Using Illness Scripts to Teach Clinical Reasoning Skills to Medical Students

Anna Lee, PhD; Gavin M. Joynt, MB, BCh; Alex K.T. Lee, MPhil; Anthony M.H. Ho, MD; Michele Groves, PhD; Alexander C. Vlantis, MB, BCh; Ronald C.W. Ma, MB, BChir; Colman S.C. Fung, MB,BS; Cindy S.T. Aun, MD

Background and Objectives: *Most medical students learn clinical reasoning skills informally during clinical rotations that have varying quality of supervision. We conducted a randomized controlled trial to determine if a workshop that uses "illness scripts" could improve students' clinical reasoning skills when making diagnoses of patients portrayed in written scenarios.* Methods: *In 2007–2008, 53 fourth-year medical students were randomly assigned to either a family medicine (intervention) or psychiatry (control) clerkship at The Chinese University of Hong Kong. Students in the intervention group participated in a 3-hour workshop on clinical reasoning that used illness scripts. The workshop was conducted with small-group teaching using a Web-based set of clinical reasoning problems, individualized feedback, and demonstration of tutors' reasoning aloud. The effectiveness of the intervention was assessed using the Diagnostic Thinking Inventory (DTI) and the measurement of individual students' performance in solving clinical reasoning problems (CRP).* Results: *The post-intervention overall DTI scores between groups were similar (mean difference 0, 95% confidence interval [CI]= -7.4 to 7.4). However, the total scores on the CRP assessment were 14% (95% CI=8% to 21%) higher in the intervention group than in controls.* Conclusion: *A workshop on illness scripts may have some benefit for improving diagnostic performance in clinical reasoning problems.*

(Fam Med 2010;42(4):255-61.)

The ability to diagnose effectively and accurately requires an appropriate clinical knowledge and sound clinical reasoning skills. Medical educators agree that clinical reasoning is a central component of physician competence.[1] Currently, most students learn clinical reasoning skills informally in clinical rotations with varying quality of supervision.

When making a diagnosis, students often reason by an iterative process of hypothesis generation and testing of one symptom at a time.[2] Since they may be unable to synthesize clinical features into meaningful clusters or syndromes, their clinical reasoning styles are typically characterized by poorly organized knowledge structure. In comparison, expert clinicians reason by organizing and prioritizing syndrome recognition through comparing and contrasting key clinical features in making a diagnosis ("illness scripts").[2] The illness script concept[3] provides a theoretical framework to explain how medical diagnostic knowledge can be organized for diagnostic problem solving.

Illness scripts are a product of extensive experience with patients superimposed on a formal knowledge structure.[4] Expert clinicians use illness scripts most of the time in their clinical reasoning since it involves a knowledge-driven model of pattern recognition that may be more efficient than hypothetico-deductive reasoning used by students.[5]

The Chinese University of Hong Kong undergraduate medical curriculum is a body-system-based model taught over 5 years. In response to faculty members' perception that our students had poor clinical reasoning skills when starting clinical rotations during their fifth and final year of medical school, we constructed a brief 3-hour educational intervention to improve those skills. Our hypothesis was that an educational intervention would promote and improve students' competency in

From the Department of Anaesthesia and Intensive Care (Dr Lee, Dr Joynt, Mr Lee, Dr Ho, Dr Aun), Department of Otorhinolaryngology, Head and Neck Surgery (Dr Vlantis), Department of Medicine and Therapeutics (Dr Ma), School of Public Health and Primary Care (Dr Fung), The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, NT, Hong Kong, China; and Faculty of Health Sciences, University of Queensland, Brisbane, Queensland, Australia (Dr Groves).

clinical reasoning in a systematic manner. The objective of our study was to determine if a brief workshop to teach students to refine their knowledge organization through the use of illness scripts could improve students' clinical reasoning skills for diagnosis.

Methods

The workshop started with a 20-minute lecture to all students. It introduced the key elements of the clinical diagnostic process.[6] This was followed by a 1-1/4 hour small-group tutorial on problem representation and developing an illness script and a final tutorial, also 1-1/4 hours long, in the computing laboratory on developing and selecting an appropriate illness script. A manual for teachers wishing to conduct a similar workshop is available from the authors by request.

Problem Representation and Developing an Illness Script

The objective in the workshop was for students to acquire the skill of articulating the patient's problem and to develop an illness script based on two clinical scenario articles from the *New England Journal of Medicine's* Clinical Problem-solving series.[7,8] The purpose was to help students develop a succinct and coherent case presentation from the clinical problem presented in the journal article and to organize pertinent clinical data into a structured template (illness script) to formulate the basis of reaching a diagnosis. We chose the articles based on the complexity of the scenario and the perceived interest to students.

A PowerPoint presentation of the clinical scenarios was prepared by one faculty member before the workshop. Three other physicians from the Faculty of Medicine acted as tutors for the interactive sessions.

The students were given the details of the clinical presentation, history, and examination results during the workshop. The tutor's task was to guide the students to identify the important findings and to help them develop a presentation of the patients' problems using the teaching strategies described by Bowen.[6] Students were expected to characterize the clinical problem presented in the two articles as "an elderly man with a persistent cough"[7] and "a middle-aged woman with an acute swollen and painful left leg."[8]

The tutor then reasoned aloud to compare and contrast his expert problem representation with those of the students. Students then developed an illness script by listing the enabling (predisposing conditions or risk factors), fault (pathophysiological insult), and clinical consequences (key signs, symptoms, complaints) with feedback from the tutor. A sample problem illness script for pertussis in "an elderly man with a persistent cough"[7] is shown in Appendix 1.

Developing and Selecting an Appropriate Illness Script

The objective of this session was for students to acquire the skill of developing and selecting an appropriate illness script in formulating the most probable diagnosis for a given clinical problem. The purpose was to help students prioritize multiple diagnoses by identifying discriminating features for each diagnostic consideration. We developed a Web-based program that was comprised of a set of 20 clinical reasoning problems (CRPs) for teaching, combined with a scoring system for assessment. Our Web-based program was based on previously validated paper-based CRPs developed by Groves et al in conjunction with 22 family medicine physicians at the University of Queensland.[9] Details of the scoring system for CRP are described elsewhere.[9] The reliability of the paper-based CRP for medical students ranged from 0.61 to 0.83.[9]

Prior to the implementation of the workshop, we tested the Web-based program on eight final-year students who were not involved in the study, to identify computing errors in the program and scoring system. The Web-based CRPs for post-workshop assessment had acceptable reliability (Cronbach's alpha=0.80).

For each CRP, students had to nominate the two most likely diagnoses, list the clinical features that supported or opposed each diagnosis, and weigh each of the clinical features (slightly relevant, somewhat relevant, very relevant) from a series of drop-down menus. For the first CRP, two tutors interacted with students by asking them to list all important findings from the case, create a problem representation based on those findings, generate and prioritize diagnostic considerations that identify discriminating features for each consideration, identify findings from the case to support the diagnosis, and lastly to identify and compare alternative diagnoses.[6] Students were assisted if necessary at each step to reach reasonable conclusions and, when appropriate, positive feedback for correct responses was provided. This was then supplemented by a tutor reasoning aloud to summarize the process and illustrate key components such as the relevance of supporting clinical features associated with the correct diagnoses.

Students were encouraged to gradually work on four other CRPs individually during the session, with individual feedback from the tutors as they reasoned aloud. Thus, students were developing their own illness scripts for each CRP attempted. At the end of each completed CRP, the student was given an automated score by the program that gave immediate quantitative feedback on their clinical reasoning skills. Toward the end of the session, the tutors reasoned aloud for the selected CRPs to illustrate the gap between actual and desired expert performance. A sample CRP is shown in Appendix 2. The overall score for CRPs from the responses shown in the tables was 28/37, as visceral

trauma was considered to be a more likely diagnosis than gastroenteritis.

Evaluation Methods

Our evaluation methods were approved by The Chinese University of Hong Kong Survey and Behavioural Research Ethics Committee, Hong Kong. All students gave written informed consent and were assured that their scores would not be used in any part of their formal end-of-year assessment.

The overall study design and participant group assignment are shown in Figure 1. The participants were students completing the family medicine (intervention) and psychiatry (control) rotations toward the end of their fourth year. Students were originally allocated to either family medicine or psychiatry by the Faculty of Medicine office at the beginning of the academic year by a computer-generated random table. The students in the control group had previously finished their family medicine rotation before their psychiatry rotation; they were given no illness script-related material during their psychiatry rotation. Both pre-workshop and post-workshop assessments were held outside scheduled curriculum time, and participants were informed that they would receive a small honorarium if they completed both assessments.

Two weeks before the workshop, students were invited to participate in a study "to help develop and refine a Web-based teaching package for medical students aimed at improving their way of thinking in solving clinical problems." The explanation of the study, as communicated to the students, was purposely nonspecific so that they could not otherwise prepare for the new material.

Instruments

To provide different perspectives of the student's diagnostic skills ability after the intervention, two methods of assessment were chosen—the previously validated Diagnostic Thinking Inventory (DTI)[10] and a Web-based series of a previously validated CRP assessment tool.[9] A week before the workshop, both groups completed the DTI[10] within 20 minutes. At assessment, 8 weeks after the workshop, both groups completed the DTI and a set of 10 CRPs under examination conditions for up to 2 hours. We also assessed students' satisfaction with the educational intervention.

DTI

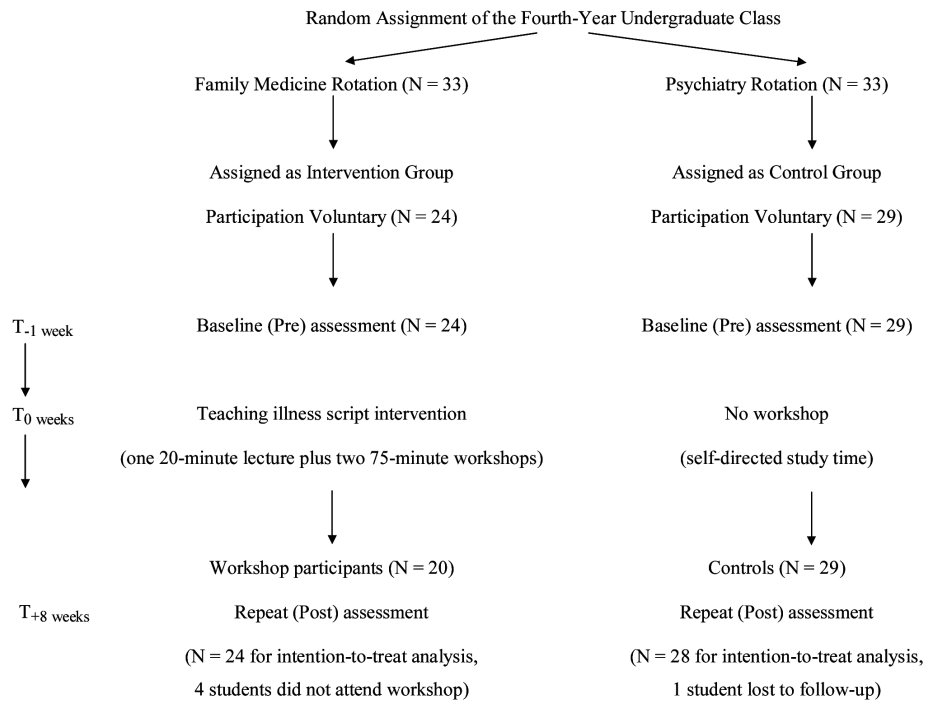
The DTI probes the subject's clinical reasoning style and attitude, has acceptable reliability (Cronbach's alpha=0.83),[10] and the ability to discriminate between degrees of diagnostic expertise at both the group and individual levels.[9] The DTI has two subscales: (1) flexibility in thinking (extent to which means or processes can be applied during the diagnostic process) and (2) structure in memory (availability of knowledge, stored in memory during the diagnostic process).[11] There is a moderate correlation between DTI and CRP scores,[9] implying that different aspects of clinical reasoning are assessed. The DTI focuses only on the process of diagnostic thinking since there is no knowledge measurement component.

CRP

The CRP score measures knowledge directly and also rewards the progressive integration of knowledge and process. The interface for the Web-based CRP as-

Figure 1

Study Design Used to Evaluate the Teaching Intervention for Clinical Reasoning in 53 Undergraduate Fourth-year Medical Students



sessment tool was different from that used in the tutorial to reduce any potential advantage that the intervention group might have gained from a better understanding of the interface. A drop-down menu for diagnosis included all the possible answers for 10 CRPs among several additional plausible, but incorrect, diagnoses. There was no drop-down menu for listing the supporting clinical features, and students were encouraged to type accurately or to cut and paste the text from the scenarios where relevant and to use a spell-check feature that was added to the program to match the program's predefined answers with that of the students' answers. The CRPs used during the workshop were not used for the assessment; a different set of 10 CRPs were used, with the maximum score set at 353.

Satisfaction

Student satisfaction was assessed through an anonymous 10-item written questionnaire administered at the end of the workshop. For each item, students rated the statement using a Likert scale (1=strongly disagree, 6=strongly agree).

Data Analysis

An intention-to-treat analysis was performed on all results (Figure 1). Multiple linear regressions were performed with adjustment to the repeated measures of DTI. Student's *t* test was used to compare CRP performance between groups. Stata Version 10.0 statistical software (StataCorp, College Station, TX) was used for all analyses.

Table 1

Comparison of DTI and Overall CRP Scores Between Fourth-year Medical Students Attending and Not Attending (Controls) a Workshop on Illness Scripts Using Intention-to-Treat Analyses

	Workshop Attendees (n=24)	***Controls (n=29)***
	Mean score (SD)	*Mean score (SD)*
Baseline		
Flexibility subscale	84.3 (10.7)	81.6 (11.8)
Structure subscale	77.7 (11.6)	79.8 (12.4)
Total DTI score	161.9 (20.7)	161.4 (21.7)
Follow-up*		
Flexibility subscale	83.1 (8.7)	83.2 (13.1)
Structure subscale	79.2 (8.1)	80.1 (9.0)
Total DTI score	162.3 (16.0)	163.3 (20.0)
Total CRP score	153.2 (36.7)	102.7 (43.7)

DTI—Diagnostic Thinking Inventory
CRP—Clinical Reasoning Problem

One student in the control group was lost to follow-up

Results

Of the 53 participating students, 26 (49%) were female. There was no difference in the participation rate of students in the intervention and control groups (P=.12). Also, there was no difference in baseline scores between those who did and did not complete the post-workshop assessment (P=.73).

The mean DTI scores in the two groups are shown in Table 1. The percentage of maximum DTI post-score in both groups was similar (both 66%). There was no significant difference between the intervention and control groups for post-flexibility subscale (mean difference 0, 95% CI= -5.1 to 5.0), post-structure subscale (mean difference 0.2, 95% CI= -3.5 to 3.9), or overall DTI score (mean difference 0, 95% CI= -7.4 to 7.4) after adjusting for baseline scores.

However, students attending the workshop scored significantly higher in the overall CRP assessment score than did students in the control group (mean difference 50, 95% CI=28 to 73) (Table 1). The percentage of maximum CRP scores in the workshop attendees and controls were 43% and 29% respectively, equivalent to an improvement of 14% (95% CI=8% to 21%). The effect size was large (1.24). There was no difference in overall CRP score between male (135 ± 46) and female (117 ± 48) students (P=.19).

Our satisfaction measure indicated that the intervention was generally well received by students (Table 2). When prompted for their views on the worst aspects of the course, concerns were expressed about timing (held outside of curriculum time, n=3) and brevity of the workshop (n=4). The best aspects of the workshop were reported to be: small-group teaching with interactive discussion (n=2), Web-based program (n=2), and relevant and practical sessions (n=2). The main suggestion for improvement was that the workshop should be of longer duration (n=4). Although the tutors' feedback was not formally assessed because of the small number of tutors involved, the overall feedback from them was positive.

Discussion

While there is a strong and continuing emphasis on understanding and describing the clinical reasoning process, reports of substantial and successful educational interventions to improve clinical reasoning are sparse. Despite the relative brevity of the intervention, our course appeared to be effective in improving diagnostic performance in simulated clinical reasoning skills problems, as measured by the CRP, and was generally well received by students. Given that the post-assessment was 8 weeks after the workshop and that

Table 2

Results of the Fourth-year Medical Students' Ratings of the Workshop on Illness Scripts

Statement	*Median (Interquartile Range)*
I gained a good understanding of concepts/principles in this field.	5 (5–5)
It deepened my interest in the subject matter of this workshop.	5 (5–6)
The objectives of the workshop were clearly defined.	5 (5–6)
The amount of material covered in the workshop was reasonable.	5 (5–6)
The level of difficulty of the workshop material was appropriate.	5 (5–6)
The "thinking aloud" approach was helpful for improving my clinical reasoning skill.	5 (5–5)
The illness script model was useful for improving my clinical reasoning skill.	5 (5–5)
Completing the Web-based clinical scenarios helped improved my clinical reasoning.	5 (5–6)
Attending a workshop to improve clinical reasoning would be time wasted in the medical curriculum.	2 (1–2)
Overall, I am satisfied with the course.	**5 (5–6)**

Rating scale for each statement: 1=strongly disagree, 2=disagree, 3=slightly disagree, 4=slightly agree, 5=agree, and 6=strongly agree.

the CRP covered distinctly different clinical content, the workshop appears to have provided students with a useful and effective approach to clinical reasoning rather than simply adding knowledge content for the particular cases included in the learning session.

There was a significant educationally important improvement in the overall CRP score following the workshop on illness scripts, with improvements up to 21% of the maximum score. Compared with the previous development and validation study in Australia,[9] a low overall CRP score in the control group was apparent and provides some evidence to support the perception that our undergraduate medical students in the body-system-based curriculum have relatively poor clinical reasoning skills immediately prior to entry into their fifth year of study. It is remarkable, then, that the overall CRP score from the intervention group was much improved and comparable to previously published data from second- and third-year medical students in a problem-based learning curriculum at the University of Queensland.[9] Although the student year number is different, the different systems in the two institutions would place these students at a similar level of seniority—the University of Queensland has a system of admitting graduates to their medical course that is 4 years in duration. Our students enter directly from high school and complete a 5-year curriculum.

On the other hand, the overall DTI scores of our students at baseline were broadly comparable to second- and third-year medical students in the development and validation study.[9] We detected little change in the overall post-intervention DTI score between groups. The DTI is a self-reporting instrument originally designed for use with clinicians, and its validity depends on respondents having sufficient clinical experience to develop a clear idea of how their own clinical reasoning process works in practice. It may be that this degree of self-awareness was not sufficiently developed in our medical students. Another possible explanation for the lack of change in DTI scores between groups is that the way knowledge is structured and accessed in memory is highly idiosyncratic. Therefore, while the intervention provided new knowledge and one approach to the clinical reasoning process, the DTI scores would be likely only to reflect the internalization of a long-term process in which students use their clinical experience to create an awareness of their knowledge organization and recognize a refinement of their process for accessing and using it during a clinical encounter. Interestingly, Beullens et al showed that a series of 70 problem-solving clinical seminars over 2 months was associated with significantly higher DTI Structure subscale score but not with the DTI Flexibility subscale score.[11]

Limitations

Our study relied on voluntary participation by students, thus raising the possibility of selection bias. To counter this possibility, the students were effectively randomized to the intervention or control groups. Further, the groups had comparable baseline characteristics, further decreasing the chance of selection bias between the two groups. For several logistical reasons, however, we were unable to administer the CRP to all students at baseline, so we cannot be certain that the groups did not differ in reasoning skills at baseline.

A second limitation is that this was a single-center study with a small number of participants. It is not possible, therefore, to confidently generalize the results of our study to other settings.

Conclusions

The key elements in the intervention involved the use of integrated small-group teaching, individual feedback to students when developing appropriate illness scripts, highly motivated tutors who reasoned aloud to students to highlight the desired performance, and the development of a Web-based set of clinical reasoning problems for teaching and assessments. Our initial results are encouraging. They can serve as a rationale for implementation of a one-morning clinical reasoning module to improving clinical diagnostic skills of medical students.

Acknowledgments: The work described in this paper is supported by a grant from the University Grants Committee Teaching Development Triennium Grant of the Hong Kong Special Administrative Region, China (Project No. 4170255).

This work was presented at The Chinese University of Hong Kong, Faculty of Medicine Curriculum Day meeting in September 2008 in Hong Kong.

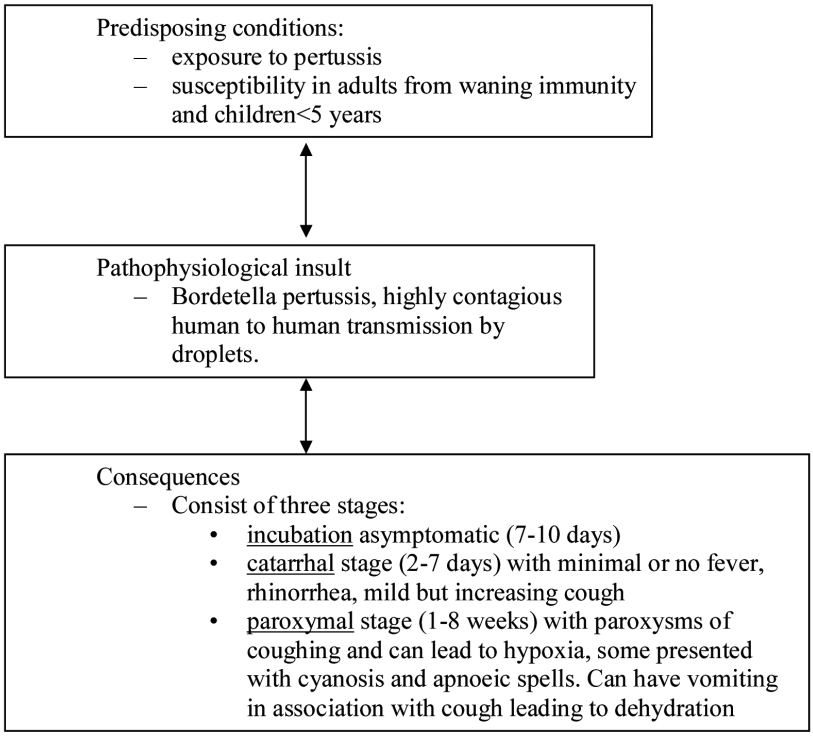
Corresponding Author: Address correspondence to Dr Lee, The Chinese University of Hong Kong, Prince of Wales Hospital, Department of Anaesthesia and Intensive Care, Shatin, NT, Hong Kong, China. +852-2632-2735. Fax: +852-2637-2422. annalee@cuhk.edu.hk.

References

1. Norman G. Research in clinical reasoning: past history and current trends. Med Educ 2005;39(4):418-27.
2. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. Acad Med 1990;65(10):611-21.
3. Barrows HS, Feltovich PJ. The clinical reasoning process. Med Educ 1987;21(2):86-91.
4. Norman G. Building on experience—the development of clinical reasoning. N Engl J Med 2006;355(21):2251-2.
5. Charlin B, Tardif J, Boshuizen HP. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. Acad Med 2000;75(2):182-90.
6. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. N Engl J Med 2006;355(21):2217-25.
7. Cornia PB, Lipsky BA, Saint S, Gonzales R. Clinical problem-solving. Nothing to cough at—a 73-year-old man presented to the emergency department with a 4-day history of nonproductive cough that worsened at night. N Engl J Med 2007;357(14):1432-7.
8. Fazel R, Froehlich JB, Williams DM, Saint S, Nallamothu BK. Clinical problem-solving. A sinister development—a 35-year-old woman presented to the emergency department with a 2-day history of progressive swelling and pain in her left leg, without antecedent trauma. N Engl J Med 2007;357(1):53-9.
9. Groves M, Scott I, Alexander H. Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. Med Teach 2002;24(5):507-15.
10. Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. Med Educ 1990;24(5):413-25.
11. Beullens J, Struyf E, Van DB. Diagnostic ability in relation to clinical seminars and extended matching questions examinations. Med Educ 2006;40(12):1173-9.

Appendix 1

Sample Problem Illness Script for Pertussis in "An Elderly Man With a Persistent Cough"[7]*



* The illness script forms the logical construct underlying the symptoms and signs making up the recognizable patterns for making a diagnosis.

Appendix 2

Sample Clinical Reasoning Problem From the Web-based Program

Margaret P. arrives in the A&E department of a public hospital at 9 o'clock one Sunday morning, accompanied by her 8 year old son, Joshua. Joshua is complaining of stomach ache and nausea. His mother says that he first complained of a colicky pain the previous night after dinner, but that now it has become much worse and has developed into a steady ache. He is also feverish. She adds that Joshua had been playing rugby for his local club yesterday, but had had to come off the ground after he was winded in a tackle. She now wonders whether the two events might be connected.

Other than giardiasis when he was at pre-school, Joshua has no significant medical history and is on no medication currently.

Joshua has a temperature of 38.1C. He complains of feeling cold. Abdominal examination reveals diffuse tenderness with guarding, but no distension. There is a large bruise immediately below his sternum and several smaller bruises on his legs and arms.

1. What do you think is the most likely diagnosis in this patient?

Appendicitis

2. Please list the features of the case which you consider support your diagnosis and also those which oppose it, giving an appropriate sign [positive (+) or negative (-)] and weighting to each. Please answer each feature in a separate row.

Feature	Supports (+) or Opposes (-)	Weighting 1: slightly relevant 2: somewhat relevant 3: very relevant	
Abdominal tenderness	+	2	delete
Guarding	+	3	delete
Fever	+	3	delete
Initial colicky pain	+	2	delete
Pain is increased	+	3	delete

Click to Add New Feature

3. If this diagnosis proved incorrect, what would your next choice be?

Gastroenteritis

4. Please list the features of the case which you consider support your diagnosis and also those which oppose it, giving an appropriate sign [positive (+) or negative (-)] and weighting to each. Please answer each feature in a separate row

Feature	Supports (+) or Opposes (-)	Weighting 1: slightly relevant 2: somewhat relevant 3: very relevant	
Nausea	+	2	delete
Fever	+	3	delete
Initial colicky pain	+	3	delete

Click to Add New Feature