,42,43], and Table 8, the physical examination[21,26,44, 45,46,41,47,42,10].

Table 7

Appendicitis: Key Points of History

Sensitivity/Specificity

Author	Migration of Pain **Sn/Sp**	Anorexia **Sn/Sp**	Pain a vomit **Sn/Sp**	Pain < 48 hours **Sn/Sp**	No Similar sxs. previously **Sn/Sp**
Staniland		78/30		91/36	82/50
Brewer		72/39	100/64		91/32
Alvarado	69/84	61/72			
Izbicki	43/84				
Nauta	63/67	80/28			

Positive Likelihood Ratios

Author	Migration of Pain	Anorexia	Pain a vomit	Pain < 48 hours	No Similar sxs. prev.
	LR+	**LR+**	**LR+**	**LR+**	**LR+**
Staniland		1.1		1.4	1.6
Brewer		1.18	2.8		1.3
Alvarado	2.7	2.2			
Izbicki	2.7				
Nauta	1.9	1.1			

Several points from the tables should be emphasized. First, the most sensitive findings in the history include three points: 1) a history of pain for less than 48 hours, 2) if vomiting is present, then pain preceded the vomiting, and 3) the absence of similar symptoms previously. Although the power of these three findings in combination has not been assessed, individually they remain powerful. The absence of any of these findings makes it easier to rule out a diagnosis of appendicitis.

The most specific point of the history is the classic history of pain migration. J.B. Murphy originally described the order of symptom occurrence as follows:

1. Pain, usually epigastric or umbilical
2. Anorexia, nausea, or vomiting
3. Tenderness-somewhere in the abdomen or pelvis
4. Fever
5. Leukocytosis.

Murphy stated, "the symptoms occur almost without exception in the above order, and when that order varies I always question the diagnosis"[2]. Cope describes the late tenderness as being localized to the right lower quadrant[2]. Indeed, the experience of these revered physicians has withstood scientific evaluation thus far.

Several key points can be derived from the studies of the physical examination, summarized in Table 8. The most sensitive signs are right lower quadrant tenderness and leukocytosis. Murphy and Cope warn us, however, that leukocytosis occurs late. This dictum is in conflict with what studies have shown, but several of these studies have methodological flaws (to be discussed later) and the anecdotal advice of the experts should not yet be discarded. Nonetheless, it is unusual for appendicitis to occur in the absence of right lower quadrant tenderness. To recap, a sensitive test that is negative influences the physician toward ruling out the disease.

The most specific tests are Rovsing's, the Psoas, and the Obturator Signs. The Obturator sign's characteristics are not shown in the table, but were also evaluated in Izbicki's study and are similar to the other two mentioned. Although the techniques used were not specifically outlined in the study, definitions from Dorland's Illustrated Medical Dictionary should suffice:

Rovsing's Sign: Pressure on the left side over the point corresponding to McBurney's point will elicit the typical pain at McBurney's point in appendicitis.
Psoas Sign: flexion of or pain on hyperextension of the hip due to contact between an inflammatory process and the psoas muscle; a sign often seen in appendicitis. Also called Cope's sign.
Obturator Sign: 1. hypogastric or adductor pain elicited by passive internal rotation of the flexed thigh, due to contact between an inflammatory process and the internal obturator muscle; a sign of appendicitis. 2. a widening and change in contour of the normal obturator x-ray shadow, indicative of pathologic condition of the hip joint; also called the Obturator Sign.
Rebound Tenderness: a sensation of pain felt on the release of pressure. (In this case, pressure in the right lower quadrant of the abdomen)
McBurney's Point: a point between 1 1/2 and 2 inches above the anterior superior spine of the ileum, on a straight line joining that process and the umbilicus, where pressure of the finger elicits tenderness in acute appendicitis.

Table 8

Appendicitis: Key Points of Exam

Sensitivity/Specificity

Author	RLQ Tender	Rebound	Rovsing's Sign	Psoas Sign	Increased WBC
	Sn/Sp	**Sn/Sp**	**Sn/Sp**	**Sn/Sp**	**Sn/Sp**
Staniland	75/95	33/61			
Brewer		84/87			91/64
Nase	90.2/	68/			
Fenyo	94/				
Bongard	72.2/				
Alvarado	100/12	55/78			93/38
Liddington		78/42			
Izbicki	94/2	76/56	22/96*	15/97	72/65
Nauta		67/61			
Berry	96/4	70/39		13/91	

Appendicitis: Key Points of Exam

Positive Likelihood Ratios

Author	RLQ Tender	Rebound	Rovsing's Sign	Psoas Sign	Increased WBC
	LR+	**LR+**	**LR+**	**LR+**	**LR+**
Staniland	15	0.85			
Brewer		6.7			2.5
Alvarado	1.1	2.5			1.5
Liddington		1.9			
Izbicki	0.96	1.7	5.5	5.0	2.1
Nauta		1.7			
Berry	1.0	1.1		1.6	

Appendicitis in Women of Child-Bearing Age

Pelvic Inflammatory Disease (PID) is the most frequent confounder of a diagnosis of appendicitis, as evidenced by the decreased accuracy of the clinical examination for appendicitis in women of child-bearing age[10,48,49,50,51]. Three authors have specified the characteristics of the clinical examination that help discern between the two diagnoses.

Lewis[50] suggested in 1975 that PID can be distinguished from appendicitis by four characteristics: 1) PID has a longer duration of symptoms (2-3 days as opposed to 17 hours), 2) Gastrointestinal symptoms are much less common in PID (patients may even be hungry), 3) Last menstrual period within one week of admission to the hospital, and 4) Chills and fever are twice as common in PID. Unfortunately, the study was retrospective and these results were not reported in a manner that allowed testing of the statistical significance of these findings.

In 1985, the same group, lead by Bongard[46], performed a prospective analysis of 118 patients in whom the diagnosis of PID versus appendicitis was an issue. The clinical findings that occurred in statistically different proportions **favoring appendicitis** were:

1) the presence of nausea, vomiting, or both,
2) the absence of a history of venereal disease,
3) isolated right lower quadrant pain,
4) cervical motion tenderness on exam, and
5) bilateral adnexal tenderness on exam.

Neither the timing of the last menstrual period nor the presence of fever or chills was found to be a predictor in this study. Sensitivity, specificity, and likelihood ratios could not be calculated from the data as presented.

Webster[52] lead the most recent group to address the differences in clinical examination of these two diseases. This study confirmed the findings of Bongard, and adds 1) a history of vaginal discharge or the presence of vaginal discharge on exam, 2) urinary symptoms, and 3) an abnormal urinalysis as predictors of PID. This study was not as methodologically sound as the Bongard study, so the best recommendations to date would be from that study as outlined above.

Appendicitis in the Elderly

Several studies have documented the different presentation and course of appendicitis in the elderly[53,45,25]. All have demonstrated that although the probability of appendicitis declines with age, the sensitivity and specificity also declines, and the morbidity and mortality increases. For these reasons, extreme prudence should be exercised when evaluating elderly patients with abdominal pain, recognizing that benign presentations may often be misleading.

Cholecystitis

In general, cholecystitis is less common than appendicitis. Most studies indicate the probability of cholecystitis in patients presenting in various settings with abdominal pain ranges from 1% to 8%[21,22,26,27,29,53,23,25]. In the elderly, however, the incidence surpasses that of

appendicitis and ranges from 9% to 41%[45,53,25]. Interestingly, the sensitivity and specificity of clinical findings has not been observed to decline as for appendicitis.

The clinical features of cholecystitis in the general population are outlined in Tables 9 and 10. Again, the historical elements are featured in the first[21,26], and the physical examination findings in the second table[21,26,53]. The first impression of these tables is how less well studied this disorder is when compared to appendicitis. Several of the classic signs of cholecystitis like Murphy's Sign, Charcot's triad, Reynold's Pentad have not been rigorously evaluated to support or refute their putative value.

Several points can be derived from the available data, however. Again, pain before vomiting (when vomiting present) is a sensitive indicator of serious disease. The remainder of the historical findings examined are less impressive. The physical examination has also been topically examined and reveals unimpressive result, summarized in Table 10. All of these studies were done with the general evaluation of abdominal pain as the goal. Until a specific study of cholecystitis is performed using strong methodology, the precision, accuracy, and usefulness of the clinical examination will remain unmeasured.

Table 9
Cholecystitis: Key Points of History

Sensitivity/Specificity

	Anorexia	Nausea	Vomiting	Pain a Vomiting
	Sn/Sp	**Sn/Sp**	**Sn/Sp**	**Sn/Sp**
Staniland	60/26	80/36	65/29	
Brewer	69/40		77/51	100/64

Cholecystitis: Key Points of History

Positive Likelihood Ratios

	Anorexia	Nausea	Vomiting	Pain a Vomiting
	LR+	**LR+**	**LR+**	**LR+**
Staniland	0.8	1.25	0.9	
Brewer	1.1		1.6	2.8

Table 10
Cholecystitis: Key Points of Exam

Sensitivity/Specificity

	RUQ Tender	Rebound	Rigidity	Fever	Elev. WBC	Sxs. Prev.
	Sn/Sp	**Sn/Sp**	**Sn/Sp**	**Sn/Sp**	**Sn/Sp**	**Sn/Sp**
Staniland	39/89	25/61	43/60			71/60
Brewer		35/82	58/80	31/84	77/64	12/71
Fenyo	91/		14/			

Positive Likelihood Ratio

	RUQ Tender	Rebound	Rigidity	Fever	Elev. WBC	Sxs. Prev.
Staniland	3.5	0.6	1.1			1.8
Brewer		1.9	2.9	1.9	2.1	0.4

Assessing Spleen Size

Anatomically, the spleen is located inside the posterior segment of the 10th rib, as depicted in Figure 5. It's narrow, posterior pole points toward the spine (Figure 6), and the blunt, anterior pole approaches the mid-axillary line. As the spleen enlarges, the anterior pole heads toward the right iliac fossa (Figure 7).

Figure 5

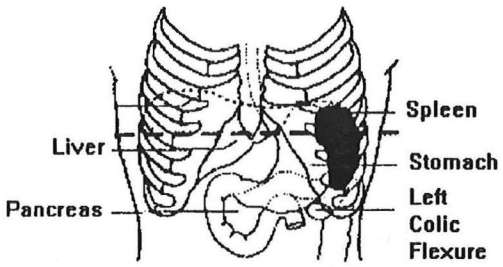
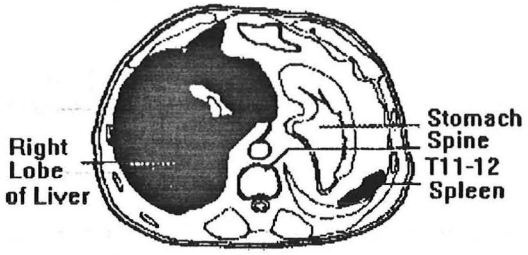
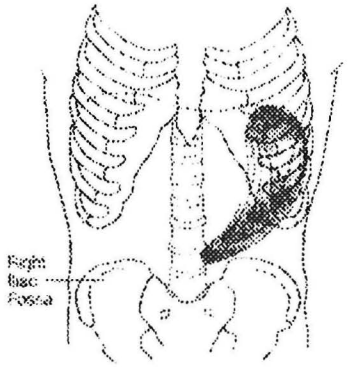


Figure 6



There are many techniques described in the literature for estimating spleen size. There are many more variations on these techniques, depending on site of training, clinical experience, and instructor. The most ingenious, but least practical of approaches was to suggest the patient "jump up and down for 20-25 times" and subsequently examining the patient from behind while both physician and patient remain standing[54].

Figure 7



Many of the major physical examination textbooks conflict on their recommendations for assessing spleen size[55,56,57]. Fortunately, this maneuver is the best studied of the physical examination. Halpern[58], Sullivan[59], Barkun[60], and Tamayo[61] are of the highest quality. A summary of these studies is provided in Tables 11 and 12.

Several points can be derived from the tables. First, most techniques evaluated have a high specificity, making a positive finding useful for ruling in splenomegaly (SpPin). The sensitivity is less impressive, making it more difficult for ruling out an enlarged spleen (SnNout). Another point of the tables is that a rigorous comparison of the techniques has been only attempted a few times, with some techniques performing better.

Note: the figures from this section were borrowed from Grover[62]

Table 11
Palpation for Splenomegaly

Author	Yr.	Technique	n	Sens.	Spec.	LR+*
Holzbach	'62	undefined	16	-	0.62	-
Halpern{34}	'74	undefined	214	0.28	0.99	28
Sullivan{38}	'76	undefined	65	0.71	0.90	7.1
Barkun{40}	'91	Supine/RLD	118	0.39	0.97	13
		Middleton's		0.39	0.97	13
Tamayo{39}	'93	Bimanual/ Supine/RLD	27	-	-	1.69†
		Ballottement		-	-	2.22†
		Middleton's		-	-	2.66†

*LR+ = Likelihood Ratio = Sn/1-Sp
† Confidence Intervals exclude 1.0

Three palpation techniques have been the most thoroughly evaluated in the literature, and found to be useful: Middleton's, Ballottement, and the Supine/Right Lateral Decubitus maneuvers. Two recent reviews have assessed these maneuvers and the studies that have evaluated them[63,62].

Middleton's Maneuver[64] has the most powerful likelihood ratio, and therefore the most effective profile. The maneuver is performed by having the patient lie on their left fist just inferior to their left scapula (Figure 5). The physician approaches the patient from their left shoulder, hooking his/her fingertips beneath the left costal margin as the patient inspires deeply.

Figure 5

Ballottement may be performed with a single hand or bimanually[57]. Tamayo's study uses bimanual ballottement, which was the second most powerful palpation technique. Performed from the patient's right side, the physician's left hand reaches over the patient to rest posteriorly on the patient's left flank. The physician's right hand is placed passively on the left hypochondriac region. The left hand then lifts the left flank toward the right hand, which should receive an impulse from an enlarged spleen (Figure 6).

Figure 6

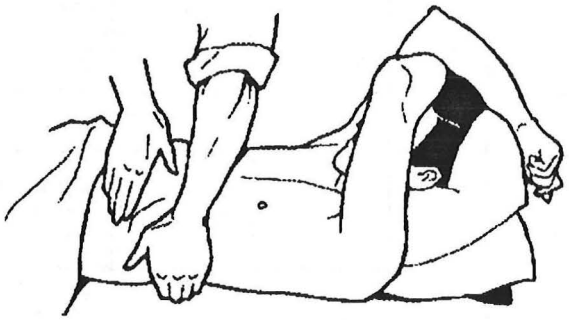


Figure 7

Tamayo studied another bimanual palpation performed from the patient's right side with the patient supine or in the right lateral decubitus position. This maneuver was initially described by **Hamilton Bailey**[1,65]. The physician's left hand is placed on the patient's lower left anterior rib cage, pulling the skin inferiorly. This serves to immobilize the rib cage to accentuate diaphragmatic movements as well as loosening the abdominal wall. This allows the fingertips of the right hand to sink under the costal margin and palpate for spleen tips upon deep inspiration (Figure 7).

Note: the figures in this section are from Yang[63]

In comparing the individual maneuvers, there is wide variation of results in the literature. Tamayo's study employed the most rigorous experimental design and analysis of the data. This was the only study that included an assessment of the study's power by including confidence intervals. This study should therefore be regarded as most representative of reality.

In summary, the best palpation technique appears to be Middleton's maneuver, although all techniques are of fairly equal efficacy. Grover cautions us that these maneuvers are of limited value when the prevalence of splenomegaly is low[62].

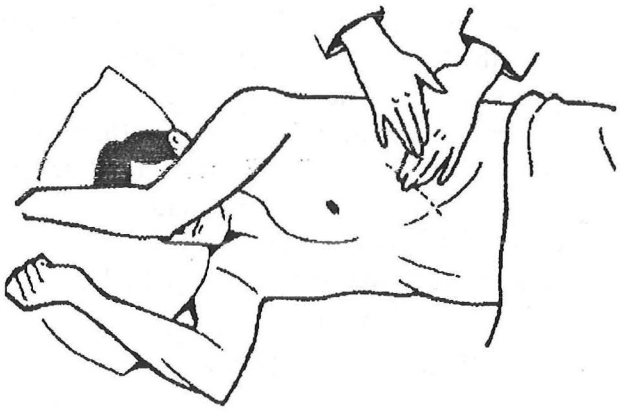
There are also three main variations for percussion to estimate the splenic size: the Nixon, Castell, and Barkun techniques. Sullivan[59], Barkun[60], and Tamayo[61] have compared these techniques, as described in Table 12.

Table 12
Percussion for Splenomegaly

Author	Yr.	Technique	n	Sens.	Spec.	LR+*
Sullivan	'76	Nixon	65	0.59	0.94	9.8
	'76	Castell	65	0.82	0.83	4.8
Barkun	'89	Barkun		0.62	0.72	2.2
Barkun	'91	Castell	118	0.79	0.46	1.5
Tamayo	'93	Nixon	27	-	-	1.74†
		Castell		-	-	1.97†
		Barkun		-	-	1.69

*LR+ = Likelihood Ratio = Sn/1-Sp
†95% confidence intervals exclude 1.0

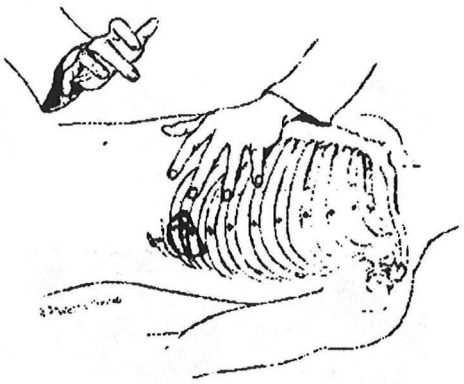
Figure 8



The **Nixon** technique[66] is depicted in Figure 8 and is described in the original article as follows: "The patient is placed in the right lateral recumbent position with the left arm extended forward and upward sufficiently to clear the left lower part of the thorax. In this position the spleen lies above both stomach and colon, permitting determination of its upper and lower borders of dullness. after palpation for the lower border on inspiration, percussion is initiated at the lower level of pulmonary resonance in approximately the

posterior axillary line and carried downward obliquely on a general perpendicular line toward the lower midanterior costal margin. Normally, the upper border of dullness is measured 6 to 8 cm. above the costal margin. Dullness increased over 8 cm. is indicative of splenic enlargement in the adult."

Figure 9



The **Castell** method, also called the "Spleen Percussion Sign"[67] is performed with the patient in the supine position. A percussion note is sounded in the lowest intercostal space (usually 8th or 9th) along the left anterior axillary line (Figure 9). Castell states that this percussion "usually produces a resonant note if the spleen is normal in size. Furthermore, the resonance persists with full inspiration. as the spleen enlarges, the lower pole of this organ is displaced inferiorly and medially. This may produce a change in the percussion note in the lowest left interspace in the anterior axillary line form resonance to dullness with full inspiration. The percussion sign is considered positive, therefore, when such a change is noted between full expiration and full inspiration."

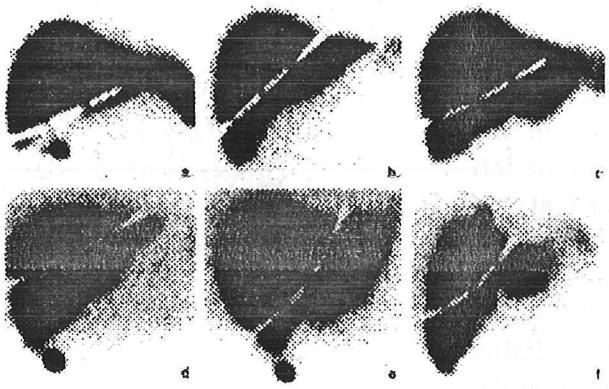
Barkun describes his method as percussion of Traube's Space[68]. Traube's space is defined as "an area delineated by the sixth rib superiorly, the midaxillary line laterally, and the left costal margin inferiorly". In this method, this area is percussed from its medial to lateral margin. Again, the percussive note is expected to be resonant throughout this course with a normal-sized spleen.

Tamayo again provides the most meaningful comparison of these findings, calculating likelihood ratios with confidence intervals. The Castell Method appears to be the most consistently powerful maneuver, although all methods are probably equally effective.

Barkun went on to compare combinations of findings, and discovered that palpation was "a better discriminator among patients with percussion dullness compared with those in whom Traube's space was tympanitic".

In summary, when assessing spleen size, the best approach is to begin with a single percussive note in the lowest intercostal space along the left anterior axillary line while the patient is supine. If this is not clearly resonant, the examiner should proceed with a palpation maneuver. The best maneuver is probably a gentle bimanual hooking under the patient's left costal margin while standing at the patient's left shoulder.

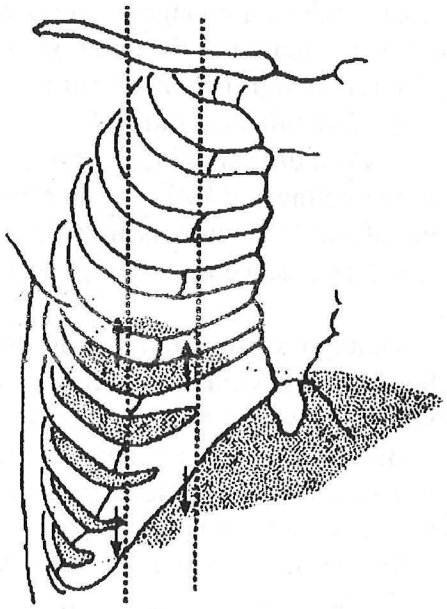
Figure 10



Assessing Liver Size

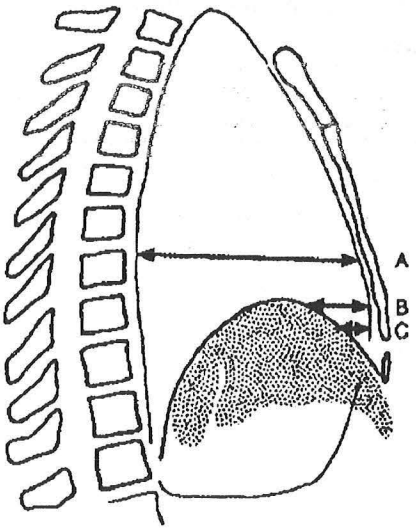
Liver morphology varies widely in normal individuals as depicted in Figure 10. An assessment of upper and lower borders is necessary for the most accurate assessment of size. Standardization of where these borders is necessary, as the liver span can vary widely when measured in different vertical planes (Figure 11). The upper edge is convex and surrounded by the thoracic ribs, so assessment of this border can be difficult (Figure 12).

Figure 11



As might be expected, the methods for assessing liver size are more variable than those for the spleen. Several reviews have recently summed up the exam of the liver[69,70]. In general, four different approaches are described in the physical examination textbooks: 1) Percussion of the upper and lower borders, 2) Palpation of the lower edge of the liver from the patient's right side with fingerpads or edge of hands **(poking)**, 3) Palpation of the lower edge from the patient's right shoulder by curling the fingertips under the costal margin **(hooking)**, and 4) Auscultory Scratching or Percussing to delineate the upper and lower borders of the liver.

Figure 12



Either palpation technique is always used in combination with the percussion techniques, as Naftalis[71] has demonstrated that simple palpation is inaccurate. This inaccuracy arises

from normal-sized livers appearing to be enlarged due to large lung volumes, or large livers pushing the diaphragm into the thorax more than extending below the costal margin. The point is an assessment of both the upper and lower borders of the liver is crucial to the most accurate clinical assessment of liver size.

The Scratch Test. This maneuver is infrequently described in textbooks, and the literature conflicts when it is described. In general, the diaphragm of the stethoscope is placed either over the Xiphoid process or just superior to the costal margin along the mid-clavicular line. The examiner then gently scratches the skin along the right mid-clavicular line, starting in the lower abdomen and advancing toward the head. The sound produced by the scratching changes in quality and intensity when over the liver, as sounds are much more easily transmitted through the solid organ.

Despite the more protean approach to the assessment of liver size, the literature comparing the methods is less developed than for the spleen. The techniques are inadequately defined and/or standardized in the studies, the experimental design is of variable quality, and the data analysis is either not described or inappropriate. For these reasons, there is not much to report from an Evidence-Based Medicine standpoint. A summary of the best studies is provided in Table 113.

Table 13
Liver Span for Hepatomegaly

Author	Technique	n	Acc.†	Sn.	Sp.	LR+§
Naftalis{149}	Perc./Palp. undefined	39	0.97	-	-	-
Riemen-schneider{32}	Perc./Palp. undefined	47	-	0.50	0.77	2.17
Peternel{31}	Perc./Palp. undefined	43	0.42	0.64	0.75	2.56
Blendis{33}	Perc./Palp. undefined	32	0.48	0.50	0.47	0.94
Halpern{34}	undefined	214	0.66	0.71	0.62	1.87
Sullivan{29}	Scratch	50	0.42	-	-	-
	Perc./Palp. undefined		0.40	-	-	-
Fuller{27}	Scratch	24	0.78	-	-	-
	Perc./Palp. undefined		0.44	-	-	-

† Accuracy defined as within 2 cm of standard used (e.g. US, Scintillation Scan, Autopsy)
§ LR+ = Likelihood Ratio = Sn/1-Sp

One consistent finding across studies not reflected in Table 3 is percussion's underestimation of liver size. Presumably this is due to the insensitivity of percussion to the height of the dome of the liver that projects into the thorax. This fact is reflected in the mediocre sensitivities across studies.

In summary, the precision and accuracy of the assessment of liver size has not been adequately evaluated. Some studies even imply that the physical examination is of little worth in this regard (with a LR < 1.0). The best recommendation at present is for a re-assessment of the scratch test, for which a trend for more accurate diagnosis of hepatomegaly may exist.

Limitations of the Data

Although these studies are the best attempt available at scientific evaluation and validation of prior anecdotal doctrines, the data and analyses in these studies is far from perfect. If strong recommendations are to be made to the physician evaluating the individual patient or to society when making Health Care Policy, much more detail and planning is necessary. Several recommendation can be proposed, based on the available data and the current state of the art of clinical research.

The first recommendation is for more emphasis on proper study design. For example, most studies reviewed in this grand rounds were retrospective chart reviews. Prospective study design would eliminate much bias introduced. In addition, study design should take into account the proper population to be studied. Many of the studies cited were based on the population of patients known to have a disorder via radiologic or operative findings. More studies should be based in the emergency department or the clinic. Proper study design would also collect and report the patient demographics and other pertinent characteristics, then control for possible confounding by these variables. In summary, the best design for the evaluation of diagnostic procedures would be a prospective controlled cross-sectional study.

The second recommendation is for more sophisticated data analysis. All studies reviewed reported the univariate analysis of the clinical findings. For example, the positive likelihood ratio for Ballottement is 2.22 and for Middleton's maneuver is 2.66 when assessing spleen size. Are these ratios additive, that is, if a patient has positive Ballottement and Middleton's maneuver, is their positive likelihood ratio 4.88? or higher? Multivariate analyses using regression techniques or recursive partitioning could easily answer these questions.

The third recommendation is for a standardized examination in studies. The spleen studies were most conscientious in describing the technique used, and had the most impressive outcomes. Most of the other studies relied on chart reviews to derive the findings of the exam, so we can assume the technique was of heterogeneous description and quality. Inter- and intra-observer reliability is also of major concern in these type of studies, and are important to report along with the results.

Summary

The Evaluation of Abdominal Pain:

For Non-Specific Abdominal Pain:

Wasson's Rule has a high specificity, SpPins in a diagnosis of non-serious disease
Remember the Abdominal Wall Tenderness Test, especially in chronic pain

For Appendicitis:

On History: Pain < 48 hours, pain preceding vomiting (if vomiting is present), and the absence of similar pain previously are the most sensitive elements. The classic history of pain migration is the most specific element.

On Exam: right lower quadrant tenderness and leukocytosis are the most sensitive elements. Rovsing's, the Psoas and Obturator signs are the most specific elements.

In a Woman of Childbearing Age, factors that favor a diagnosis of appendicitis over PID are:

1) the presence of nausea, vomiting, or both,
2) the absence of a history of venereal disease,
3) isolated right lower quadrant pain,
4) cervical motion tenderness on exam, and
5) bilateral adnexal tenderness on exam.

Remember that although the probability of appendicitis is less in the elderly, the signs and symptoms are manifested less frequently, and the morbidity and mortality is greater.

For Cholecystitis:

On History: pain before vomiting is again a good predictor, being the most sensitive of the elements.

On Exam: right upper quadrant tenderness is the most specific element.

When Evaluating Spleen Size:

The best approach is to screen with a Castell's Maneuver. If the note is not clearly resonant, the examiner should proceed with a Middleton's maneuver.

When Evaluating Liver Size:

The current literature advocates palpation/percussion from the patient's right, hooking/percussing, and the scratch test equally. There is a trend for the scratch test to be superior.

Conclusion

This grand rounds has documented the "state of the science" for the clinical examination. Although it is useful to know which maneuvers contribute the most to the clinical assessment, individually, each maneuver has only modest operating characteristics. No single finding can conclusively rule in or out disease. The skeptic would claim this adds to the evidence that the physical examination is a waste of time and physician's time is better spent analyzing radiological and laboratory data. Even the harshest skeptic, however, could not deny that a major therapeutic intervention could not be prescribed without first "seeing" the patient. There is no substitution for laying eyes on the patient to assess how "sick" they appear. Other subtle clues are gleaned from the doorway, across the room, or during history taking that also count as "physical examination".

The literature thus far has not yet been able to capture the value of these clinical clues. As mentioned previously, several studies have documented that the value of the physical examination in general is at least equal to the contribution of laboratory and radiological evaluation[19,18] The source of this general value of the physical examination has as yet been undiscovered, as only individual findings have been evaluated.

It is clear that much more needs to be done in determining the most and least useful components, most powerful combination of findings, and teaching them in a standardized manner. Although most of the clinical exam remains "experience based", the artists at the beside must strive for a healthy infusion of science. The beneficiaries of such an infusion would be the patient, the physician, and society in general.

Finally, this grand rounds is not meant to decry the "artistic" approach to the physical examination. Many intangibles are the result of the laying on of hands. They help establish the rapport of the physician-patient relationship, at which its contribution to therapy can only be speculated.

References

1. Clain Allan. Hamilton Bailey's Demonstrations of Physical Signs in Clinical Surgery. Bristol, England: IOP Publishing Limited, Techno House, 1986: 1-622.

2. Silen W. Cope's Early Diagnosis of the Acute Abdomen. New York: Oxford University Press, Inc., 1987: 1-290.

3. Feinstein AR. Clinical Judgement. Huntington, N.Y.: Robert E. Krieger Publishing Co., 1967: 1-414.

4. Fitzgerald FT. Physical Diagnosis Versus Modern Technology. A Review. West J Med 1990; 152:377-382.

5. Nardone DA, Lucas LM,Palac DM. Physical Examination: A Revered Skill Under Scrutiny. Southern Med J 1988; 81(6):770-773.

6. Sackett KL, Haynes RB, Guyatt GH, Tugwell P. Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston, Mass: Little Brown & Co Inc., 1991: 173-186.

7. Weinstein MC, Fineberg HV, Elstein AS, et al. Clinical Decision Analysis. Philadelphia,Penn.: W.B.Saunders Co, 1980: 1-351.

8. Evidence-Based Medicine Working Group. Evidence-Based Medicine. A New Approach to teaching the Practice of Medicine. JAMA 1992; 268(17):2420-2425.

9. Sackett DL. A Primer on the Precision and Accuracy of the Clinical Examination. JAMA 1992; 267(19):2638-2644.

10. Berry J,Malt RA. Appendicitis Near its Centenary. Ann Surg 1984; 200(5):567-575.

11. Neutra RR. Appendicitis: Decreasing Normal Removals Without Increasing Perforations. Med Care 1978; XVI(11):956-961.

12. White JJ, Santillana M,Haller JA. Intensive In-Hospital Observation: a Safe Way to Decrease Unnecessary Appendectory. Am Surgeon 1975; 41:793-798.

13. StClair EN, Oddone EZ, Waugh RA,et al. Assessing Housestaff Diagnostic Skills. Ann Intern Med 1992; 117:751-756.

14. Johnson JE,Carpenter JL. Medical House Staff Performance in Physical Examination. Arch Intern Med 1986; 146:937-941.

15. Koran LM. The Reliability of Clnical Methods, Data, and Judgements (Part 1). N Engl J Med 1975; 293:642-646.

16. Wiener S,Nathanson M. Physical Examination. frequently Observed Errors. JAMA 1976; 236(7):852-855.

17. Bayes T. A Problem in the Doctrine of Change. Philosophic Trans R Soc 1763; 53:370-418.

18. Peterson MC, Holbrook JH, Hales DV, Smith NL,Staker LV. Contributions of the History, Physical Examination, and Laboratory Investigation in Making Medical Diagnoses. West J Med 1992; 156:163-165.

19. Hampton JR, Harrison MJG, Mitchell JRA, Prichard JS,Seymour C. Relative Contributions of History-taking, Physical Examination, and Laboratory Investigation to Diagnosis and Management of Medical Outpatients. Br Med J 1975; 2:486-489.

20. Coulehan JL, Block MR. The Medical Interview: A Primer for Students of the Art. Philadelphia,PA: F.A.Davis Company, 1987: 1-221.

21. Staniland JR, Ditchburn J,De Dombal FT. Clinical Presentation of Acute Abdomen: Study of 600 Patients. Br Med J 1972; 3:393-398.

22. De Dombal FT, Leaper DJ, Horrocks JC, Staniland JR,McCann AP. Human and Computer-aided Diagnosis of Abdominal Pain: Further Report with Emphasis on Performance of Clinicians. Br Med J 1974; 1:376-380.

23. Jess P, Bjerregaard B, Brynitz S, et al. Prognosis of Acute Nonspecific Abdominal Pain. A Prospective Study. Am J Surg 1982; 144:338-340.

24. Thomson HJ,Jones PF. Active Observation in Acute Abdominal Pain. American Journal of Surgery 1986; 152:522-525.

25. Irvin TT. Abdominal Pain: a Surgical Audit of 1190 Emergency Admissions. Br J Surg 1989; 76:1121-1125.

26. Brewer RJ, Golden GT, Hitch DC, Rudolf LE,Wangensteen SL. Abdominal Pain. An Analysis of 1000 consecutive Cass in a University Hospital Emergency Room. Am J Surg 1976; 131:219-223.

27. Wilson DH, Wilson PD, Walmsley RG, Horrocks JC,De Dombal FT. Diagnosis of Acute Abdominal Pain in the Accident and Emergency Department. Br J Surg 1977; 64:250-254.

28. Simmen HP, Decurtins M, Rotzer A, Duff C, Brutsch HP,Largiader F. Emergency Room Patients with Abdominal Pain Unrelated to Trauma: Prospective Analysis in a Surgical University Hospital. Hepato-Gastroenterol 1991; 38:279-282.

29. Wasson JH, Sox Jr HC,Sox CH. The diagnosis of Abdominal Pain in Ambulatory Male Patients. Med Decision Making 1981; 1(3):215-224.

30. Orient JM, Kettel LJ,Lim J. A Test of a Linear Discriminant for Identifying Low-Risk Abdominal Pain. Med Decision Making 1985; 5(1):77-87.

31. Britt H, Bidges-Webb C, Sayer GP, Neary S, Traynor V,Charles J. The Diagnostic Difficulties of Abdominal Pain. Australian Fam Physician 1994; 23(3):375-381.

32. Lukens TW, Emerman C,Effron D. The Natural History and Clincal Findings in Undifferentiated Abdominal Pain. Ann Emerg Med 1993; 22(4):690-696.

33. Hall PN,Lee APB. Rectus Nerve Entrapment Causing Abdominal Pain. Br J Surg 1988; 75:917-917.

34. Greenbaum DS, Greenbaum RB, Joseph JG,Natale JE. Chronic Abdominal Wall Pain. Diagnostic Validity and Costs. Digest Dis Sciences 1995;

35. Norfleet RG. More Causes of Abdominal Wall Pain. J Clin Gastroenterol 1993; 16(1):82-82.

36. Thomson H,Francis DMA. Abdominal-Wall Tenderness: A Useful Sign in the Acute Abdomen. Lancet 1977; 1053-1054.

37. Hershfield NB. The Abdominal Wall. A Frequently Overlooked Source of Abdominal Pain. J Clin Gastroenterol 1992; 14(3):199-202.

38. Carnett JB. Intercostal Neuralgia as a Cause of Abdominal Pain and Tenderness. Surg Gynecol Obstetr 1926; 42:625-632.

39. Anonymous. Abdominal Wall Tenderness Test: could Carnett Cut Costs?. Lancet 1991; 337:1134-1134.

40. Fitz RH. Perforating Inflammation of the Vermiform Appendix: with Special Reference to its Early Diagnosis and Treatment. Am J Med Sci 1886; 92:321-346.

41. Alvarado A. A Practical Score for the Early Diagnosis of Acute Appendicitis. Ann Emerg Med 1986; 15(5):557-564.

42. Izbicki JR, Wolfram TK, Dietmar KW, et al. Accurate Diagnosis of Acute Appendicitis: a Retrospective and Prospective Analysis of 686 Patients. Eur J Surg 1992; 158:227-231.

43. Nauta RJ,Magnant C. Observation Versus Operation for Abdominal Pain in the Right Lower Quadrant. Roles of the Clinical Examination and the Leukocyte Count. American Journal of Surgery 1986; 151:746-748.

44. Nase HW, Kovalcik PJ,Cross GH. The Diagnosis of Appendicitis. Am Surgeon 1980; 46:504-507.

45. Fenyo G. Diagnostic Problems of Acute Abdominal Diseases in the Aged. Acta Chir Scand 1974; 140:396-405.

46. Bongard F, Landers DV,Lewis F. Differential Diagnosis of appendicitis and Pelvic Inflamatory Disease. American Journal of Surgery 1985; 150:90-95.

47. Liddington MI,Thomson WHF. Rebound Tenderness Test. Br J Surg 1991; 78:795-796.

48. Gilmore OJA, Browett JP, Griffin PH, et al. Appendicitis and Mimicking Conditions. A Prospective Study. Lancet 1975; 7932:421-423.

49. Jess P, Bjerregaard B, Brynitz S, Holst-Christensen J, Kalaja E,Lund-Kristensen J. Acute Appendicitis. Prospective Trial Concerning Diagnostic Accuracy and Complications. American Journal of Surgery 1981; 141:232-234.

50. Lewis FR, Holcroft JW, Boey J,Dunphy E. Appendicitis. A Critical Review of Diagnosis and Treatment in 1000 Cases. Arch Surg 1975; 110:677-684.

51. Teicher I, Landa B, Cohen M, Kabnick LS,Wise L. Scoring system to Aid in the Diagnosis of Appendicitis. Ann Surg 1983; 198(6):753-759.

52. Webster DP, Schneider CN, Cheche S, Daar AA,Miller G. Differentiating Acute Appendicitis from Pelvic Inflammatory Disease in Women of Childbearing Age. Am J Emerg Med 1993; 11:569-572.

53. Fenyo G. Acute Abdominal Disease in the Elderly. Am J Surg 1982; 143:751-754.

54. Breman JG. Spleen Palpation Aided by Gravity. Lancet 1973;

55. Bates Barbara. A Guide to Physical Examination and History Taking. Philadelphia:

J. B. Lippincott Company, 1987: 1-661.

56. Willms JL, Schneiderman H, Algranati PS. Physical Diagnosis. Bedside Evaluation of Diagnosis and Function. Baltimore: Williams & Wilkins, 1994: 1-804.

57. Seidel HM, Ball JW, Dains JE, Benedict GW. Mosby's Guide to Physical Examination. St. Louis,Missouri: Mosby-Year Book,Inc., 1995: 1-918.

58. Halpern S, Coel M, Ashburn W, et al. Correlation of Liver and Spleen Size. Determinations by Nuclear Medicine Studies and Physical Examination. Arch Intern Med 1974; 134:123-124.

59. Sullivan S,Williams R. Reliability of Clinical Techniques for Detecting Splenic Enlargement. Br Med J 1976; 2:1043-1044.

60. Barkun AN, Camus M, Green L, et al. The Bedside Assesment of Splenic Enlargement. Am J Med 1991; 91:-19203.

61. Tamayo SG, Richman LS, Mathews WC, et al. Examiner Dependence on Physical Diagnostic Tests for the Detection of Splenomegaly: A Prospective Study with Multiple Observers. J Gen Intern Med 1994; 8:69-75.

62. Grover SA, Barkun AN,Sackett DL. Does This Patient Have Splenomegaly?. JAMA 1993; 270(18):2218-2221.

63. Yang JC, Rickman LS,Bosser SK. The Clinical Diagnosis of Splenomegaly. West J Med 1991; 155:47-52.

64. Lipp WF, Eckstein EH,Aaron AH. The Clinical Significance of the Palpable Spleen. Gastroenterol 1944; 3:287-291.

65. Sapira Joseph D. The Art and Science of Bedside Diagnosis. Baltimore: Williams & Wilkins, 1990: 1-601.

66. Nixon RK. The Detetion of Splenomegaly by Percussion. N Engl J Med 1954; 250(4):166-167.

67. Castell DO. The Spleen Percussion Sign. A Useful Diagnostic Technique. Ann Intern Med 1967; 67(6):1265-1267.

68. Barkun AN, Camus M,Meagher T. Splenic Enlargement and Traube's Space: How Useful is Percussion?. Am J Med 1989; 87:562-566.

69. Naylor CD. Physical Examination of the Liver. JAMA 1994; 271(23):1859-1865.

70. Meidl EJ,Ende J. Evaluation of Liver Size by Physical Examination. J Gen Intern Med 1993; 8:635-637.

71. Naftalis J,Leevy CM. Clinical Estimation of Liver Size. Am J Dig Dis 1963; 8(3):236-243.