

Supplemental Digital Appendix 1

Diagnostic Justification Scoring Rubric (Legacy Rubric)

1. **Differential:** Based on the diagnostic possibilities discussed did the student consider an appropriate range of diagnostic possibilities given the findings of the case?

0

Poor

1

Borderline

2

Competent

3

Excellent

2. **Recognition and use of key findings (Pertinent positives and negatives alike) in building an argument for the final diagnosis**

0

Poor

1

Borderline

2

Competent

3

Excellent

To be completed only for students rated poor or borderline in this section

- Some key findings not noted.*
- Student reported findings that were not present in this patient.*
- Some key findings were misinterpreted.*

3. **Thought Processes and Clinical Knowledge Utilization**

0

Poor

1

Borderline

2

Competent

3

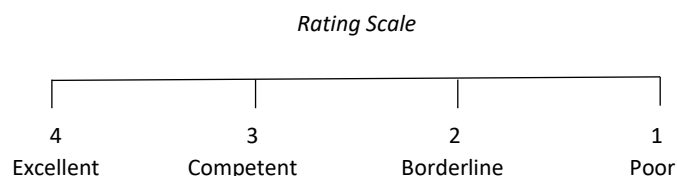
Excellent

To be completed only for students rated poor or borderline in this section

- No response or a response that re-asserted the diagnosis without providing supporting evidence.***
- Organization reflects routine unfocused data collection rather than an effort to link diagnostic models and patient findings (functioned as a reporter rather than interpreter of findings)***
- Student was overly focused on defending a single diagnosis. Student failed to consider alternative explanations.***
- Response reflects a superficial analysis and/or an oversimplified understanding of the patient's medical problems (Examples: Student used one or two pieces of information to support the diagnosis. Student failed to explicitly address disconfirming evidence and/or pertinent negatives).***
- Response reflects erroneous understanding of the constellation of findings***
- Conclusions that could be supported by data were not drawn***

Supplemental Digital Appendix 2

Overview of the Refined Essay Scoring Rubric Used to Generate Algorithm Training Data



Consideration of Differentials

Excellent – The differentials considered comprehensively cover the range of diagnostic possibilities, including the most likely and most crucial diagnoses, without being overly broad, obscure, or epidemiologically inconsistent with the patient data.

Competent – Differentials considered attempt to comprehensively cover the range of diagnostic possibilities, including the most likely and most crucial diagnoses, but exclude some common hypotheses given the patient data or include hypotheses that go well beyond the specific patient data provided, suggesting a functional, but incomplete understanding of the patient's problem.

Borderline – Differentials considered do not comprehensively cover the range of diagnostic possibilities, and are likely to be overly narrow or overly broad in a manner inconsistent with the patient data, but may include some of the most likely and most crucial diagnoses, suggesting a marginally functional understanding of the patient's problem.

Poor – Differentials considered do not comprehensively cover the range of diagnostic possibilities and suggest a dysfunctional understanding of the patient's problem; Differentials may comprise a minimal list of hypotheses that reflect an incomplete or oversimplified understanding of the chief complaint; Differentials may comprise a laundry list of possible diagnoses that include the patient's chief complaint as a symptom, but bear little to no relationship to the patient data; Differentials may or may not include the most likely and most crucial diagnoses; Differentials may include obvious incorrect, missing, or inaccurate information.

Recognition and Use of Findings

Excellent – Pertinent positive and negative findings are comprehensively identified, clearly specified, and incorporated into consideration of the diagnostic options, reflecting the appropriate range of history and physical information and a complete understanding of the chief complaint as well as the patient's medical problem.

Competent – Most pertinent findings are identified and incorporated into the consideration of diagnostic options; Pertinent negatives may be missing, as well as less salient pertinent positives; Data gathered may reflect a somewhat superficial understanding of the chief complaint or the patient's medical problem or suggest a functional, but incomplete repertoire of data gathering skill (e.g., missing physical examination data).

Borderline – Some pertinent findings are identified and incorporated into consideration of diagnostic options, but the data gathered are superficial, and key findings may be missing; Data gathered appears based on an oversimplified understanding of the chief complaint or the patient's medical problem and may reflect narrow focus on one or two differentials; Data gathered may suggest a marginally functional repertoire of data gathering skill (e.g., reliance on lab/testing results or repetitive use of the same one or two findings to rule up or down the differentials).

Poor – Few pertinent findings are identified and incorporated into consideration of diagnostic options; Data gathered may reflect narrow focus on one or two differentials; Key findings may be missing, misinterpreted, or stated as vague generalizations; Findings may be listed as pertinent without reference to diagnostic options (e.g., some basic findings may

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

be used repetitively, regardless of their utility to rule up or down the differentials); Data gathered may suggest a dysfunctional understanding of the data needed to investigate the chief complaint.

Workup

Excellent – Patient evaluation involves a comprehensive, yet targeted set of diagnostic procedures to validate a preferred diagnosis based on patient history and physical examination; Labs/tests are used for final confirmation of preferred diagnoses and not as a starting point for diagnostic thinking.

Competent – Use of labs/tests reflects a mostly targeted approach to patient evaluation, but may include labs/tests to eliminate hypotheses that could have been ruled down from history and physical information alone; Incorporation of lab/test findings into reasoning suggests a functional, but perhaps incomplete or superficial understanding of how diagnostic procedures should be used to investigate the chief complaint.

Borderline – Labs/tests may be prioritized over history and physical examination data, such that lab/test results are provided and some key H&P findings are not; Key labs/tests are ordered, but so also may be some unnecessary diagnostic procedures; Use of diagnostic procedures suggests “fishing” for data in light of a marginally functional understanding of the chief complaint or the patient’s medical problem.

Poor – Diagnostic procedures appear to have been used as an alternative to history and physical data gathering, with reasoning suggesting an almost exclusive use of diagnostic procedures to make a diagnosis; Use of diagnostic procedures may be scattershot, without clear application to a focused, confirmatory line of investigation, suggesting “fishing” for data in light of a dysfunctional understanding of the chief complaint; Labs/tests irrelevant to or unnecessary for ruling in or out the differentials may be ordered or some key confirmatory tests may be overlooked.

Diagnostic Thought Process

Excellent – Thinking represents a thorough, yet coherent and concise argument justifying the preferred diagnosis, reflecting a strong command of reasoning about the chief complaint; Pertinent positive and negative findings are linked to each diagnostic option with their implications for ruling it up or down clearly stated.

Competent – Thought process is mostly thorough, but may lack some coherence and/or concision; Some disorderly or repetitive thinking may be apparent; Pertinent positive and negative findings are linked to the diagnostic options, but some connections may be missed, unclear, or erroneously made, but still reflect a functional capacity to reason about the chief complaint.

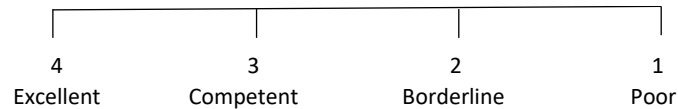
Borderline – Thought process is present, as evidenced by connection of some key findings to the diagnostic hypotheses listed, but thinking appears constrained by an oversimplified or superficial understanding of the data gathered or the chief complaint; The implications of pertinent negatives or less common pertinent positives may be missed or poorly understood; Reasoning may appear disorderly or repetitive, suggesting a marginal capacity to reason about the chief complaint.

Poor – Thought process lacks completeness and coherence; Findings are listed with mostly superficial analysis regarding their implications for ruling up/down diagnostic hypotheses; Argumentation may reflect over-focus on defending a single diagnosis or circular reasoning, where one hypothesis being preferred is used as data to rule down other hypotheses.

Supplemental Digital Appendix 3

Full Refined Essay Scoring Rubric for the Chest Pain Case (Essay Examples for Each Anchor Shown)

Consideration of Differentials



Excellent – The differentials considered comprehensively cover the range of diagnostic possibilities, including the most likely and most crucial diagnoses, without being overly broad, obscure, or epidemiologically inconsistent with the patient data.

This patient has 2 wks of worsening, exertional chest pain, located mid sternum lasting 10-15 min/episode, which occurs daily. He also has a hx of smoking, DM and HTN (with vitals suggesting poorly controlled HTN) so that makes unstable angina a high diagnostic possibility. CV exam did not demonstrate murmurs or new S3/S4, but this is mainly seen with actual MI. I ruled out respiratory causes based upon the 2 wk onset and benign lung exam. PTX or PE would be more acute presentation, plus he has had no trauma or recent immobility. Pneumonia would present with cough, sputum, fever, and lungs would likely not sound clear, so I ruled this out. He has no epigastric pain on palpation, and does not describe sx's associated with mealtime therefore I r/o GERD. Other Ddx include: -pericarditis - thought initially because of chest pain with associated dyspnea, but no recent illness and no rubs on cardiac auscultation. Also no pain on inspiration per hx. EKG did not demonstrate ST elevations in multiple leads so that confirmed for me. -drug-induced vasospasm - no hx of recreational drug use, specifically cocaine. -Aortic dissection - thought of this because he has high BP and chest pain, but no trauma, he was not in acute distress, his pulse was 90 but which AD I would expect a much higher pulse and him to report more back pain. -MSK - ruled out these etiologies based upon no hx of rash or burning/tingling prodrome, and no hx of trauma to chest. No point tenderness on exam and visual inspection did not reveal skin lesions. -MI - exertional chest pain radiating to both arms and jaw made me think of MI, but he has no associated diaphoresis, N/V, and negative cardiac enzymes helped me confirm. Additionally, his angina would not let up with rest if this were true MI. EKG demonstrated ST depression and T wave inversion which confirms angina as opposed to MI (which would show ST elevation most likely). The normal chest x ray again confirmed that his symptoms were not due to pulmonary etiology.

Competent – Differentials considered attempt to comprehensively cover the range of diagnostic possibilities, including the most likely and most crucial diagnoses, but exclude some common hypotheses given the patient data or include hypotheses that go well beyond the specific patient data provided, suggesting a functional, but incomplete understanding of the patient's problem.

This patient suffers from unstable angina. I considered a number of diagnoses when treating this patient. But my top five diagnoses I considered were, 1). Stable Angina vs Unstable Angina vs. MI. 2). Aortic dissection. 3) GERD. 4). Pneumothorax. 5). Anxiety/ Costochondritis 1). 2 weeks ago the patient had predictable chest pain that only occurred with vigorous activity such as climbing multiple flights of stairs. When he did vigorous activity he experienced dull pain 8/10 that would dissipate within a minute of resting. This describes stable angina, in that it is predictable, reproducible, and the pain is easily managed with short rest. Currently the chest pain has quickly progressed in the last week, making his clinical picture look more like unstable angina. In the last week he has chest pain 8/10 radiating to his arms, and neck with just walking across the room. He also has to rest for up to 15 mins for the pain to dissipate. It is also possible he had suffered a myocardial infarction during this time which is why I ordered an EKG and troponin level. The EKG showed some T wave inversions and ST depression showing myocardial ischemia and a normal troponin level, which makes a current MI unlikely. Another troponin should be checked. His risk factors which support the diagnosis of unstable angina include his BMI of 30, HTN treated with medication, DM type II treated with medication, 1PPD smoking habit, and a FMH of HTN and premature CAD- father had a heart attack in his 50's. These all serve to increase his risk and increase the likelihood that this patient has CAD. 2) Aortic Dissection- This patient's intense chest pain makes the inclusion of aortic dissection in the differential necessary. This patient denies any tearing, ripping, or sharp pain that radiates to his back or feet which decreases the likelihood of aortic dissection. It would also be unusual for the patient to have an aortic

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

dissection two weeks ago and still remain stable. Additionally I did not hear an early diastolic murmur of AR that is sometimes associated with Aortic dissection. 3). GERD is a common cause of chest pain and can lead to esophageal spasm which can mimic cardiac pain. This patient reports that he does not have a history of known acid reflux disease, his pain is not associated with eating or position, and he reports no cough, no metallic taste in his mouth, and no trouble swallowing food or liquids. These factors, plus the EKG findings suggest a cardiac etiology as more likely than GERD. 4). A pneumothorax may present like this patient did with chest pain and shortness of breath. However, all of his breath sounds were clear to auscultation in all lung fields. And percussion was not hyper resonant in any of the lung fields. This decreases the chance of a pneumothorax. 5). Anxiety/Costochondritis- Patient was not anxious, and palpation of chest- no pain.

Borderline – Differentials considered do not comprehensively cover the range of diagnostic possibilities, and are likely to be overly narrow or overly broad in a manner inconsistent with the patient data, but may include some of the most likely and most crucial diagnoses, suggesting a marginally functional understanding of the patient's problem.

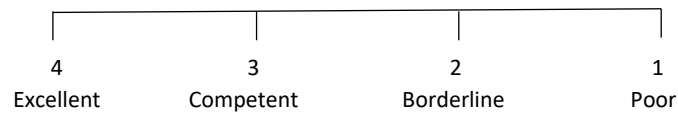
Stable Angina- For - Chest pain in center of chest radiating to neck and left arm, relieves with 10- 15 minutes of rest, brought on by activity, Chest pain as a dull ache history of diabetes and hypertension, shortness of breath with pain normal heart physical exam, normal lung physical exam, EKG, and Cardiac Enzymes confirmed MI- For-For: Chest pain in the center of chest radiating to neck and left arm, history of diabetes, and hypertension, Against- Pain is relive by rest, no N/V/ No fever, Cardiac Enzyme confirmed DVT- Against- no history long travel or sitting still, no cough, no continue shortness of breath, no blood spit up, pain is not sharp stabbing, no fever Pneumonia- Against- No fever, No cough, no continuous shortness of breath, no sick contact, normal lung physical exam- no crackles GERD- Against- No history of Heart Burn, no correlation with eating and pain, no epigastric tenderness, Psych- Against- No history of anxiety, no nervousness, no stress, patient seemed calmed Musculoskeletal- Costochondritis- For: construction worker. Against-No increase in activity Pericarditis- Against-Chest pain no continuous, no relive by leaning forward, only by rest, no fever, no signs of infection, no pericardial knock, onset of pain is by activity Asthma- For- Shortness of breath with activity , Against-No cough, No wheezing on PE, No History of asthma Unstable Angina- For: Chest pain located in the center of the chest radiating to neck and down both arms, history of diabetes and hypertension, Against- Relieve by rest by 15 minutes, Never occurs at rest

Poor – Differentials considered do not comprehensively cover the range of diagnostic possibilities and suggest a dysfunctional understanding of the patient's problem; Differentials may comprise a minimal list of hypotheses that reflect an incomplete or oversimplified understanding of the chief complaint; Differentials may comprise a laundry list of possible diagnoses that include the patient's chief complaint as a symptom, but bear little to no relationship to the patient data; Differentials may or may not include the most likely and most crucial diagnoses; Differentials may include obvious incorrect, missing, or inaccurate information.

Mr. K____ is a 45 year old obese male who presents with a 2-week history of chest pain on exertion. At this point the most likely diagnosis is unstable angina. Supporting: The chest pain is substernal, occurs with exertion, and is relieved by rest. It radiates to the left neck and down both arms. Patient is obese, and he has history of hypertension, DM, 25 pack year-Hx smoking cigarettes, FMH significant for MGF who had MI at age 50; all of these are risk factors for MI at young age. Physical exam was unremarkable, which is consistent with angina. EKG findings indicate t-wave inversion and ST depression in lateral leads 1, AVL, v5-6, which is consistent with ischemia. Detracting: none Though the pain occurs with exertion, because it is the first reported episode of chest pain, it is classified as unstable angina. After this, if he continues to have chest pain with exertion, it will be classified as stable angina. Other differentials include: stable angina- see above MI - supporting: substernal chest pain radiating to left neck and down both arms. Obese, Hx HTN, DM, 25 pack-year Hx smoking cigarettes, FMH of MI in MGF at 50. Detracting: Pain is dull ache, no diaphoresis, relieved with rest. EKG more indicative of ischemia than infarction. Troponins negative. GERD - may present as substernal chest pain, common in obese patients. Detracting: pain radiates to left neck and down both arms. Obese, Hx HTN, DM, 25 pack-year Hx smoking cigarettes, FMH of MI in MGF at 50. No epigastric tenderness on exam. No Hx heartburn. EKG indicative of ischemia. Esophageal spasm - may present as substernal chest pain. Detracting: pain radiates to left neck and down both arms. Obese, Hx HTN, DM, 25 pack-year Hx smoking cigarettes, FMH of MI in MGF at 50. No epigastric tenderness on exam. No Hx heartburn. EKG indicative of ischemia.

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

Recognition and Use of Findings



Excellent – Pertinent positive and negative findings are comprehensively identified, clearly specified, and incorporated into consideration of the diagnostic options, reflecting the appropriate range of history and physical information and a complete understanding of the chief complaint as well as the patient's medical problem.

Mr. K ___ is a 45 yr old male with a hx of DM and HTN and 25 pk-yr smoking hx who presented to the ED with chest pain. My top differentials were Angina (stable, unstable), MI, PE, pneumonia, and the rest stated above. Angina/MI: He has multiple cardiovascular risk factors for angina including hx of HTN, DM, and tobacco use. However, he is relatively wrong to have an MI. The pain happens only with exertion, has happened multiple times over the past couple weeks which points more towards angina. The pain has radiated to his arms and neck, was associated with SOB which make and he does have a hx of MI in his maternal grandfather at age 50. However, there was no sweating associated with the episode which would be the most specific associated sx with MI. In addition, the pain was characterized as dull, vs "pressure on chest" which is more typical for MI. As far as physical exam, his BP was high indicating that it isn't controlled with his BP meds. Labs showed normal EKG, Troponin and CK-MB making AN ACUTE MI less likely. 2) PE- His chest pain and SOB hx make it a possibility, however he doesn't have any of the Triad risk factors; No recent travel, no hypercoagulable state, and no endothelium injury. On exam, he was not tachypnic or tachycardia and Respiratory exam was benign. He had no edema. Spiral CT was negative. 3)Pneumonia- patients may present with chest pain, however he had no cough, fever, no leukocytosis, no fatigue. 4)Dissecting Aortic aneurysm- Is a serious and potentially fatal cause of chest pain that must be ruled out early, he did not have radiation to his back, pain was not tearing in nature and his vitals were stable. 5)GERD: Can definitely present with chest pain. however he denies history of reflux, pain not associated with food/eating, and he had no cough. 6) Costochondritis- There was no pain on palpation during exam making this less likely. Thus, with his risk factors, hx of pain only with exertion, and description of pain as dull vs. pressure like, a relatively benign exam except for HTN, and labs negative for MI, my most likely diagnosis is Stable angina.

Competent – Most pertinent findings are identified and incorporated into the consideration of diagnostic options; Pertinent negatives may be missing, as well as less salient pertinent positives; Data gathered may reflect a somewhat superficial understanding of the chief complaint or the patient's medical problem or suggest a functional, but incomplete repertoire of data gathering skill (e.g., missing physical examination data).

Mr. K ___ is a 45 year old male with a pmh significant for hypertension, diabetes, and 25 pack year history of smoking who presents with 2 weeks of progressively worsening chest pain. First noticed the pain when lifting a 2x4 at work as a construction worker. The chest pain in mid-sternal 8/10 and dull in quality. It comes and goes with exertion but recently has been at rest as well. Pain also radiates up to his neck and to his arms. He has been short of breath with this and cannot get enough breath in. Pain gets better with resting for 10-15min No fever, diaphoresis, nausea, vomiting, acid reflux. No personal or FMH of heart disease. Hypertensive at 160/90 with rest of vitals normal. EKG showed T wave inversion and ST depressions in the lateral leads of I, AVL, V5, and V6. Troponins were 0.3 (nl). Cardiac cath showed normal vascular anatomy with no stenosis, aneurysm, or AV malformation. Cardiac echo was normal. Based upon this the likely diagnosis is unstable angina. Other differentials considered were myocardial infarction (STEMI, NSTEMI), muscle strain, GERD, and esophageal spasm. MI was considered based upon the nature of the pain being substernal with radiation to the neck and arms. This was made less likely based upon the chest pain being dull instead of pressure, troponins showing normal levels, and the cath being normal. Muscle strain was also considered based upon it occurring when lifting a 2x4 but the nature of the pain being substernal and no tenderness to palpation on exam made this less likely. GERD/esophageal spasm was also considered based upon the nature of the complaint but he had no history of nausea/vomiting and no history of acid reflux.

Borderline – Some pertinent findings are identified and incorporated into consideration of diagnostic options, but the data gathered are superficial, and key findings may be missing; Data gathered appears based on an oversimplified understanding of the chief complaint or the patient's medical problem and may reflect narrow focus on one or two differentials; Data gathered may suggest a marginally functional repertoire of data gathering

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

skill (e.g., reliance on lab/testing results or repetitive use of the same one or two findings to rule up or down the differentials).

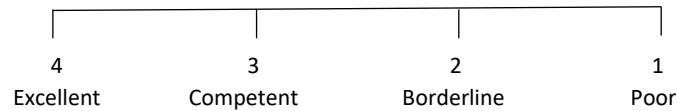
This 45 y/o male presented with chest pain which is most concerning for a myocardial infarction vs. angina. Other causes could be pulmonary embolism, GERD, esophagitis, or gastritis. 1. Stable Angina: Positive: chest pain radiating to shoulder/neck area of dull pain on exertion and relieves with rest. Onset is two weeks ago with risk factors of smoking, obesity, and diabetes mellitus. Cardiac enzymes were negative and EKG showed some ST depressions in I, AvL and V5-V6. His stress test was negative, so cardiac cath will be requested to further workup his angina. 2. Unstable Angina: Positive: chest pain radiating to shoulder/neck area of dull pain. Negative: relieves with rest and has occurred repeatedly over the last two weeks. This is still a possible diagnosis if no narrowing is found on cardiac cath and intermittent thrombus could be forming in unstable plaques. 3. Myocardial infarction: Positive: chest pain radiating to shoulder/neck area of dull pain, increased blood pressure. Negative: no diaphoresis, nausea/vomiting. There are also no elevated cardiac enzymes to support this as the current problem although some T wave inversions on EKG could signify old infarcts. 4. CAD: Positives: Risk factors of smoking, obesity, and hypertension. Negatives: No family history Treadmill stress test did not bring on chest pain, although a cardiac cath needs to be executed to rule this out. 5. Pulmonary embolism: Positive: risk factors of smoking Negatives: active occupation, low risk for PE. Negative D-dimer helps rule this down. 6. Esophagitis/gastritis/GERD: Positives: chest pain. Negatives: Pain not associated with eating, only exertion. No history of heartburn. 7. Muscle strain: Positive: chest pain Negative: no pain on palpation of chest wall.

Poor – Few pertinent findings are identified and incorporated into consideration of diagnostic options; Data gathered may reflect narrow focus on one or two differentials; Key findings may be missing, misinterpreted, or stated as vague generalizations; Findings may be listed as pertinent without reference to diagnostic options (e.g., some basic findings may be used repetitively, regardless of their utility to rule up or down the differentials); Data gathered may suggest a dysfunctional understanding of the data needed to investigate the chief complaint.

The radiation of pt's substernal CP to neck and both arms is an unusual presentation for angina, although the location, occurrence with exertion, and relief with rest are hallmarks of stable anginal pain; the patient's DM2 and self-admittedly uncontrolled blood glucose could account for atypical signs. One wishes, before diagnosing angina, to exclude ACS; the EKG's T wave inversion and ST depression painted the picture of a former infarct or an ongoing NSTEMI; the fact that cardiac enzymes were negative and the echo showed no trace of wall motion abnormality points more towards the former. The fact that he is a smoker raised serious concerns about carotid artery occlusion, which was definitively ruled out via angiogram; any lung or mediastinal masses were ruled out with the normal CXR. The stress test will provide more conclusive answers as to the anginal nature of the pain.

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

Workup



Excellent – Patient evaluation involves a comprehensive, yet targeted set of diagnostic procedures to validate a preferred diagnosis based on patient history and physical examination; Labs/tests are used for final confirmation of preferred diagnoses and not as a starting point for diagnostic thinking.

For this 45 year old male with chest pain, stable angina is my most likely diagnosis. First, his pain is exertional, is accompanied by shortness of breath, and is relieved by rest. This fits the picture of early cardiac-related chest pain. The pain also radiates up to his neck and down his arms, which can be seen in cardiac-type chest pain. His past medical history also fits the picture for someone who would have ACS/angina as he has DM, HTN, and he smokes. Some pertinent negatives that support this diagnosis are the lack of fever, pleuritic pain, or pain that is reproducible on palpation. Finally, his EKG shows some ST depression in AVL, V5, and V6 which suggest some possible cardiac ischemia. However, I would have expected his coronary angiogram to have been abnormal with at least some narrowing of arteries. My next possible diagnosis was an MI. I was thinking this for essentially the same reasons as I was thinking about stable angina. However, this diagnosis is less likely because of the negative cardiac enzymes, troponin, the EKG showing no ST elevation, and the negative coronary angiogram. My next diagnosis I considered was a pulmonary embolism. This diagnosis is less likely because of the 2 week history of chest pain, the fact that it is only exertional, and the fact that it is relieved by rest. In addition, the patient had no period of immobilization and his D-dimer was negative. I also considered a chest wall muscle strain as a diagnosis. I considered this because the pain was exertional and began after he was doing some heavy lifting. However, with a muscle strain, I would expect the pain to be reproducible and I would not expect accompanying shortness of breath. Another diagnosis I considered was an aortic dissection. I considered this because of his hypertension and smoking risk factors. However, I would expect this to be a constant pain instead of one aggravated by exertion and relieved by rest.

Competent – Use of labs/tests reflects a mostly targeted approach to patient evaluation, but may include labs/tests to eliminate hypotheses that could have been ruled down from history and physical information alone; Incorporation of lab/test findings into reasoning suggests a functional, but perhaps incomplete or superficial understanding of how diagnostic procedures should be used to investigate the chief complaint.

Mr. K___ is a 45 year old male with a history of HTN and T2DM who presents to the ED with 2 weeks of worsening chest pain and dyspnea. Differentials include MI, PE, Musculoskeletal pain, GERD, Diffuse Esophageal Spasm, Pneumonia, Unstable Angina, and Trauma. PE: This pt has some increased risk factors for PE such as obesity, HTN, and DM. however, the duration of the illness, no hx of recent travel or prolonged sitting, no pleuritic chest pain, no swollen/red/tender leg on exam, no respiratory distress makes this less likely. Due to the low likelihood of PE in this pt, I ordered D-dimer to rule it out, which it did. Pericarditis: Chest pain would be more pleuritic, pt has no history of recent viral illness, and the pain is intermittent/worse with exertion, not typically seen with pericarditis. Normal CV exam, no friction rub heard. No pericardial effusion on CXR. Musculoskeletal pain/Trauma: Costochondritis could cause central chest pain, but would not be sudden onset pain worse with exertion, no hx of trauma, no tenderness to palpation, no broken ribs on CXR rule these down. GERD: Often confused with substernal chest pain, described more as burning (this pt had dull pain), pain was not worse with meals, he has no hx of GERD. Aortic dissection is less likely due to duration of illness, no radiation to the back or tearing quality of the pain, pulses were equal in upper/lower extremities, no carotid bruits or abdominal bruits on exam. CXR did not show mediastinal widening. Pneumonia can present with only Sx being chest pain and dyspnea, and patient is also a smoker, increasing risk of resp infxn. But this is less likely in an afebrile patient with no cough, rhinorrhea, headache. Resp exam was normal, no evidence of pneumonia on CXR. Diffuse Esophageal Spasm presents with pain very similar to MI pain, intermittent and often substernal. However, this patient has worsening with exertion and dyspnea, which makes DES less likely Unstable Angina/NSTEMI is the most likely dx in this patient, due to worsening substernal dull chest pain that worsens with exertion and improves with rest. Pt also has risk factors for heart disease like HTN, obesity, DM, and smoking. Pain 2 weeks ago was sudden onset, this may have been the NSTEMI, and now he has dyspnea which could be some CHF due to the MI. EKG showed T inversion and ST dep in lateral leads, which is consistent with this, but stress testing and cardiac cath showed no evidence of ischemia or blocked vessels, so I am not 100% on this dx. Troponins and CKMB were normal which I expected with the initial event being 2 weeks ago. I would have expected exercise stress testing to reproduce the chest pain the patient reports with exertion, and expected to see some vessel blockage on catheterization. I would do further tests to rule out esophageal spasm/dysmotility or consider Prinzmetal's angina.

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

Borderline – Labs/tests may be prioritized over history and physical examination data, such that lab/test results are provided and some key H&P findings are not; Key labs/tests are ordered, but so also may be some unnecessary diagnostic procedures; Use of diagnostic procedures suggests “fishing” for data in light of a marginally functional understanding of the chief complaint or the patient’s medical problem.

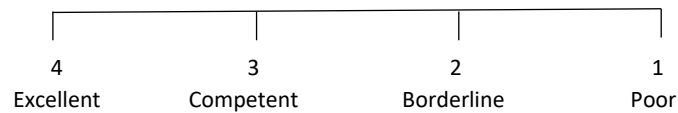
unstable angina supportive findings -dull chest pain, 8/10 that radiates to both arms and chest -pain consistently comes on with exertion, initially when carrying heavy items but now has progressed to when walking upstairs at home, pain goes away upon resting in 10-15 minutes -prior to two weeks ago, no prior episodes of chest pain like this -no prior illness, trauma, history of GERD -history of diabetes, HTN (poorly controlled) - 45 year old male, epidemiologically more likely to have CAD -25 pack year smoking history -cardiac enzymes were normal -EKG showed T wave inversion and ST segment depression in lateral leads unsupportive findings -expected to find visible stenosis on angiogram, however, a lack of stenosis in main coronary arteries is possible in unstable angina STEMI/NSTEMI supportive findings -history as mentioned above with increasing pain on exertion is consistent with STEMI/NSTEMI as well as past medical history and other risk factors for ACS mentioned above -ST segment depression and T wave inversion can be seen in NSTEMI unsupportive -normal cardiac enzymes and a negative cath makes this diagnosis unlikely. Will repeat cardiac enzymes over next few hours to see if there is a rise as it can take 4-6 hours for this to rise after myocardial damage. If enzymes remain negative then this is ruled out. -Additionally, pain goes away upon rest, with NSTEMI/STEMI pain usually remains despite resting - no signs of ST segment elevation which rules out STEMI stable angina unsupportive -recent increase in symptoms and EKG changes rule out this diagnosis and classify it into at least unstable angina GERD/gastric ulcer supportive findings -these can cause epigastric pain that can be interpreted as chest pain unsupportive -no history of acid reflux, does not awake with acidic taste in back of mouth, foods don't exacerbate pain, no pain to palpation of epigastric region, which makes diagnosis very unlikely costochondritis/trauma supportive -can cause chest pain unsupportive -no history of trauma, no pain to palpation of sternum or costochondral junctions or ribs which makes diagnoses very unlikely pleurisy/pneumonia unsupportive findings -pain not exacerbated by breathing. - no history of URI, no fever, chills, pericarditis unsupportive findings -EKG is not suggestive of pericarditis (diffuse ST elevation) - no recent infection - pain is non pleuritic

Poor – Diagnostic procedures appear to have been used as an alternative to history and physical data gathering, with reasoning suggesting an almost exclusive use of diagnostic procedures to make a diagnosis; Use of diagnostic procedures may be scattershot, without clear application to a focused, confirmatory line of investigation, suggesting “fishing” for data in light of a dysfunctional understanding of the chief complaint; Labs/tests irrelevant to or unnecessary for ruling in or out the differentials may be ordered or some key confirmatory tests may be overlooked.

This patient likely has worsening stable angina bordering on unstable angina. His pain and dyspnea occur only on exertion. Which rules up angina and down MI aortic dissection. His pain started 2 weeks ago which rules down MI, aortic dissection. His chest pain is substernal this supports angina, MI, GERD, PUD however it radiates to his arms bilaterally and neck which is uncommon for MI to radiate bilaterally and GERD and PUD do not radiate at all. Pericarditis is unlikely because the patient has had no symptoms of infection, his pain is exertional (not relieved by sitting forward), and no friction rubs were heard on cardiac auscultation. The pt denied N/V, sour taste in his mouth and heartburn making GERD unlikely. He does not regularly take NSAIDs, has had no melena or hematemesis. And his pain/dyspnea is not consistent with PUD. The pt denies cough, fatigue, and had a normal CXR ruling down pneumonia. He has a smoking history by no dx of COPD making exacerbation unlikely. Normal CV exam, along with no history of CHF make exacerbation unlikely.

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

Diagnostic Thought Process



Excellent – Thinking represents a thorough, yet coherent and concise argument justifying the preferred diagnosis, reflecting a strong command of reasoning about the chief complaint; Pertinent positive and negative findings are linked to each diagnostic option with their implications for ruling it up or down clearly stated.

Final Diagnosis: Unstable Angina Pt presented with a 2 week hx of sudden onset exertional CP with no prior occurrence that had worsened to CP with minimal exertion. The pain was relieved by rest and associated with exertional dyspnea. Pain was described as dull, sub-sternal, 8/10 in severity, with radiation down both arms and into the neck. Pertinent negatives were no diaphoresis, no N/V, and no syncope. Pt has a PMH of HTN and DM as well as hx of smoking which also predispose to ACS. Pt has a FMH of HTN, DM, and MI @ 50 which further supports a diagnosis of ACS. Physical exam was unremarkable for any cardiac pathology. Pt denied any history of illicit drug use which helps rule out a tox cause of his chest pain which a negative drug screen supports. EKG showed T wave inversion and ST depression in the lateral leads which, along with a negative troponin and the history of physical data support a diagnosis of unstable angina I also considered PE as a diagnosis but the pt has no history of immobilization and does not describe the pain as pleuritic. He also states that the dyspnea is not constant but is related to exertion with chest pain. Physical examination did not reveal any evidence of PE/DVT and as the pre-test probability of a PE/DVT was low, I ordered a confirmatory D-dimer which was also negative I also considered pneumothorax as a diagnosis as the pt describes a history of sudden onset chest pain and dyspnea, but he relates this to exertion and it has progressively gotten worse over the last two weeks. Physical examination did not reveal any decreased breath sounds and the patient was saturating well on room air with no increased work of breathing. A CXR did not reveal any evidence of pneumothorax I also considered pericardial tamponade as a diagnosis but a history of a 2 week duration, no preceding trauma, and a non-acute presentation along with no evidence of JVD, muffled heart sounds, or electrical alternans on EKG allowed me to rule down this diagnosis Aortic dissection was an item high on my differential list and as the patient has a history of HTN, remained high on my list throughout the interview. The history which indicated a progressing worsening of CP and SOB with exertion with no hx of tearing CP radiating to the back helped me rule down this item MSK pain was initially considered as the patient reported that the pain began while he was lifting as his job, but the patient reports that he did not feel any initial trauma during the incident and reports that the pain has occurred with subsequent minimal exertion allows me to rule down this item. The physical examination showed no reproducible chest pain with gentle pressure of the stethoscope.

Competent – Thought process is mostly thorough, but may lack some coherence and/or concision; Some disorderly or repetitive thinking may be apparent; Pertinent positive and negative findings are linked to the diagnostic options, but some connections may be missed, unclear, or erroneously made, but still reflect a functional capacity to reason about the chief complaint.

Substernal chest pain on exertion is concerning for acute coronary syndrome, in the forms of angina, unstable angina, STEMI and NSTEMI. He has several risk factors for ACS: current smoker, DM, HTN, and FH of early MI. This patient has had recent intermittent exertional substernal chest pressure and dyspnea that is alleviated with rest. However, the amount of exertion required to produce pain has decreased to 20 feet. He is currently hypertensive, and reports his BP is normally well controlled with medications. He took his medications today. Pain radiates into the jaw and arms bilaterally. No palpitations, nausea, or diaphoresis occurred with the pain, as is classically seen in MI. Pain was relieved by rest in the past, but currently is not relieved by rest. This lack of resolution makes stable angina not a possibility. His decreased exercise tolerance makes unstable angina a likely cause, but the EKG changes seen at rest indicate unstable angina is no longer his diagnosis. His EKG showed ST depression and T wave inversions in the lateral leads suggesting he is having an NSTEMI. His troponin was WNL, but his pain onset less than 4 hours ago so a repeat troponin is needed. Pneumonia is an unlikely cause of his chest pain without fever/chills, productive cough, or recent sick contacts. His lungs were clear to auscultation bilaterally and his CXR showed no sign of infiltrate. His CBC showed no elevated WBC count. Aortic Dissection is an unlikely cause of his chest pain as he has strong pulses bilaterally, reports no weakness or presyncope/syncope, and pain is not a tearing quality radiating to the back. Additionally, there was no widening of the mediastinum on CXR and he had intact strength and sensation in his upper extremities bilaterally. Pulmonary embolism is an unlikely cause of his symptoms due to his prolonged course of chest pain and intermittent pain prior to today. PE pain is acute, pleuritic and associated with cough or hemoptysis. PE was excluded due to his negative D-dimer. With pain not being positional or pleuritic, pericarditis is an unlikely etiology. He reports no recent URI or GI symptoms to suggest viral infection, the most

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

common cause of pericarditis, and there was no friction rub present on auscultation. Pain was not reproducible on palpation of chest wall and pain is not diffuse across rib cage, making costochondritis an unlikely source of his pain.

Borderline— Thought process is present, as evidenced by connection of some key findings to the diagnostic hypotheses listed, but thinking appears constrained by an oversimplified or superficial understanding of the data gathered or the chief complaint; The implications of pertinent negatives or less common pertinent positives may be missed or poorly understood; Reasoning may appear disorderly or repetitive, suggesting a marginal capacity to reason about the chief complaint.

The initial presentation was chest pain. My differential for chest pain was Myocardial infarction, Pulmonary Embolism, GERD, Costochondritis, Stable Angina, pneumonia, esophageal spasm, and gastroparesis. The initial history gave me that the chest pain was dull and severe and started 2 weeks ago. This lead me to think that GERD, Stable Angina, and costochondritis were higher on my differential. Then he stated that the pain goes away with rest and returns on activity. This made me think more of Stable angina, MI and PE. He stated that he smoked 1 pack per day of cigarettes which raised my suspicion more for MI, PE, and Stable angina. He then went on to state he had a history of diabetes and hypertension. These diagnoses raised my suspicion more for MI, PE and stable angina. He stated the pain was not worse with food and that he hasn't had heartburn. This lowered my suspicion of GERD, esophageal spasm, and gastroparesis considerably. He then said that he had not pain when pressing on his chest which ruled down costochondritis. His family history was significant for MI, diabetes and hypertension again raising my suspicion for MI, stable angina, and PE. On exam his lungs were clear to auscultation and his heart sounds were normal with no murmurs. This ruled out pneumonia as I expected to hear an area of consolidation. He also didn't have a fever which ruled out most infectious processes. This left me with MI, stable angina, and PE. His labs showed T wave inversion and ST depression, which lead me to MI and stable angina as my differential. Then his troponins and CK-MB came back in the normal range ruling out an acute MI and leading me to stable angina as a diagnosis.

Poor— Thought process lacks completeness and coherence; Findings are listed with mostly superficial analysis regarding their implications for ruling up/down diagnostic hypotheses; Argumentation may reflect over-focus on defending a single diagnosis or circular reasoning, where one hypothesis being preferred is used as data to rule down other hypotheses.

Mr. K___ presented with sternal chest pain and shortness of breath that is unremitting, is exacerbated by movement and relieved when he rests. The pain started 2 weeks ago and has been getting progressively worse since that time with less movement causing the pain as time went on. His EKG was significant for inverted T waves in lead one and two indicating myocardial ischemia. He also had elevated QRS complexes in the anterior leads. Mr. Kessler denied an inciting event and thus trauma was ruled down. CHF was ruled down by the lack of edema and pulmonary congestion and his normal cardiac enzymes. Lung cancer was ruled down by his normal CXR. Unstable Angina, pericarditis and myocarditis are ruled down by his normal cardiac enzymes and his normal cardiac catheterization and his normal swanz gantz catheter. He also had a normal event monitor and a normal echo. AAA was ruled down by normal CXR and History and physical.

Supplemental digital content for Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: Initial validity evidence. Acad Med.

Supplemental Digital Appendix 4

Senior Medical Student and Faculty Rating Data^a for the Five Cases Selected for Study

Case	Abdominal Pain #1		Abdominal Pain #2		Chest Pain		Edema		Fever		
N	144		138		138		132		148		
Rater	Student	Faculty	Student	Faculty	Student	Faculty	Student	Faculty	Student	Faculty	
Average Criterion Rating	Differentials	2.4 (.74)	3.3 (.43)	2.7 (.65)	3.3 (.48)	2.7 (.77)	3.3 (.53)	3.0 (.61)	3.1 (.54)	2.7 (.94)	3.3 (.75)
	Findings	2.5 (.76)	3.1 (.51)	2.7 (.65)	2.9 (.56)	2.7 (.77)	3.1 (.44)	2.5 (.62)	2.9 (.56)	2.2 (.86)	2.6 (.70)
	Workup	2.7 (.80)	NA ^a	2.5 (.76)	NA ^a	2.8 (.80)	NA ^a	3.1 (.84)	NA ^a	2.2 (.92)	NA ^a
	Thought Process	2.3 (.81)	3.0 (.58)	2.3 (.66)	3.2 (.52)	2.5 (.75)	3.1 (.44)	2.6 (.64)	3.2 (.58)	2.2 (.87)	2.7 (.70)
Overall Percent Score ^b	61.8% (15.3)	78.3% (10.8)	63.4% (11.9)	78.4% (10.3)	66.3% (15.3)	79.4% (9.8)	69.7% (10.9)	77.1% (12.3)	58.5% (18.5)	71.9% (13.9)	
Inter-rater Correlation ^c	.97	.31 / .36	.97	.37 / .61	.97	.44 / .65	.97	.82 / .39	.96	.69 / .67	
Student-faculty correlation ^{a,c}	.53		.53		.60		.47		.69		

^aPrior to this study, faculty used a legacy rubric to assign ratings for summative assessment purposes. The legacy rubric featured three assessment criteria: *Differential Diagnosis*, *Recognition and Use of Findings*, and *Thought Process* (see Digital Supplemental Appendix 1). Medical students used a refined rubric designed for research purposes that featured four assessment criteria: *Differential Diagnosis*, *Recognition and Use of Findings*, *Workup*, and *Thought Process* (see Digital Supplemental Appendix 2).

^bBecause the refined rubric contained an additional assessment criterion (*Workup*), the maximum overall rating was 16, whereas the maximum overall rating for the legacy rubric is 12. For this reason, percentage scores were used for comparison.

^cCorrelation coefficients were calculated using Overall Sum Ratings. Faculty inter-rater correlation was calculated separately for each year of administration because case author differed across years.