Interdisciplinary Journal of Problem-Based Learning

Volume 13 | Issue 1 Article 5

Published online: 1-11-2019

Developing Clinical Reasoning Skills Through Argumentation With the Concept Map Method in Medical Problem-Based Learning

Jihyun Si
Dong-A University, jennyhan0306@gmail.com

Hyun-Hee Kong
Dong-A University

Sang-Hwa Lee
Dong-A University

Recommended Citation

Si, J. , Kong, H. , & Lee, S. (2019). Developing Clinical Reasoning Skills Through Argumentation With the Concept Map Method in Medical Problem-Based Learning. *Interdisciplinary Journal of Problem-Based Learning, 13*(1).
Available at: https://doi.org/10.7771/1541-5015.1776

The Interdisciplinary Journal of
Problem-based Learning

ARTICLE

Developing Clinical Reasoning Skills Through Argumentation With the Concept Map Method in Medical Problem-Based Learning

Jihyun Si, Hyun-Hee Kong, and Sang-Hwa Lee

Abstract

This study aims to explore the effects of argumentation with the concept map method during medical problem-based learning (PBL) on individual clinical reasoning. Individual clinical reasoning ability was assessed through problem-solving performance and arguments that students constructed during individual clinical reasoning processes. Toulmin's model of argument was utilized as a structure for arguments. The study also explored whether there would be any differences between the first- and second-year medical students. Ninety-five medical students participated in this study, and they took two PBL modules. During PBL, they were asked as a group to construct concept maps based on their argumentation about a case under discussion. Before and after each PBL, they were asked to write individual clinical problem-solving tests. One-way, within-subjects ANOVAs were conducted to examine the quality of arguments and clinical problem-solving performance in three individual tests. The results provided evidence that utilizing argumentation with the concept map method during PBL positively affects the development of clinical reasoning skills by individual students.

Keywords: clinical reasoning, problem-based learning, argumentation, concept map, Toulmin's model of argument

Introduction

Clinical Reasoning Skills and Problem-Based Learning

Clinical reasoning processes are the problem-solving process used by doctors (Barrows & Feltovich, 1987). Clinical reasoning is a central component of doctors' competence, so developing students' clinical reasoning skills is a central goal in medical education. Problem-based learning (PBL) as a teaching approach was introduced to foster clinical reasoning skills, particularly hypothetico-deductive reasoning skills based on clinical problems similar to real practice (Barrows & Feltovich, 1987; Patel, Arocha, & Zhang, 2005; Rochmawati & Wiechula, 2010). Hypothetico-deductive reasoning is known to be used especially by novices with limited knowledge of patient disease (Groves, 2012; Harasym, Tsai, & Hemmati, 2008; Patel et al., 2005). However, despite its wide acceptance, evidence is lacking on how effective PBL is in fostering clinical reasoning skills (Kong, Qin, Zhou, Mou, & Gao, 2014; Rochmawati & Wiechula, 2010; Taylor & Miflin, 2008; Yuan, Williama, & Fan, 2008).

In PBL, clinical reasoning skills are often assessed with clinical problem-solving performance through such things as written tests or concept map approaches (Kassab et al., 2016; Kassab & Hussain, 2010; Kreiter & Bergus, 2009), under the premise that clinical problem-solving performance reflects clinical reasoning skills (Harasym et al., 2008; Kreiter & Bergus, 2009; Wu, Wang, Grotzer, Liu, & Johnson, 2016). However, these assessments mainly focused on knowledge structures or products that students produce as a result of the instructional interventions to enhance students' clinical reasoning skills. In these assessments, as long as the right answer is provided, reasoning is considered acceptable, and no further regard is paid to how the answer is obtained.

http://dx.doi.org/10.7771/1541-5015.1776

Clinical Reasoning Skills and Argumentation

Clinical reasoning skills require a dynamic interaction between medical knowledge structures, thinking strategies, and clinical experiences (Groves, 2012; Harasym et al., 2008), which implies that it is reasonable to use multiple forms of assessment to evaluate clinical reasoning skills. Jonassen (2011, p. 354) insists that "adequate assessment of problem-solving skills requires more than one form of assessment" and that problem-solving skills can be assessed in many ways, including problem-solving performance and ability to construct arguments in support of the solutions to problems (Jonassen, 2011). Argumentation is a process of making claims and providing explanation for the claims using evidence (Siegel, 1995; Toulmin, 1958). In PBL, students start to learn with a clinical problem that stimulates their learning process in small collaborative groups (Rochmawati & Wiechula, 2010). Clinical problems that usually begin with the symptoms of a sick person do not have all the information available at the outset (Barrows & Feltovich, 1987). Students must formulate multiple hypotheses, search for relevant information, and justify their decisions with that information. The approaches that lead to the solutions are generally not standardized (Oh & Jonassen, 2007). These processes involve argumentation. Thus, argumentation is essential for clinical problem solving in PBL (Jonassen & Kim, 2010).

Empirical evidence has also confirmed the close relationship between argumentation and problem-solving skills, especially in ill-structured problems (Cho & Jonassen, 2002; Ju, 2016). Clinical problems are regarded as ill-structured problems whose alternative solutions and interpretations necessitate argumentation (Barrows & Feltovich, 1987; Cho & Jonassen, 2002; Jonassen, 2011). Cho and Jonassen (2002) utilized computer- and constraint-based argumentation scaffolds in PBL and showed that better production of argumentation during PBL has a direct positive effect on individual problem-solving activities.

Argument is an essential component of clinical problem solving, so it provides critical evidence about students' problem-solving ability (Jonassen, 2011). Thus, arguments that students construct during clinical problem solving could provide evidence about their critical reasoning skills as well. In that sense, assessing arguments for thinking strategies and problem-solving performance for the knowledge structure obtained through PBL is thought to be a reasonable way to assess clinical reasoning skills in PBL.

Argumentation and Concept Maps

Argumentation can be the means by which students rationally solve clinical problems in PBL, but students are not often capable of constructing cogent argumentation (Cerbin, 1988; Cho & Jonassen, 2002; Ju, 2016; Kuhn & Udell, 2007; Simon, Erduran, & Osborne, 2006; Yeh, 1998). Ju (2016) analyzed how medical students engage in argumentation in PBL and found that they often omitted data or did not connect evidence via warrants.

Two different teaching approaches have been utilized in order to facilitate students' argumentation skills. Several researchers suggested that explicit instructions requiring particular patterns of argumentation such as Toulmin's model of argument enhanced students' argumentation skills (Saunders, 1994; Toulmin, Rieke, & Janik, 1984; Yeh, 1998). Toulmin's model shows a structure of argument that includes five major components such as claims, grounds, warrants, backings, and rebuttals (Toulmin et al., 1984). However, explicit instructions do not always improve students' argumentation skills (Knudson, 1991; Sanders, Wiseman, & Gass, 1994).

Another approach to support argumentation is to scaffold argumentation through visualizing arguments (Jonassen & Kim, 2010; Kirschner, Buckingham Shum, & Carr, 2003). Visualizing arguments has advantages in making students see the structure of an argument, thus facilitating its more coherent construction and subsequent communication among students (Jonassen, 2011). Among different forms of visual representation of arguments, concept map approaches are useful. Concept maps are tools for organizing and representing knowledge (Novak & Cañas, 2006). Their implementation involves new concepts enclosed in circles or boxes, creating hierarchical arrangements between concepts and subconcepts, and identifying relationships between concepts and subconcepts by a connecting line with a linking word on it (Figure 1, Novak & Cañas, 2006).

In fact, concept maps have been utilized as teaching and learning strategies to help students to represent and organize their knowledge in PBL (Charlin et al., 2012; Kassab et al., 2016; Kassab & Hussain, 2010; Rendas, Fonseca, & Pinto, 2006). However, concept mapping alone is found to be insufficient in supporting complex problem-solving processes, especially for eliciting students' reasoning process (Stoyanov & Kommers, 2008; Wu et al., 2016). The graphical convention in the instruction for drawing a concept map is not a sufficient condition for making concept mapping an effective tool in support of complex cognitive processes. Stoyanov and Kommers (2008) empirically demonstrated that the concept mapping group with explicit problem-solving support outperformed the concept mapping group without such support on problem-solving performance. It was a set of concrete instructions that led to significant differences in the problem-solving performance (Stoyanov & Kommers, 2008). Thus, the instruction on graphical conventions for drawing a concept map should be coupled with an instruction that includes a set of concrete instructions regarding complex

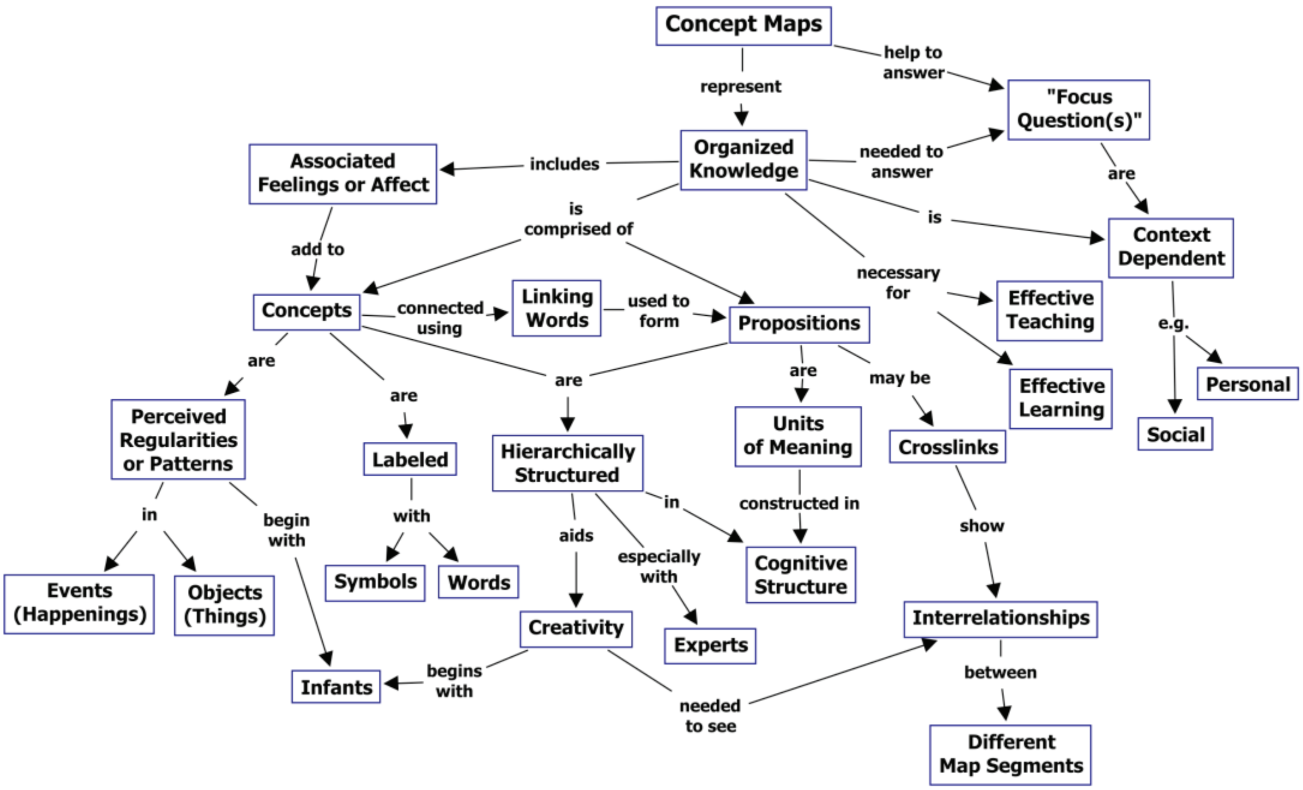

Figure 1. A concept map showing the key features of concept maps (cited from Novak & Cañas, 2006, with permission).

cognitive processes (Stoyanov & Kommers, 2008; Wu et al., 2016). Therefore, to facilitate students' argumentation, combining the concept map method with explicit instructions on the requirements of particular patterns of arguments is thought to be effective.

The ability to argue effectively for a different perspective is equivalent to solving clinical problems (Oh & Jonassen, 2007). However, very few studies examine the effects of argumentation on clinical problem solving, and there is not even research about the effects of argumentation with the concept map method on clinical problem solving in medical PBL.

Thus, for the purpose of developing medical students' clinical reasoning skills in PBL, this study aims to explore the effects of argumentation using the concept map method during PBL on individual clinical reasoning skills of medical students. Toulmin's model of argument will be utilized as a structure of argument because it is useful to teach students how to construct arguments as well as evaluate arguments. Individual clinical reasoning skills will be assessed with problem-solving performance and arguments students construct during individual problem solving after PBL. This study will also explore whether there are any differences between the first- and second-year medical students. The second-year medical students took typical PBL courses while they were in their first year without any scaffolding to support the clinical reasoning process, and the recall tests were administered after the PBL. Thus, comparing the two groups may provide important insights regarding the effect of argumentation using the concept map method on clinical reasoning skill development. The general research question and the specific research questions for this study are as follows:

How does the argumentation using the concept map method during PBL affect medical students' clinical reasoning skills?

1. How does the argumentation using the concept map method during PBL affect the production of medical students' arguments in individual clinical problem solving?
2. How does the argumentation using the concept map method during PBL affect medical students' problem-solving performance in individual clinical problem solving?
3. Are there any differences between the first-year and second-year students in terms of the production of arguments and clinical problem-solving performance?

Method

Participants

The subjects of this study were medical students enrolled in PBL courses at a medium-sized university in Korea during the fall semester of 2015. Two groups of medical students (first-year and second-year students) participated in this study. Among 49 first-year students, 5 students did not complete all three individual problem-solving tests, so they were excluded. In contrast, 51 second-year students finished three individual problem-solving tests. Thus, ultimately, data from 95 students were analyzed. Among the first-year participants, 29 students were male and 15 were female. Among the second-year students, 27 were male and 24 were female. The first-year students ranged from 21 to 33 years of age (M = 24, SD = 2.43), and the second-year students ranged from 23 to 37 years of age (M = 27, SD = 2.69). The students were randomly assigned to discussion groups of seven or eight in PBL classes. The first-year students had not previously worked in PBL classes, but the second-year students had worked in PBL classes on five previous occasions between the fall semester of 2014 and the spring semester of 2015. The PBL classes that the second-year students had previously experienced were the same as the PBL classes in this study except for the treatment of argumentation with the concept map method. That is, in this study, students constructed concept maps according to Toulmin's structure of arguments instead of just engaging in verbal discussion, presented their concept maps as a group to the class instead of just listening to a lecture about the case from a professor in the third meeting of PBL, and were assessed with their concept map presentations as a group for their PBL credit instead of taking a recall test.

PBL

One PBL module consisted of three meetings over three weeks (one meeting each week) in this study. In the first two meetings, seven to eight students worked together in one team under the guidance of a tutor for about 120 minutes at a time. Tutors played facilitator roles and encouraged students to get involved in discussions and build cogent arguments. Typical PBL sessions started with students being exposed to a patient's main symptoms. Through discussion in a small group, they first made use of their existing knowledge to generate multiple hypotheses, and then gathered relevant data from a tutor only when they had asked for specific data and identified issues for self-directed learning at the end of the first meeting. In the second meeting, they shared information they had gathered during the self-directed learning, reviewed their prior discussion based on new information, and made a decision about which hypothesis would be accepted as a final diagnosis. In doing so, the students were asked to construct concept maps about their argumentation based on Toulmin's model of argument. In the third meeting, all students gathered in one classroom with a professor who developed the PBL module discussed, and each group presented its concept map developed during the last two sessions. The professor gave feedback on their presentations and delivered a final mini-lecture. The concept map presentations were evaluated by a rubric assessing their reasoning processes and presentation styles and took 20% of their PBL credit.

Argumentation with the Concept Map Method

Students were asked to construct a concept map according to the structure of Toulmin's model of argument (Toulmin, 2003; Toulmin et al., 1984) in relation to a case the students had discussed during the PBL sessions. Toulmin's model of argument has been widely used to support the development of student argumentation (Erduran, Simon, & Osborne, 2004).

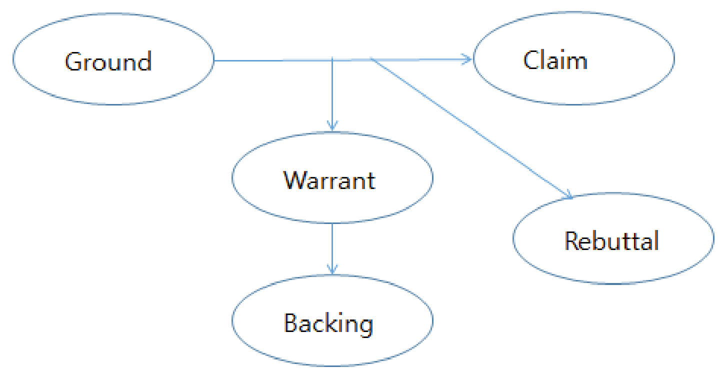

Figure 2. The structure of Toulmin's arguments (Toulmin, 1958).

The structure of argumentation by Toulmin includes claims, grounds, warrants, backings, and rebuttals (Toulmin, 1958; Toulmin et al., 1984) and helps students engage in specific forms of argumentation (Figure 2). That is, arguers justify their claims using supporting grounds through warrants that provide a link between grounds and the claim. Backings strengthen the warrants, and rebuttals point to the constraints that contradict claims (Jonassen & Kim, 2010; Toulmin et al., 1984).

Using the definition of "concept" by Novak and Cañas (2006) "as a perceived regularity in events or objects, or records of events or objects" (p. 1), we can regard each category of the argument as a concept. The students were asked to draw each category of argument in boxes and represent it in a hierarchical fashion. In addition, they were asked to connect concepts with lines and write a linking word on the line

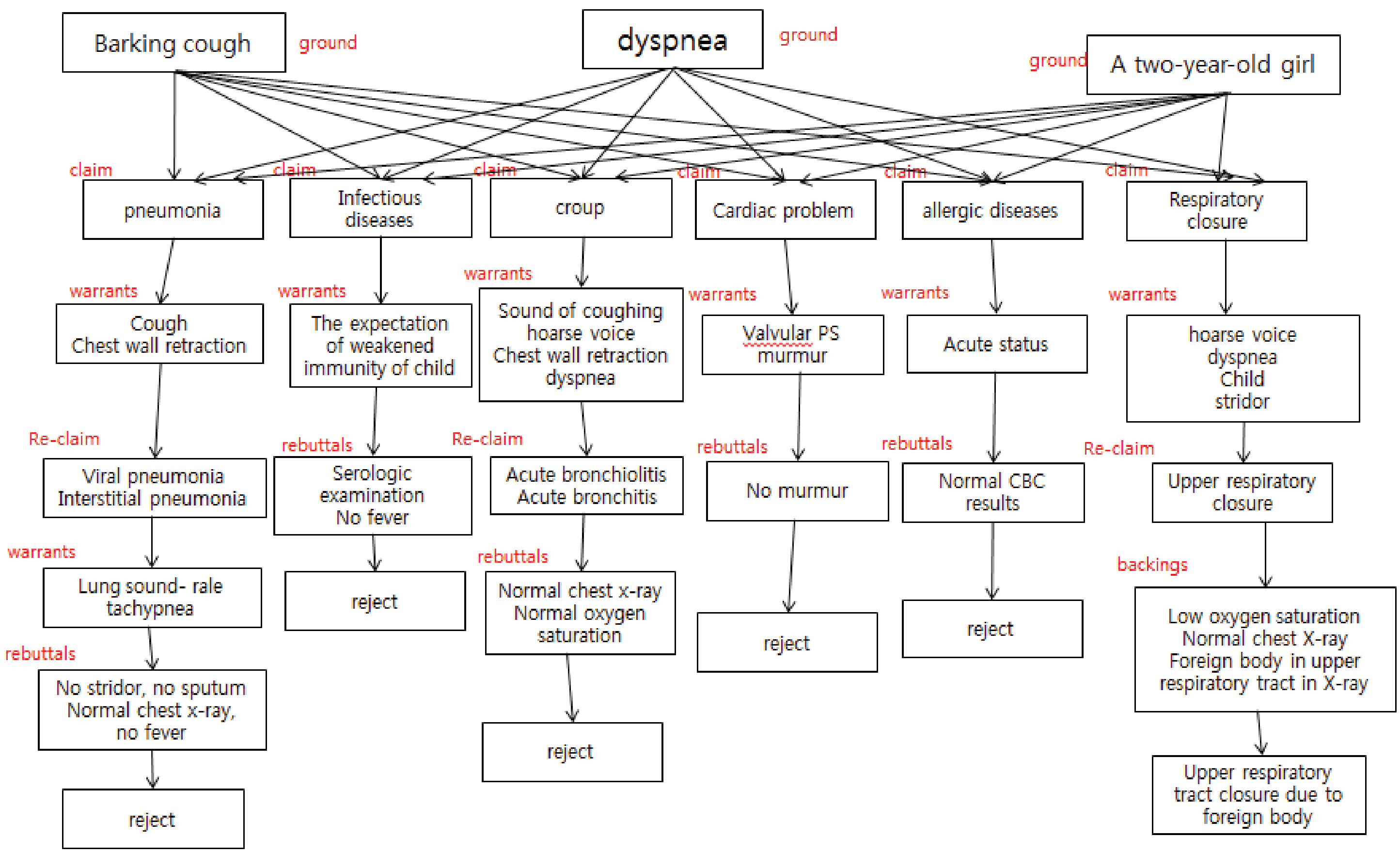

Figure 3. A concept map example constructed by the first-year students.

to specify the relationship between two concepts. Figure 3 is an example of a concept map that the first-year students constructed as a group during the first PBL session. The patient in this case was a two-year-old girl who came to an emergency room because of a barking cough, hoarseness, and difficulty breathing. The concept map showed students' reasoning processes through argumentation according to Toulmin's model of argument from generating multiple hypotheses or claims (pneumonia, infectious disease, croup, cardiac disorder, allergic diseases, and respiratory closure) based on the first data available when they met a patient to offering the final diagnosis, upper respiratory closure. Figure 4 is also an example of a concept map that the second-year students constructed during the second PBL session. It shows how they went from multiple hypotheses including hydrocephalus and the like to the final diagnosis, hypertrophic pyloric stenosis.

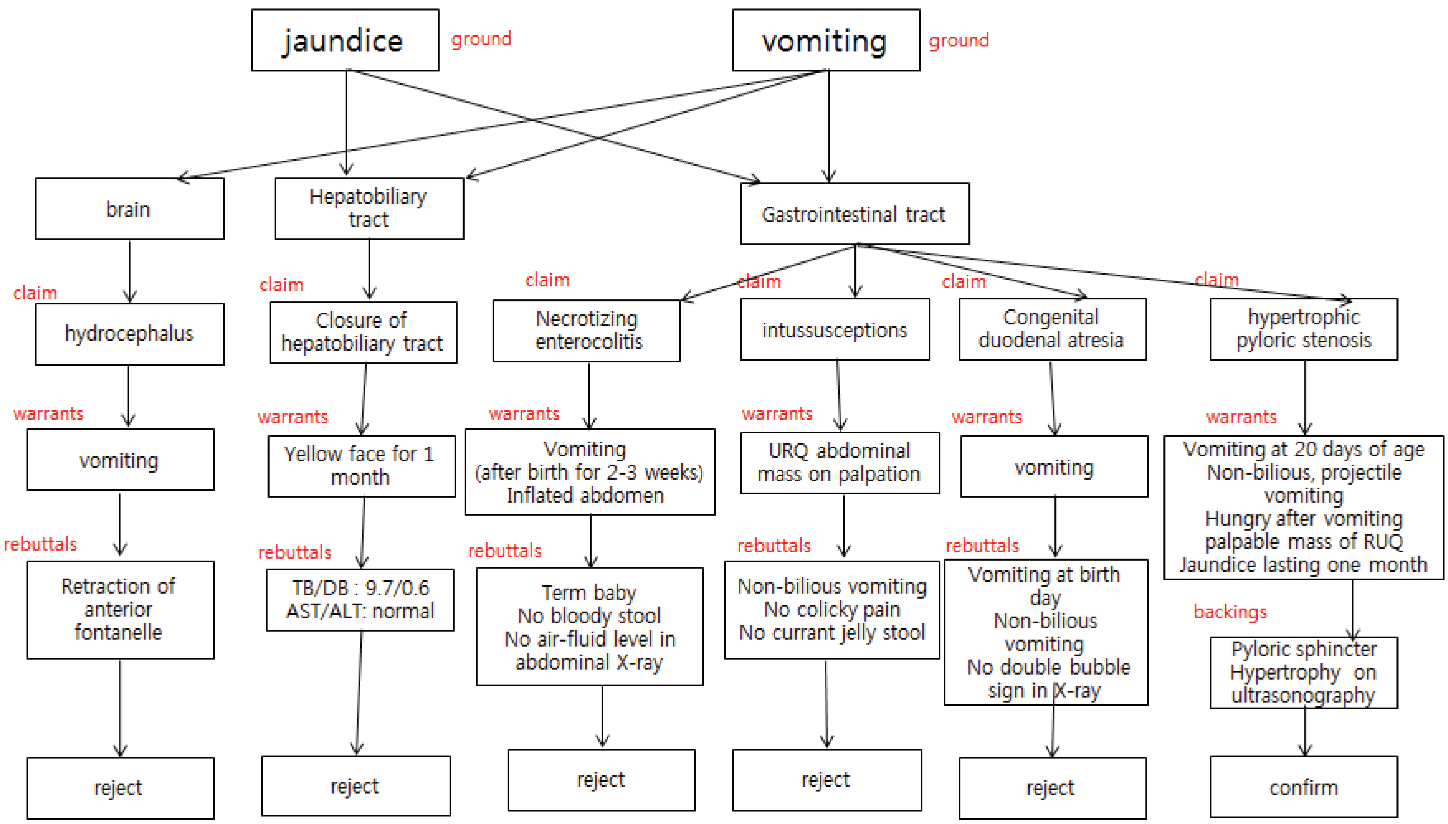

Figure 4. A concept map example constructed by the second-year students.

A 26-year-old man reported that his skin grew more and more yellow as his eyes had turned yellow over three days. Previously, he had been tired and felt nauseated for two weeks, and he did not take any medication for that. He had never had hepatitis before and was vaccinated against hepatitis B. He was currently taking no medicines or health foods. He had not traveled recently. There was a slight feeling of warmth from one week ago, and there was persistent discomfort in the upper part of the stomach and right upper stomach. He felt nausea but didn't vomit. He had lost about 5 kg of weight in 1–2 years. He worked for a credit card company. His current job situation had not been good, so he was stressed. He thought weight loss was due to stress as well as exercise and diet and that there was no particular problem. One year ago, he had had a medical check-up and heard that there was no abnormality. There was currently no particular medicine for him to take regularly. He drank 1/2 a bottle of soju about 1–2 times a week, did not smoke at all, and ate well. But for about a year, he had eaten less during dinner to lose weight. He walked 2–3 times a week for 30 minutes.

Figure 5. A case description of an individual problem-solving test.

Instruments

Individual Problem-Solving Tests

An individual problem-solving test was administered three times: before PBL, after the first PBL session, and after the second PBL session. These paper-and-pencil tests took 30 minutes each, and the students wrote individual essays regarding the case presented. Students solved a new clinical problem that was similar to the cases that they had solved during the PBL sessions. They were the cases used in the previous PBL classes. Students were asked to explain for individual problem-solving cases, using evidence found in a case description, what the possible diagnoses were, why they chose a certain diagnosis, what the alternative diagnoses were, why they were excluded, and what their diagnostic plan was to confirm their diagnosis. Figure 5 is an example of a case description of an individual problem-solving test for the second-year students.

Assessment of the Quality of Arguments

The individual problem-solving tests used as a measure of the quality of individual arguments were scored by two assessors without knowing the identity of the student, using the scoring rubric developed by Cho and Jonassen (2002) based on Toulmin's model of argument (Toulmin et al., 1984). Table 1 shows the rubric for assessing the quality of claims of student arguments (see Cho & Jonassen, 2002, for rubrics for other categories such as grounds, warrants, backings, and rebuttals). Individual scores were achieved by summing up numbers of points (0–30) earned in each argument category (claims, grounds, warrants, backings, and rebuttals). Pearson's r was calculated to assess interrater reliability of the scores. The interrater agreements for the quality of arguments were .81 (the first test), .86 (the second test), and .93 (the third test) for the first-year student tests, and .91 (the first test), .95 (the second test), and .82 (the third test) for the second-year student tests. The final scores were determined based on consensus between the two assessors.

Table 1. Rubric for assessing the quality of claims of arguments in individual students' problem solving (Cho & Jonassen 2002, p. 12)

Quality	Criteria
6	The writer states generalizations that are related to the proposition and that are clear and complete.
4	The writer states generalizations that are related to the proposition, but the assertions are not complete. Enough information is available to determine the writer's intent, but much is left to the reader to determine.
2	The writer makes generalizations that are related to the proposition, but the assertions lack specificity or offer unclear reference. The writer leaves much for the reader to infer in order to determine the impact of the claim.
0	No claim related to the proposition or unclear assertions.

Assessment of Clinical Problem-Solving Performance

Each individual problem-solving test was also analyzed for problem-solving performance by using a rubric. The quality of the problem-solving performance can be assessed using a rubric that describes desirable performance (Jonassen, 2011). The rubric for problem-solving performance assessed accuracy, relevance, completeness, and specificity in this study. In each individual problem-solving test, the assessors looked for students' claims (possible disease), supporting data (patients' symptoms), explanations about how the supporting data were relevant to the claims, other specific backings, and systematic identification of constraints of alternative diagnoses that had been ruled out. The assessors were all professors. They evaluated one-third of the students of each group in each test and a different student each time without knowing their identities. The scores were 6, 4, 2, and 0.

Table 2. The descriptive statistics and the F-test values of the quality of arguments of the second-year and first-year students

		Max.	Min.	M	SD	F	p
First-year Students (N = 44)	Pretest	14.85	9.79	12.32	8.32	13.18	p<.001
	Posttest 1	16.00	12.63	14.32	5.54		
	Posttest 2	21.71	17.74	19.73	6.54		
Second-year students (N = 51)	Pretest	16.41	12.85	14.63	6.33	40.98	p<.001
	Posttest 1	23.25	19.9	21.57	5.96		
	Posttest 2	24.67	21.13	22.90	6.29		

Procedures

The students registered in PBL took a PBL orientation session and received an explanation on argument categories and how to construct concept maps as a part of a PBL orientation session one week before the PBL sessions started in this study. A sample concept map was also provided for their better understanding of concept map construction. Tutors also took a PBL tutor orientation session. They received an explanation on how to construct concept maps and argument categories as part of the orientation program as well. These orientation sessions were mandatory for both students and tutors. After the orientation session, the students worked in two PBL sessions. The first PBL session took place in the beginning of the fall semester, and the second session took place five weeks later for both the first- and second-year students. One day before the first PBL session started, both the first- and second-year students took an individual problem-solving test. On the last day of the first PBL session and the second PBL session, they took an individual problem-solving test again after the PBL class. Thus, in total, they took three individual problem-solving tests. The cases given for the three tests were all different, and the three tests (pretest, posttest 1, and posttest 2) were administered by the researcher.

Analysis

One-way, within-subjects ANOVAs were conducted to explore students' quality of arguments and clinical problem-solving performance. To detect significant differences among the three tests (pretest, posttest 1, and posttest 2), adjusted Bonferroni post hoc comparisons were also conducted. The quality of each argument category (claim, ground, warrant, backing, and rebuttal) was analyzed with one-way, within-subjects ANOVAs and adjusted Bonferroni post hoc comparisons as well. For the comparison of the second- and first-year students, independent-sample t-tests were conducted. SPSS 23 was utilized for statistical analyses.

Results

The Quality of Individual Arguments

To examine whether the participants showed differences in the quality of arguments, the individual problem-solving tests of 95 medical students were analyzed (see Table 2). Among second-year students, one-way, within-subjects ANOVAs revealed a statistically significant difference among the three test scores with a large effect size, $F(2, 100) = 40.98$, $p < .001$ $\eta_p^2 = .45$. According to adjusted Bonferroni post hoc comparisons, the mean scores of posttest 1 and posttest 2 were significantly higher than the pretest. There was no significant difference between posttest 1 and posttest 2. In the case of the first-year students, Mauchly's test for sphericity was found to be significant, $\chi^2(2) = 16.51$, $p < .001$, so degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = .78$). The results of one-way, within-subjects ANOVAs for the first-year students also showed a statistically significant difference with a large effect size, $F(1.56, 66.76) = 13.18$, $p < .001$ $\eta_p^2 = .24$. Adjusted Bonferroni post hoc comparisons revealed that the mean scores of posttest 1 and posttest 2 were significantly higher than the pretest. The mean score of posttest 2 was also significantly higher than posttest 1.

The Quality of Each Category of Individual Arguments

Each category of arguments was also analyzed using one-way, within-subjects ANOVAs (Tables 3 and 4). In the case of the second-year students, in an examination of the claim scores, Mauchly's test for sphericity was found to be significant, $\chi^2(2) = 7.32$, $p < .05$, so the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .91). The results showed a statistically significant difference among the three claim scores with a large effect size, $F(1.82, 90.76) = 34.02$, $p < .001$ $\eta_p^2 = .41$. According to adjusted Bonferroni post hoc comparisons, the mean scores in posttest 1 and posttest 2 were significantly higher than in the pretest.

In an examination of the ground score, Mauchly's test for sphericity was also found to be significant, $\chi^2(2) = 10.73$, $p < .01$, so the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .86). The statistical results showed a significant difference among the three ground scores with a large effect size, $F(1.72, 86.12) = 56.82$, $p < .001$ $\eta_p^2 = .53$. According to adjusted Bonferroni post hoc comparisons, the mean scores in posttest 1 and posttest 2 were significantly higher than in the pretest.

In an examination of the warrant score, the results revealed a statistically significant difference with a large effect size, $F(2, 100) = 26.51$, $p < .001$ $\eta_p^2 = 35$. According to adjusted Bonferroni post hoc comparisons, the mean scores in posttest 1 and posttest 2 were significantly higher than in the pretest. In an examination of the backing score, the results revealed a statistically significant difference with a large effect size, $F(2, 100) = 8.70$, $p < .001$ $\eta_p^2 = 15$. According to adjusted Bonferroni post hoc comparisons, the mean scores in posttest 1 and posttest 2 were significantly higher than in the pretest. In an examination of the rebuttal scores, the results revealed a statistically significant difference with a large effect size, $F(2, 100) = 14.41$, $p < .001$ $\eta_p^2 = 22$. According to adjusted Bonferroni post hoc comparisons, only the mean score in posttest 2 was significantly higher than in the pretest. However, the mean score in posttest 2 was also significantly higher than in posttest 1.

In the case of the first-year students, in an examination of the claim score, Mauchly's test for sphericity was found to be significant, $\chi^2(2) = 13.77$, $p < .01$, so the degrees of

Table 3. The descriptive statistics and the F-test values of each category of arguments among the second-year students (N = 51)

		Max.	Min.	M	SD	F	p
Claims	Pretest	4.00	3.16	3.57	1.46	34.02	p<.001
	Posttest 1	5.59	5.00	5.29	1.05		
	Posttest 2	5.42	4.86	5.14	1.00		
Grounds	Pretest	4.20	3.41	3.80	1.4	56.82	p<.001
	Posttest 1	5.75	5.15	5.45	1.01		
	Posttest 2	5.80	5.33	5.57	0.83		
Warrants	Pretest	3.69	2.82	3.25	1.55	26.51	p<.001
	Posttest 1	5.29	4.44	4.86	1.51		
	Posttest 2	5.42	4.46	4.94	1.71		
Backings	Pretest	2.75	1.80	2.27	1.70	8.70	p<.001
	Posttest 1	4.08	2.90	3.49	2.11		
	Posttest 2	4.36	3.02	3.69	2.38		
Rebuttals	Pretest	2.23	1.22	1.73	1.79	14.41	p<.001
	Posttest 1	3.08	1.86	2.47	2.18		
	Posttest 2	4.23	2.91	3.57	2.34		

Table 4. The descriptive statistics and the F-test values of each category of arguments among the first-year students (N = 44)

		Max.	Min.	M	SD	F	p
Claims	Pretest	3.53	2.47	3.00	1.75	9.39	p<.001
	Posttest 1	4.49	3.42	3.96	1.75		
	Posttest 2	5.03	4.07	4.54	1.58		
Grounds	Pretest	3.55	2.55	3.05	1.64	7.23	p<.001
	Posttest 1	3.56	2.71	3.14	1.39		
	Posttest 2	4.60	3.68	4.14	1.52		
Warrants	Pretest	3.15	1.94	2.55	2.00	7.72	p<.001
	Posttest 1	3.17	2.46	2.82	1.17		
	Posttest 2	4.34	3.38	3.86	1.58		
Backings	Pretest	2.69	1.49	2.09	1.97	4.27	p<.05
	Posttest 1	2.65	1.81	2.23	1.28		
	Posttest 2	3.68	2.59	3.14	1.80		
Rebuttals	Pretest	2.26	1.02	1.64	2.04	19.30	p<.001
	Posttest 1	2.85	1.52	2.18	2.19		
	Posttest 2	4.56	3.53	4.05	1.70		

freedom were corrected using Huynh-Feldt estimates of sphericity ($\boldsymbol{\varepsilon}$ = .81). The results of one-way, within-subjects ANOVAs revealed a significant difference with a large effect size, F (1.61, 69.30) = 9.39, p < .001 η_p^2 = .18. According to adjusted Bonferroni post hoc comparisons, the mean score in posttest 2 was significantly higher than in the pretest. In an examination of the ground scores, Mauchly's test for sphericity was also found to be significant, χ^2 (2) = 8.96, p < .05, so the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\boldsymbol{\varepsilon}$ = .87). The results showed a statistically significant difference with a large effect size, F (1.74, 74.74) = 7.23, p < .001 η_p^2 = .14. According to adjusted Bonferroni post hoc comparisons, the mean score in posttest 2 was significantly higher than in the pretest and posttest 1. In an examination of the warrants score, Mauchly's test for sphericity was found to be significant, χ^2 (2) = 15.26, p < .001, so the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\boldsymbol{\varepsilon}$ = .79). The results showed a statistically significant difference with a large effect size, F(1.58, 67.87) = 7.72, p < .001 η_p^2 = .15. According to adjusted Bonferroni post hoc comparisons, the mean score in posttest 2 was significantly higher than in the pretest and posttest 1. In an examination of the backing scores, the results revealed a statistically significant difference with a medium effect size, F(2, 86) = 4.27, p < .05 η_p^2 = .09. According to adjusted Bonferroni post hoc comparisons, the mean score in posttest 2 was significantly higher than in the pretest. In an examination of the rebuttal scores, the results revealed a statistically significant difference with a large effect size, F(2, 86) = 19.30, p < .001 η_p^2 = .31. According to adjusted Bonferroni post hoc comparisons, the mean score in posttest 2 was significantly higher than in the pretest and posttest 1.

Problem-Solving Performance

The descriptive statistics and the F-test values of one-way, within-subjects ANOVAs of problem-solving performance are summarized in Table 5. One-way, within-subjects ANOVAs revealed a statistically significant difference among three test scores of the second-year students with a large effect size, $F(2, 100) = 14.44$, $p < .001$ $\eta_p^2 = .22$. According to adjusted Bonferroni post hoc comparisons, the mean score in posttest 1 was significantly higher than in the pretest. There was also significant difference in the performance scores between posttest 1 and posttest 2. In an examination of the scores of the first-year students, Mauchly's test for sphericity was found to be significant, $\chi^2(2) = 9.37$, $p < .01$, so the degree of freedom was corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = .86$). The results also revealed a statistically significant difference with a large effect size, $F(1.73, 74.21) = 8.04$, $p < .001$ $\eta_p^2 = .15$. Adjusted Bonferroni post hoc comparisons revealed that the mean score in posttest 2 was significantly higher than in the pretest and posttest 1.

Comparisons Between First-Year and Second-Year Students

For the comparisons of two groups in terms of the quality of arguments, independent-sample t-tests were conducted. According to the results, there was no significant difference in the pretest, but the second-year students showed significantly higher mean scores than the first-year students in posttest 1 ($t(93) = -6.11$, $p < .001$, $d = 1.26$) and posttest 2 ($t(93) = -2.41$, $p < .05$, $d = 0.49$).

In each category of arguments, first, regarding the claim scores, the assumption of equal variance was not met, so Welch's t-test was used both in posttest 1 and posttest 2. In the pretest, there was no significant difference, but the second-year students showed significantly higher mean scores than the first-year students in posttest 1 ($t(50) = -22.51$, $p < .001$, $d = -0.97$) and posttest 2 ($t(71) = -2.15$ $p < .05$, $d = 0.45$). In the ground scores, the second-year students showed significantly higher mean scores than the first-year students in the pretest ($t(85) = -2.40$, $p < .05$, $d = 0.49$), posttest 1 ($t(69) = -4.94$, $p < .001$, $d = -1.04$), and posttest 2 ($t(64) = -5.58$, $p < .001$, $d = 1.17$). In the remaining three tests, as the assumption of equal variance was not met, Welch's t-test was used. In terms of the warrant scores, in the pretest, there was no significant difference, but the second-year students showed significantly higher mean scores than the first-year students in posttest 1 ($t(93) = -5.76$, $p < .001$, $d = 1.18$) and posttest 2 ($t(93) = -3.17$, $p < .001$, $d = 0.66$). In terms of the backing and rebuttal scores, there were no significant differences between the two groups. Regarding the problem-solving performance, the two groups showed a significant difference only in posttest 1 ($t(93) = -5.73$, $p < .001$, $d = 1.18$).

Table 5. The descriptive statistics and the F-test values of the problem-solving performance of the first- and second-year students

		Max.	Min.	M	SD	F	p
First-year students (N = 44)	Pretest	3.61	2.58	3.09	1.70	8.04	p<.001
	Posttest 1	3.60	2.85	3.23	1.24		
	Posttest 2	4.59	3.77	4.18	1.35		
Second-year students (N = 51)	Pretest	3.93	3.13	3.53	1.42	14.44	p<.001
	Posttest 1	5.18	4.40	4.78	1.35		
	Posttest 2	4.40	3.68	4.04	1.30		

Discussion

This study examined the effects of the argumentation with the concept map method during PBL on individual clinical reasoning skills. Specifically, it explored how the argumentation with the concept map method affects the production of medical students' arguments and problem-solving performance in individual clinical reasoning processes. It was expected that the educational intervention would have transfer effects, resulting in enhancement of their reasoning skills measured by the quality of arguments and problem-solving performance.

As expected, use of argumentation with the concept map method positively affected first-year and second-year students' ability to individually construct arguments. In particular, the first-year students showed more improvement even after the second PBL session, while the second-year students did not show further significant improvement after the first PBL. These findings are consistent with the results of Cho and Jonassen (2002), who reported that the argumentation scaffold during PBL had a transfer effect on the quality of arguments in individual problem solving. Although they used the constraint-based scaffold for computer-supported collaborative argumentation, the constraints they used are very similar to the argument categories of Toulmin's model. Thus, it is thought that both studies show very similar results.

The findings of this study are also in line with those of Wu et al. (2016). They examined the effects of computer-based cognitive mapping approaches that help students to externalize their reasoning processes and the knowledge underlying their reasoning processes when they work with clinical cases. They asked the students to report their learning processes involving five elements, including data capture, hypotheses formulation, reasoning with justification, concept identification, and concept relationships, by using the cognitive mapping tool. They insisted that the first three elements are related to the reasoning process, and the latter two reflect knowledge construction. Their results showed that the cognitive mapping group outperformed the control group in the reasoning process. Although the instructional interventions for their study are different from those of this study, there are some similarities in terms of using visual figures (cognitive mapping vs. concept map) to support students' reasoning processes and giving an attention to the key aspects of reasoning (data capture, hypotheses formulation, reasoning with justification vs. data, claim, ground) in order to elicit their reasoning processes. Thus, it is fair to say that both studies show similar results.

As for the quality of each category, for the second-year students, the quality of all categories except the rebuttals significantly increased in posttest 1 and posttest 2. The first-year students showed a similar trend, but they needed more time than the second-year students. The quality of each category improved significantly after the second PBL session. These results provided evidence that the argumentation with the concept map method also positively affects the quality of each category in the individual clinical reasoning process among both the first- and second-year students.

There was no significant difference in the pretest between first- and second-year students, but the quality of arguments constructed by second-year students was significantly higher in posttest 1 and posttest 2 than that of their counterparts. In addition, regarding each category of arguments, for claims, grounds, and warrants, the second-year students showed significantly higher-quality scores than the first-year students in posttest 1 and posttest 2. These results indicate that the argumentation with concept map method had more impact on the second-year students than on the first-year students. Obviously, clinical reasoning involves understanding the pathology of disease process (Diemers, Wiel, Scherpbier, Baarveld, & Dolmans, 2015; Patel et al., 2005), and the superior medical knowledge of the second-year students seemed to facilitate the development of their clinical reasoning skills.

The second-year students had experienced PBL five times before this study. However, there were no significant differences in the pretest including the argument scores, any other categories except the ground score, and problem-solving performance. Although they showed significantly more improvement in all the scores after the first PBL session than their counterparts, this finding indicates the need for scaffolding to support students' clinical problem-solving process in PBL.

The rebuttal scores of the second-year students did not significantly improve after the first PBL session, but significantly improved after the second PBL session. In addition, although the backing scores of the second-year students improved significantly after the first PBL session, the scores are still very low compared with other scores. In the case of the first-year students, the backing and rebuttal scores are generally the lowest scores. In addition, there were no significant differences in any tests between first- and second-year students. Thus, it is thought that backings and rebuttals are the most difficult ones to construct for both groups of students. These results are consistent with Cho and Jonassen (2002), where students almost did not provide backings and rebuttals. However, these results could also be explained by the fact that no data about physical examination of the patient were provided in clinical cases for individual clinical problem solving. Thus, students might have focused on claims, grounds, and warrants rather than on backings and rebuttals as there were not enough data to construct them. Further studies are needed to explore the reasons for the lack of backing and rebuttal categories in individual clinical reasoning.

Regarding the problem-solving performance, first-year students' problem-solving performance significantly improved after the second PBL session. Second-year students showed higher performance in posttest 1 than in the pretest and also showed higher performance in posttest 1 than the first-year students. These results are generally not consistent with Cho and Jonassen's (2002) results, which showed no significant difference between the argumentation scaffold group and the control group in terms of problem-solving performance. Successful solution of a particular clinical problem does not guarantee the successful solution of another clinical problem (Eva, 2005; Kreiter & Bergus, 2009). However, in this study, comparing the pretest results with those of other tests, the performance scores significantly increased. Further research is necessary to confirm this, but these findings show a trend whereby the argumentation with the concept map method has a positive impact on problem-solving performance as well. In posttest 2, the second-year students' performance significantly decreased. This finding may result from the difficulty of the patient case provided to the second-year students in that test. The cases utilized in this study were the cases used in the previous PBL classes, but it seems that the case given to the second-year students was particularly difficult for them.

Conclusion

This study provided evidence that argumentation with the concept map method during PBL positively affects the development of individual students' clinical reasoning skills. Specifically, utilizing argumentation with the concept map method positively affected both first-year and second-year students' ability to individually construct arguments. Furthermore, the results indicated that the argumentation with the concept map method had more of an impact on the second-year students than on the first-year students. They also show a trend whereby the argumentation with the concept map method has a positive impact on clinical problem-solving performance.

These findings are particularly relevant for medical students, but have implications for all health care professionals, including nursing students, who are required to develop clinical reasoning skills. PBL is now used as a teaching approach by a number of different faculties, so it may also be relevant to other faculties that use scientific reasoning to solve problems.

Developing students' clinical reasoning skills appears to be a critical goal in most medical curricula, but developing appropriate instructional approaches to facilitate clinical reasoning skills proves to be a difficult challenge (Wu et al., 2016). Furthermore, a typical written assessment of clinical reasoning assesses only the outcome of clinical reasoning, and the clinical reasoning process itself is not even assessed. However, this study shows that the argumentation using the concept method can be utilized during PBL to develop students' clinical reasoning skills and that arguments and performance of clinical problem solving in PBL could assess both the outcome and process of clinical reasoning.

The results of this study also showed the potential to provide information about medical students' clinical reasoning skills as early as the preclinical period. Kassab et al. (2016), Kassab and Hussain (2010), and Humbert et al. (2011) utilized concept mapping and script concordance assessments to evaluate clinical reasoning skills and successfully differentiated second-year medical students from fourth-year students. Although their assessment methods are different from those in this study, they showed that clinical reasoning could be assessed at the early stages of medical education. Likewise, the assessment methods used in this study could provide informative evaluations of clinical reasoning skills of students in PBL. This information can be utilized as remedial information for some students. Remediation activities, particularly for rebuttals and backings, can then be designed to target these specific areas of difficulty.

This study was limited in several respects. It measured the quality of arguments and problem-solving performance rather than students' actual clinical reasoning performance such as case simulations. To back up the findings of this study, research is necessary to measure actual clinical reasoning performance using the same instructional intervention. In addition, a longitudinal study might shed further light on the effect of the instructional intervention used in this study. This study was based on a medium-sized and specific population sample, which limits the generalizability of the findings. Further studies with a larger sample size are necessary. Three professors assessed the quality of problem-solving performance, but they evaluated one-third of the students each time. Although they assessed different students in each test to minimize raters' errors, the assessment of the quality of problem-solving performance was conducted by one assessor. In addition, the three clinical cases presented to each group of students were the cases used in the previous PBL classes in the university where the study occurred, but their level of difficulty did not seem totally even. These factors should be taken into account in interpreting the findings.

Acknowledgments

This work was supported by the Dong-A University research fund.

References

Barrows, H. S., & Feltovich, P. J. (1987). The clinical reasoning process. *Medical Education, 21*(2), 86–91.

Cerbin, B. (1988). The nature and development of informal reasoning skills in college students. ERIC Document Reproduction Service No. ED 298 805.

Charlin, B., Lubarsky, S., Millette, B., Crevier, F., Audétat, M. C., Charbonneau, A., . . . Bourdy, C. (2012). Clinical reasoning processes: Unravelling complexity through graphical representation. *Medical Education, 46*(5), 454–463.

Cho, K. L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development, 50*(3), 5–22.

Diemers, A. D., Wiel, M. W., Scherpbier, A. J., Baarveld, F., & Dolmans, D. H. (2015). Diagnostic reasoning and underlying knowledge of students with preclinical patient contacts in PBL. *Medical Education, 49*(12), 1229–1238.

Erduran, S., Simon, S., & Osborne, J. (2004). Tapping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88*(6), 915–933.

Eva, K. W. (2005). What every teacher needs to know about clinical reasoning. *Medical Education, 39*(1), 98–106.

Groves, M. (2012). Understanding clinical reasoning: The next step in working out how it really works. *Medical Education–Oxford, 46*(5), 444.

Harasym, P. H., Tsai, T. C., & Hemmati, P. (2008). Current trends in developing medical students' critical thinking abilities. *Kaohsiung Journal of Medical Sciences, 24*(7), 341–355.

Humbert, A. J., Johnson, M. T., Miech, E., Friedberg, F., Grackin, J. A., & Seidman, P. A. (2011). Assessment of clinical reasoning: A script concordance test designed for pre-clinical medical students. *Medical Teacher, 33*(6): 472–477.

Jonassen, D. H. (2011). *Learning to solve problems: A handbook for designing problem-solving learning environments.* New York: Routledge.

Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development, 58*(4), 439–457.

Ju, H. J. (2016). *Enhancing medical students' argumentation during hypothetico-deductive reasoning (hdr) in problem-based learning (PBL)*. PhD diss., University of Georgia.

Kassab, S. E., Fida, M., Radwan, A., Hassan, A. B., Abu-Hijleh, M., & O'Connor, B. P. (2016). Generalizability theory analyses of concept mapping assessment scores in a problem-based medical curriculum. *Medical Education, 50*(7), 730–737.

Kassab, S. E., & Hussain, S. (2010). Concept mapping assessment in a problem-based medical curriculum. *Medical Teacher, 32*(11), 926–931.

Kirschner, P. A., Buckingham Shum, S. J., & Carr, C. S. (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making*. London: Springer.

Kong, L. N., Qin, B., Zhou, Y. Q., Mou, S. Y., & Gao, H. M. 2014. The effectiveness of problem-based learning on development of nursing students' critical thinking: A systematic review and meta-analysis. *International Journal of Nursing Studies, 51*(3), 458–469.

Kreiter, C. D., & Bergus, G. (2009). The validity of performance-based measures of clinical reasoning and alternative approaches. *Medical Education, 43*(4), 320–325.

Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument." *Thinking & Reasoning, 13*(2): 90–104.

Knudson, R. E. (1991). Effects of instructional strategies, grade, and sex on students' persuasive writing. *Journal of Experimental Education, 59*(2), 141–152.

Novak, J. D., & Cañas, A. J. (2006). *The theory underlying concept maps and how to construct them*. Technical report IHMC Cmap Tools 2006-01. Pensacola, FL: Institute for Human and Machine Cognition.

Oh, S., & Jonassen, D. H. (2007). Scaffolding online argumentation during problem solving. *Journal of Computer Assisted Learning, 23*(2), 95–110.

Patel, V. L., Arocha, J. F., & Zhang, J. (2005). Thinking and reasoning in medicine. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 727–750). New York: Cambridge University Press.

Rendas, A. B., Fonseca, M., & Pinto, P. R. (2006). Toward meaningful learning in undergraduate medical education using concept maps in a PBL pathophysiology course. *Advances in Physiology Education, 30*(1), 23–29.

Rochmawati, E., & Wiechula, R. (2010). Education strategies to foster health professional students' clinical reasoning skills. *Nursing & Health Sciences, 12*(2), 244–250.

Sanders, J. A., Wiseman, R. L., & Gass, R. H. (1994). Does teaching argumentation facilitate critical thinking? *Communication Reports, 7*(1), 27–35.

Saunders, K. M. (1994). Law as rhetoric, rhetoric as argument. *Journal of Legal Education, 44*(4), 566–578.

Siegel, H. (1995). Why should educators care about argumentation? *Informal Logic, 17*(2), 159–176.

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education, 28*(2–3), 235–260.

Stoyanov, S., & Kommers, P. (2008). Concept mapping instrumental support for problem solving. *International Journal of Continuing Engineering Education and Lifelong Learning, 18*(1), 40–53.

Taylor, D., & Miflin, B. (2008). Problem-based learning: Where are we now? *Medical Teacher*, *30*(8), 742–763.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Toulmin, S. E. (2003). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Toulmin, S. E., Rieke, R. D., & Janik. A. (1984). *An introduction to reasoning* (2nd ed.). New York: Macmillan.

Wu, B., Wang, M., Grotzer, T. A., Liu, J., & Johnson, J. M. (2016). Visualizing complex processes using a cognitive-mapping tool to support the learning of clinical reasoning. *BMC Medical Education, 16*(1), 216–224.

Yeh, S. S. (1998). Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English*, 49–83.

Yuan, H., Williams, B. A., & Fan, L. (2008). A systematic review of selected evidence on developing nursing students' critical thinking through problem-based learning. *Nurse Education Today,* 28(6), 657–663.

Jihyun Si is an assistant professor of medical education at Dong-A University College of Medicine in Korea. She completed her MA at the University of Toronto in Canada and her EdD at Hanyang University in Korea in educational technology with concentrations in instructional design and e-learning. Her research interests include PBL, medical education, collaborative learning, e-learning, multimedia learning, and cognitive load theory.

Hyun-Hee Kong, MD, PhD, is a professor and in charge of medical education at Dong-A University. She has been interested in patient safety and clinical communication.

Sang-Hwa Lee is a professor in the Department of Medical Education at the College of Medicine, Dong-A University. He earned his MD and PhD from Kyungpook National University in 1991. His interests are problem-based learning, competency-based education in medicine, clinical skill training, and educational assessment.