Clinical Manifestations of Disease

Andrew Jull, Alba DiCenso, and Gordon Guyatt

The following Editorial Board members also made substantive contributions to this chapter: Rien de Vos, Teresa Icart Isern, and Mary van Soeren.

We gratefully acknowledge the work of Scott Richardson, Mark Wilson, John Williams, Virginia Moyer, and David Naylor on the original chapter that appears in the Users’ Guides to the Medical Literature, edited by Guyatt and Rennie.

In This Chapter

Are the Results Valid?
- Were the Right Patients Enrolled? Was the Patient Sample Representative of Those With the Disorder?
- Was the Definitive Diagnostic Standard Appropriate? Was the Diagnosis Verified Using Credible Criteria That Were Independent of the Clinical Manifestations Under Study?
- Were Clinical Manifestations Sought Thoroughly, Carefully, and Consistently?
- Were Clinical Manifestations Classified by When and How They Occurred?

What Are the Results?
- How Frequently Did the Clinical Manifestations of Disease Occur?
- How Precise Were These Estimates of Frequency?
- When and How Did These Clinical Manifestations Occur in the Course of Disease?

How Can I Apply the Results to Patient Care?
- Are the Study Patients Similar to Those in My Clinical Setting?
- Is It Unlikely That the Disease Manifestations Have Changed Since This Evidence Was Gathered?
- How Can I Use the Results in Generating a Differential Diagnosis?
Two types of systematic investigations can inform the process of generating a differential diagnosis. One type of study addresses the presenting manifestations of a disease or condition. The second, and more important, type of study directly addresses the underlying causes of a presenting symptom, sign, or constellation of symptoms and signs. This chapter focuses on how to critique studies that address the former. Chapter 21 focuses on how to critique studies that address the latter. Table 20-1 summarizes the criteria for interpreting an article about the clinical manifestations of disease.

**ARE THE RESULTS VALID?**

**Were the Right Patients Enrolled? Was the Patient Sample Representative of Those With the Disorder?**

Ideally, a study sample will mirror the population of patients with the target condition so that the frequency of clinical manifestations in the sample approximates that in the
underlying population. Such a patient sample is termed representative. The more representative a sample is, the more accurate the resulting frequencies of clinical findings will be.²

To judge the representativeness of a study sample, we suggest three tactics. First, examine the study setting. Patients seen in referral care settings may have higher proportions of unusual findings or illnesses that are harder to diagnose and thus will have different frequencies of clinical manifestations than patients diagnosed in community practice.³ Second, examine the methods used to identify and include study patients and to exclude others. Ask yourself if they included all important demographic groups (e.g., those characterized by age, sex, and race) or excluded important subgroups. Third, examine the description of study patients’ illnesses. Are patients with mild, moderate, and severe symptoms included? If different clinical patterns of disease are known, does the sample include patients with each pattern?

Combining these three considerations, you can judge whether the spectrum of included patients is broad enough for the study to yield valid results about the clinical manifestations of the disease. For instance, a study of the clinical findings in patients with thyrotoxic periodic paralysis included 19 patients who were hospitalized during an episode of paralysis and excluded 11 patients who were diagnosed during the study period but were not admitted.⁴ To the extent that clinical manifestations differed in hospitalized and nonhospitalized patients, such a restriction could introduce bias into the study.

Investigators may deliberately choose to describe the manifestations of a disease in a purposefully narrow target population defined by demographics (e.g., a study of the

<table>
<thead>
<tr>
<th>Are the Results Valid?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Were the right patients enrolled? Was the patient sample representative of those with the disorder?</td>
</tr>
<tr>
<td>• Was the definitive diagnostic standard appropriate? Was the diagnosis verified using credible criteria that were independent of the clinical manifestations under study?</td>
</tr>
<tr>
<td>• Were clinical manifestations sought thoroughly, carefully, and consistently?</td>
</tr>
<tr>
<td>• Were clinical manifestations classified by when and how they occurred?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What Are the Results?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• How frequently did the clinical manifestations of disease occur?</td>
</tr>
<tr>
<td>• How precise were these estimates of frequency?</td>
</tr>
<tr>
<td>• When and how did these clinical manifestations occur in the course of disease?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How Can I Apply the Results to Patient Care?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Are the study patients similar to those in my clinical setting?</td>
</tr>
<tr>
<td>• Is it unlikely that the disease manifestations have changed since this evidence was gathered?</td>
</tr>
<tr>
<td>• How can I use the results in generating a differential diagnosis?</td>
</tr>
</tbody>
</table>
findings of myocardial infarction in elderly patients\textsuperscript{5}, prognosis (e.g., a study of the clinical findings before death in patients with fatal pulmonary embolism\textsuperscript{6}), or site of care (e.g., a study of the findings in patients with ruptured abdominal aortic aneurysm who present to internists in their offices, rather than to emergency departments\textsuperscript{7}). In such situations, you can assess whether the study sample is representative of the limited target population.

**Was the Definitive Diagnostic Standard Appropriate? Was the Diagnosis Verified Using Credible Criteria That Were Independent of the Clinical Manifestations Under Study?**

These questions address two closely linked issues. First, how sure are the investigators that patients in the study really did have this particular disease, rather than another disease? Clinicians often encounter patients with tentative diagnoses. However, in studies of disease manifestations, such diagnostic uncertainty could introduce bias because the sample could include patients with other diseases. To minimize this threat, investigators can use a set of explicit diagnostic criteria and include only patients who meet these criteria. Ideally, for every disease, there would be a set of published, widely accepted, diagnostic criteria, including one or more well-established reference, gold, or criterion standard tests that could be applied in a reproducible way. Reference standards can be anatomic, physiologic, radiographic, or genetic, to name a few. To judge how the presence of disease was verified, identify the standards that were used for disease verification, how they were used, and whether the standards were clinically credible.

When no reference standards exist, the degree of diagnostic certainty is much lower. In such situations, known sometimes as syndrome diagnosis\textsuperscript{8}, diagnostic criteria usually rely on a list of clinical features required for diagnosis. For example, the definition of chronic fatigue syndrome uses an explicit set of clinical features as diagnostic criteria.\textsuperscript{9} Such explicit criteria represent an advance over an implicit, haphazard approach.

However, problems arise when investigators use these clinical manifestations to make the syndrome diagnosis, select the patient sample, and then examine the frequency of these same clinical findings in the study patients. This situation creates a form of circular reasoning that can bias upward the frequencies of these findings in the study sample. For example, a study of the clinical features of 36 patients with relapsing polychondritis had this incorporation bias, because the investigators used diagnostic criteria that were based primarily on characteristic clinical findings.\textsuperscript{10} Although this may be the best available method for clinical diagnosis, incorporation bias limits the inferences we can draw about the frequency of clinical manifestations. To judge the independence of verifying criteria, compare the list of these criteria with the list of clinical manifestations studied.

**Were Clinical Manifestations Sought Thoroughly, Carefully, and Consistently?**

This criterion addresses three closely related issues. First, were study patients evaluated thoroughly enough to detect clinical findings if they were present? Within reason, the more comprehensive the workup is, the lower the chance of missing findings and drawing invalid conclusions about their frequency. Second, how did the investigators ensure that
the information they collected was correct and free of distortion? Were patients asked about symptoms in neutral, nonjudgmental ways, or were leading questions asked that could have suggested symptoms? Were patients examined by skilled examiners? The more carefully data are gathered, the more credible the resulting frequencies will be. Third, how consistently was the evaluation performed? Variable assessments could yield erroneous frequencies of disease manifestations.

You may find it relatively easy to judge the thoroughness, care, and consistency of the search for manifestations if clinicians evaluated patients prospectively using a standardized diagnostic approach. It becomes harder to judge when investigators use information previously collected by unstandardized methods. For example, a retrospective analysis of disease manifestations in 68 patients with lumbar spinal stenosis did not include a description of the search for clinical findings in sufficient detail to judge how well the investigators protected against biased ascertainment. Ordinarily, a prospective study of clinical manifestations of disease provides more credible results than a retrospective study.

**Were Clinical Manifestations Classified by When and How They Occurred?**

Clinical manifestations of disease can range from permanent to fleeting. They can occur early, late, or throughout the course of disease. Investigators would obtain the most complete information about the timing of disease manifestations if they could begin collecting data the instant the disease begins and continue to the end of the illness. Because knowing this “zero time” with certainty is impossible for most diseases, investigators can use the next strongest approach of targeting all findings that occur from the onset of first symptoms of an illness episode. Studies that do not start collecting information at the beginning of the episode or do not report the timing of evaluation relative to symptom onset may miss transient findings, and our confidence in their validity decreases. For instance, in a study of the clinical manifestations before death in 92 patients with fatal pulmonary embolism, investigators recorded findings for only the final 24 hours before death and thus may have missed transient but important diagnostic clues that occurred before that time.

Sometimes, studies describe qualitative findings that may be useful in clinical diagnosis, particularly in triggering initial diagnostic hypotheses. For instance, patients often describe the pain of aortic dissection as a “tearing” or “ripping” sensation located in the center of the torso and reaching maximal intensity quite quickly. Just as with the temporal aspects, these qualitative descriptions are more credible if practitioners gathered them deliberately and carefully.

**Using the Guide**

Spittell and colleagues studied patients from the Mayo Clinic in Rochester, Minnesota, which provides both community hospital care and tertiary referral care. The sample included 158 patients (67%) with acute aortic dissection (duration less than 2 weeks) and 78 patients (33%) with chronic aortic dissection (duration at least 2 weeks). In 60 patients, the initial clinical impression was a diagnosis other...
Using the Guide—cont’d

than aortic dissection. The sample included patients who had died suddenly, including 10 patients with out-of-hospital cardiac arrests and five deaths that occurred in hospital. It also included 11 patients without pain but with other symptoms and 33 patients without pain or other symptoms, who had abnormal chest radiographs. Thus, the sample displays a wide array of clinical presentations likely to be representative of the full spectrum of this disorder.

The investigators studied 235 patients who had 236 aortic dissections confirmed by surgical intervention (in 162), autopsy (in 27), or radiographic studies (in 47). They excluded patients with aortic dissection that occurred intraoperatively or during invasive catheterization procedures. Thus, the diagnoses of study patients appear to have been verified using clinically credible methods that were independent of the clinical manifestations.

Spittell and colleagues1 retrospectively reviewed patient charts after clinical evaluations were completed. They did not explicitly describe the diagnostic evaluation. The tables of results include detailed information about the clinical examination, which suggests a careful approach. Uncertainty remains, however, about the extent of standardization during the workup.

The investigators described the clinical manifestations of aortic dissection at presentation for patients with acute (duration less than 2 weeks) and chronic (duration at least 2 weeks) illness from aortic dissection. They also described the location of pain in relation to the site of dissection and clustering of pain with other findings, along with unusual findings, such as hoarseness and dysphagia. Thus, despite the retrospective study design, the investigators classified the temporal and qualitative features sufficiently to provide valid results for patients with acute aortic dissection. We may be less confident of the results for chronic dissection because early transient findings may not have been detected.

WHAT ARE THE RESULTS?

How Frequently Did the Clinical Manifestations of Disease Occur?

Studies of clinical manifestations of disease often display the main results in a table listing the clinical findings and the number and percentages of patients with each of those manifestations. Because patients usually have more than one finding, these proportions are not mutually exclusive. Some studies also report the number of patients with any of the findings, either as a total or separately by particular group.

Textbook descriptions can emphasize the presence of particular classic findings that are proved uncommon by systematic study. If clinicians rely on such findings, they will miss many cases. For example, although experts previously considered hemoptysis to be a hallmark of acute pulmonary embolism, only 30% of 327 patients with angiographically proved pulmonary emboli had hemoptysis.13 Thus, it would be unwise to use the absence of hemoptysis to exclude a diagnosis of pulmonary embolism.

Systematic studies of disease manifestations may also prove some findings to be more common than expected. For instance, murmur of aortic regurgitation was detected in
40 (32%) of 124 patients with confirmed aortic dissection, a finding suggesting that clinicians should purposefully seek this finding in suspected cases.12

How Precise Were These Estimates of Frequency?

Even when valid, frequencies of clinical manifestations are only estimates of the true frequencies. You can determine the precision of these estimates by examining their confidence intervals (CIs). If the authors do not provide CIs, you can calculate them using the following formula (for 95% CI):

\[
95\% \ CI = P \pm 1.96 \times \sqrt{\frac{P(1-P)}{n}}
\]

where \(P\) is the proportion of patients with the finding of interest and \(n\) is the number of patients in the sample. This formula becomes inaccurate when the number of cases is five or fewer, so approximations have been developed for this situation.14,15

For example, consider the clinical finding of pulse deficit found in 14 (6%) of 217 patients in the study by Spittell and colleagues.1 Using the formula, we would start with \(P = 0.06\) (14/217), \((1 - P) = 0.94\), and \(n = 217\). Working through the arithmetic, we find the CI to be 0.06 ± 0.03. Thus, although the most likely frequency of pulse deficit is 6%, it may range from 3% (i.e., 6% − 3%) to 9% (i.e., 6% + 3%).

Whether you consider the CI sufficiently precise depends on how you expect to use the information. For example, if a finding occurs in 50% of cases, you may plan to look for it on examination but not to use the presence or absence of this finding to exclude the diagnosis. If the CI for this estimate ranged from 30% to 70%, it would not change your expected use of the information, so the result may be precise enough. Conversely, if a finding occurs in 97% of patients, you may hope to use its absence to help you rule out the diagnosis. If the CI for this estimate ranged from 60% to 100% (the same 40-point range as before), using this finding to exclude the diagnosis could lead you to miss up to 40% of patients with the disorder. Such a result would be too imprecise to be used to rule out the disorder of interest.

When and How Did These Clinical Manifestations Occur in the Course of Disease?

Some studies report the temporal sequence of symptoms in sufficient detail to characterize symptoms as presenting (i.e., the symptoms prompted patients to seek care), concurring (i.e., the symptoms did not prompt patients to seek care but were present initially), or eventual (i.e., the symptoms were not present initially but were found subsequently). For example, in 100 patients with pancreatic cancer, investigators described weight loss and abdominal pain as presenting manifestations in 75 and 72 patients, respectively, whereas jaundice, commonly taught as a key presenting sign, was found in only 24 patients.16 In addition to reporting the chronology of events, such studies can also describe the location, quality, intensity, situational context, aggravating and alleviating factors, and associated findings for important features of the disorder.
Spittell and colleagues\(^1\) reported that 168 (74%) patients initially had acute onset of severe pain, 33 (15%) patients were asymptomatic but had abnormal chest radiographs, and 15 (6%) patients experienced cardiac arrest or sudden death. Of the 217 of 235 (92%) patients with a record of the cardiac examination, murmurs of aortic regurgitation were detected in 22 (10%). Pulse deficits were uncommon, occurring in 14 (6%) patients.

The investigators did not provide the 95% CIs for the probabilities they found. However, as illustrated, if you are concerned about how close the probabilities are to your thresholds, you can calculate the 95% CIs.

Spittell and colleagues\(^1\) described in detail the symptoms at initial assessment, both as individual findings and in clusters. They also described the location of pain and its association with the site of aortic dissection. They did not describe delayed manifestations in as much detail.

**HOW CAN I APPLY THE RESULTS TO PATIENT CARE?**

**Are the Study Patients Similar to Those in My Clinical Setting?**

The closer the match is between patients in your clinical setting and those in the study, the more confident you can be in applying the results. We suggest that you ask yourself whether the setting or the patients are so different from yours that you cannot use the results.\(^17\) You could consider whether patients in your clinical setting come from a geographic, demographic, cultural, socioeconomic, or clinical group that could be expected to differ substantively in the ways in which the disorder is expressed. For instance, the presenting symptoms of acute myocardial infarction were found to differ with advancing patient age. A study of 777 elderly hospitalized patients with myocardial infarction found that syncope, stroke, and acute confusion were common and sometimes the sole presenting symptoms.\(^5\)

**Is It Unlikely That the Disease Manifestations Have Changed Since This Evidence Was Gathered?**

As time passes, evidence about the clinical manifestations of disease can become obsolete. New diseases can emerge, and old diseases can present in new ways. New disease taxonomies can be built, thus changing the distinctions between disease states. Such events can so alter the clinical manifestations of disease that previously valid studies may no longer be applicable to current practice. For example, the arrival of acquired immunodeficiency syndrome dramatically changed our concept of pneumonia caused by *Pneumocystis carinii*, a fungus that causes disease most commonly when cell-mediated immunity is depressed.\(^18,19\)

Similar changes can occur as the result of progress in health science or medical practice. For instance, early descriptions of *Clostridium difficile* infection emphasized severe cases of life-threatening colitis. As diagnostic testing improved and awareness of the infection became widespread, milder cases were documented, and the presenting manifestations...
Clinical Manifestations of Disease

Chapter 20

were recognized to vary widely. Treatment advances can change the course of disease, so previously common eventual clinical manifestations may occur with less frequency. Moreover, new treatments bring the possibility of new iatrogenic disease, which may coexist and combine with underlying diseases in new ways.

How Can I Use the Results in Generating a Differential Diagnosis?

For studies that prove valid and applicable to your patients, knowing the clinical manifestations of possible conditions will help you to generate a differential diagnosis. A few findings may occur in almost all patients with the disease. As this proportion nears 100%, the absence of these findings allows you to omit the disease from your differential diagnosis. The presence of findings that occur in the range of 10% to 90% of patients with the disease suggests that the condition should remain among those you are considering as an explanation for your patient's presentation. Some manifestations occur seldom enough—in fewer than 10% of patients—that their presence would not prompt consideration of the illness in your differential diagnosis.

Using the Guide

Spittell and colleagues did not describe the referral filters through which their patients arrived. However, the Mayo Clinic provides community hospital care for 125,000 local residents along with referred care for many others. Of the 235 patients, 158 (67%) were men, and their mean age was close to that of the patient presenting in the emergency department. The authors did not describe patients’ comorbid conditions, socioeconomic status, race, or cultural background. Thus, although some uncertainty remains, these patients are sufficiently similar to the patient presenting in the emergency department to allow application of the results.

The study was published in 1993 and reported on patients seen from 1980 to 1990. You know of no new diseases emerging since then that would change the clinical features of aortic dissection. Both diagnostic testing for suspected dissection and treatment of hypertension (a major risk factor for aortic dissection) have changed during this period, but you expect that they would not change the presenting clinical features of acute dissection.

Clinical Resolution

Based on the evidence from Spittell and colleagues, you realize that absence of pulse asymmetry does not rule out a diagnosis of aortic dissection. Given the presence of the aortic regurgitation murmur and diastolic hypotension, along with the patient’s known risk and the absence of findings for myocardial infarction, you become more confident that the patient’s diagnosis may be proximal aortic dissection. Indeed, when you return to work the next day, you learn that the patient’s aortogram confirmed aortic dissection of the ascending aorta and arch, complicated by aortic regurgitation and that the patient was taken to the operating room for emergency surgery.
REFERENCES


