Research Article

KME Review

Korean Medical Education Review 2019; 21(1): 31-40
https://doi.org/10.17496/kmer.2019.21.1.31

pISSN 2092-5603 | eISSN 2093-6370

The Effect of the Dr. LEE Jong-wook–Seoul Project on Medical Student Academic Performance in Laos

Jungho Kim[1], Taejong Kim[2], Jwa-Seop Shin[3,4]

[1]Department of Economics, Ajou University, Suwon; [2]KDI School of Public Policy and Management, Sejong; [3]Department of Medical Education, Seoul National University College of Medicine; [4]WHO Collaborating Center for Educational Development, Seoul, Korea

The Dr. LEE Jong-wook—Seoul Project (DLSP) in the Lao People's Democratic Republic (PDR) is a long-term development cooperation initiative aimed at upgrading the medical education capacity at the University of Health Sciences (UHS) in the Lao PDR. Under the project, eight faculty members from UHS were trained at the Seoul National University College of Medicine for a 12-month period during 2010–2011. Using a training-of-trainers model, we estimated the effect of DLSP on the learning outcomes of students, which was measured using the standard test scores of the students. The questionnaire developed by the Medical Education Assessment Consortium in Korea was given to the students and both the percent-correct and standardized scores were examined. The evaluation strategy was to compare the learning outcomes of the students taught by the trainees with those that were not and to compare the medical topics taught by the trainees with other topics. It was found that the first group of trainees in the DLSP was associated with an increase of 1.5 in the percent-correct test scores for basic medicine. While the interpretation requires caution, it is consistent with the view that a training program for faculty members could be effective in improving the teaching capacity of medical schools in developing countries.

Keywords: Developing countries, Educational measurement, Teacher training

Corresponding author
Jwa-Seop Shin
Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea
Tel: +82-2-740-8175
Fax: +82-2-741-1186
E-mail: hismed1@snu.ac.kr
https://orcid.org/0000-0002-6251-3616

Received: October 16, 2018
1st revised: January 17, 2019
2nd revised: February 2, 2019
3rd revised: February 14, 2019
Accepted: February 14, 2019

INTRODUCTION

The goal of the Dr. LEE Jong-wook—Seoul Project (DLSP) is to improve the quality of healthcare in Lao People's Democratic Republic (PDR) by upgrading the capacity of faculty members at the University of Health Sciences (UHS), which is the sole institution for advanced medical education in the country. The project was named after the late Dr. Jong-wook Lee, the former director general at the World Health Organization from 21 July 2003 to 22 May 2006. The project reflects the idea that focusing on the elite in the medical profession could be a highly effective strategy through the spill-over effects into primary and secondary healthcare and the ensuing rollout of new medical doctors. This idea originates from Korea's earlier experiences with the Minnesota Project in which the professors at Seoul National University, including those at the College of Medicine (SNUCM), received training at the University of Minnesota from 1954 to 1961 with the support of the United States government. While no rigorous quantitative evaluation of the Minnesota Project has been produced, qualitative studies have suggested that the project played a crucial role in dramatically improving the capacity of the medical profession in teaching, research, and practice in Korea [1].

The purpose of the study is to evaluate the impact of the DLSP in the first 2 years with respect to UHS students' learning outcomes. The nature of DLSP is to improve the capacity of faculty members at UHS in both teaching and research, and we focused on the former because its outcome can be measured clearly over a relatively short time period. Subsequently, the main characteristics of DLSP fall into the category of the training-of-the-trainers model in medical education. The conventional evaluation method for these programs is to collect the self-reported skill levels of the trainees before and after the program [2-4] or to survey the follow-up activities of the participants [5].

While a survey has the advantage of providing a direct measure of a program's effects, its application to DLSP is limited for two

reasons. First, the program does not envision a single uniform curriculum to be applied to all participants, rather it is designed to provide training customized to each participant. In addition, the output of the project is not confined to medical knowledge but includes the motivation for teaching and research activity. Hence, it is not easy to specify a set of skills that all the participants are expected to learn. Second, the participants may have the incentive to overstate the effectiveness of the program if they are afraid that a poor evaluation result might lead to early termination from the program.

Therefore, in this study, we alternatively propose measuring the impact of the project with respect to the students' academic performance. The idea is to compare the performance of the students for two kinds of subjects: those taught by trainees and others taught by non-trainees. However, the distinction between treatment and control subjects is not clear-cut since the curriculum at UHS consists of modules on various topics with each taught by a group of faculty members. Hence, the level of treatment for each subject in the curriculum is measured by the share of teaching hours given by the trainees and is related to the level of the student's performance.

The design of the study is novel in that a training-of-trainers model is applied to the highest level of education in the medical profession. It is also new to the literature in that a standardized test is utilized in evaluating a development program. One advantage of this approach is that it provides a way of directly measuring the educational effect of the program. Another is to assess the level of academic performance of students by medical subject, which could be a useful piece of information for upgrading the curriculum. Furthermore, one could compare the effectiveness of the education program in different countries in a systematic manner. While there are some limitations on its interpretation, it is thought that the study demonstrates the use of a standardized test of medical knowledge as a useful tool in program evaluation in the context of developing countries.

METHODS

1. Program description

Under DLSP, UHS professors are invited to receive a 1-year long training program at SNUCM. The first group of trainees selected by UHS included five professors in basic medicine and three in clinical medicine. Their specialties were anatomy, microbiology, pathology, pharmacology, physiology, hematology, obstetrics & gynecology, and endocrinology. They were 36.9 years old on average and mostly junior faculty members. The first group entered the program in November 2010. DLSP was designed as a 9-year long project subject to renewal every 3 years.

The program consists of orientation, basic training, and specialty training. The basic training included a Korean language course and workshops on various topics like medical education, medical research, epidemiology, public health, and so on. As for the specialty training, each participant was assigned an advisor who set up a curriculum based on the individual's needs. The main content of the common and specialty training for each participant is summarized in Table 1. The program of specialty training varied widely from intensive courses in microbiology to practice in software skills in physiology. The teaching methods included lectures, laboratory work, clinical practice, seminars,

Table 1. The training program of Dr. LEE Jong-wook-Seoul Project by specialization

Specialization	Educational content
Common curriculum	Disease epidemiology and management: (1) Understand the epidemiological and clinical and epidemiological characteristics of malaria, dengue fever, tuberculosis, diarrheal diseases including cholera, and measles. (2) Develop a textbook with the title, "Control and prevention of disease outbreak (in general and for six specific diseases)" for medical students
	Learning and teaching in medicine
	Understand the community health administration system: disease surveillance and response, the center for disease prevention, the national surveillance policy, and the community healthcare service in Korea
	Korean language course
Anatomy	Topics in gross anatomy, histology, and neuroanatomy
	Acquisition of techniques for morphological studies
	Cellular and molecular biological techniques with theories
	Changes of cranial indexes in Korean from the Yi-dynasty to the present
	The immunohistochemical study of some markers in animal disease models

(Continued to the next page)

Table 1. Continued

Specialization	Educational content
Microbiology	Intensive course on bacteriology, virology, and immunology for 3 months each
	Participation in research activities for 3 months
	Participation in the weekly lab meetings of each academic advisor and a monthly seminar for infectious diseases
Pathology	Examine and practice the whole process from specimen reception to diagnosis
	Examine and practice the process of slide manufacturing and specimen preparation using appropriate medical facilities
	Examine and practice frozen diagnosis, immunohistochemistry, molecular pathology, and electron microscopy
Pharmacology	Develop teaching materials based on advanced knowledge of pharmacology for use in pharmacology teaching sessions in his/her country, such as lectures and group discussions
	Obtain an understanding of the current methods used in pharmacology research laboratories to plan for and set up his/her own laboratory in the future
Physiology	Train in commercially available, non-profit home-made software for practicing medical physiology
	Train in a commercially available software for a lecture series in medical physiology
	Train in the experimental study of vascular contractility control
Pediatric hematology	Experience the variety of pediatric hematologic disease/pediatric oncologic disease and management with updated knowledge
	Obtain knowledge from comprehensive lectures on pediatric hematology-oncology
	Perform clinical research on pediatric hematology under the supervision of a director
Obstetrics & gynecology	Obtain knowledge and practical skills about fetal sonography
	Study antenatal diagnosis and obstetric procedures
	Experience fetal surgery with fetoscopy, several procedures of ART (assisted reproductive technology), and gynecologic surgery
Internal secretion	Improve clinical practice and research skills for various endocrine diseases, including diabetes mellitus, thyroid diseases, metabolic bone diseases, and pituitary and adrenal gland diseases

and academic conferences. Each trainee also received a research grant of 30 million Korean won (around USD 26,700) upon completion of the program. The program committee set up by the Korea Foundation for International Healthcare (KOFIH) and SNUCM had a meeting on a monthly basis to monitor the progress.

The existing evidence suggests that DLSP in the first year induced favorable changes among the trainees. According to the surveys and interviews during and after DLSP, the first group of participants had positive reactions to the program and assessed that their knowledge and skill had improved [6]. It has also been reported that they shared their experience, knowledge, and skills with their colleagues after returning to work [6].

2. Study design

The first group of DLSP trainees returned to UHS and started teaching in February 2012, and in this study, we compared the academic performances of fourth-year students in 2012 and 2013: the former did not receive any lectures from the DLSP trainees in basic medicine whereas the latter were taught in their third year by the first group of DLSP trainees. Basic medicine is covered in the first 3 years of education at UHS. Both groups of students took a comprehensive test on medical knowledge

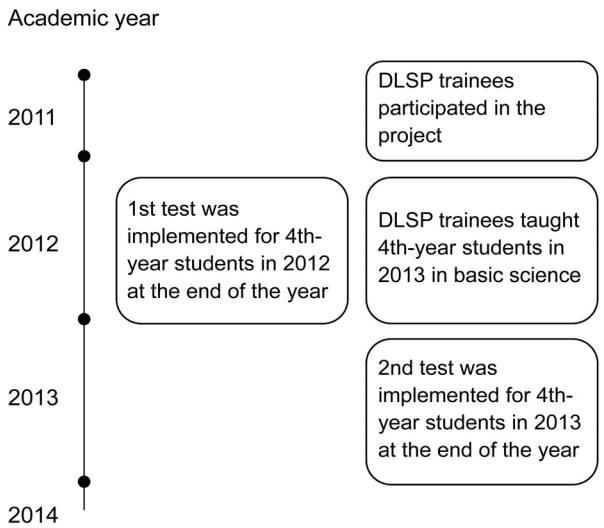

Figure 1. Timing of DLSP and medical knowledge test. DLSP, Dr. LEE Jong-wook-Seoul Project.

in the subjects of basic medicine in January 2013 and February 2014. As is presented in Figure 1, the tests coincided with the end of their fourth academic year at UHS. At the end of the test, a short survey was conducted to collect information on the students' demographic, socioeconomic, and academic backgrounds. In addition, information on the students' grades in their fourth academic year was obtained from UHS.

3. Participants

The study targeted all of the fourth-year students in 2012 and 2013, more than 80% of whom were reached. A total of 315 students out of 351 fourth-year students in 2012 (89.7%) participated in the test whereas 305 students out of 368 fourth-year students in 2013 (82.9%) participated. Demographic, socioeconomic, and other characteristics of the study participants are listed in Table 2. The fourth-year students in 2012 and 2013 were around 24 years old with the share of female students being higher in 2013 (63.7%) than in 2012 (50.3%). The share of students whose mothers had a university degree was 19.5% and 17.4% for the fourth-year students in 2012 and 2013, respectively. In both groups, around four out of 10 students considered his/her English proficiency as normal or above. The grade point average was slightly higher for the fourth-year students in 2013 (2.36) than 2012 (2.28).

4. Test development and application

The undergraduate medical education course in Korea consists of 2 years of premedical education, 2 years of basic and clinical science (including an organ system integrated curriculum), and 2 years of clinical clerkship whereas at UHS, it comprises 1.5 years in premedical education, 1.5 years of basic and clinical science, 2 years of clinical clerkship (combined with clinical lectures), and 1 year of practical experience. UHS also employed an integrated curriculum in basic and clinical science. Clinical lectures offered at UHS are structured as a clinical presentation curriculum and may be viewed as an extension of basic and clinical science. We further reviewed the academic curriculum at both UHS and colleges in Korea and found that the first 4 years of education are roughly comparable with each other. For example, a curriculum integrating basic and clinical medicine is offered to fourth-year students in medical schools in Korea whereas those in UHS participate in clinical practice in the morning and learn the integrated curriculum in the afternoon. For the test instrument, we employed the questionnaire developed by the Medical Education Assessment Consortium (MEAC) in Korea. MEAC administrates the testing of basic medicine for students who have completed 4 years at medical colleges in Korea each year. The type and level of the test items are similar to the national medical license exam and is designed to be comparable from year to year. The fourth-year students at UHS in 2012 and 2013 were given the full set of the test which had been given to the fourth-year students in Korea in 2011 and 2012, respectively. The tests were administered in the same manner as for Korean students in terms of time, and the students were provided with both English and the Laotian versions.

Each test consists of 260 multiple choice questions, although it is unlikely that all of them were relevant for students at UHS.

Table 2. Description of test-participants' characteristics and medical knowledge test

Variable	4th-year students in 2012 (N = 315)	4th-year students in 2013 (N = 305)
Participants		
Age (yr)	23.7 ± 5.1	23.9 ± 5.2
Sex (male:female)	156:159	111:194
Mother's education: no schooling	36 (11.3)	69 (22.6)
Mother's education: primary	65 (20.5)	51 (16.7)
Mother's education: secondary	104 (32.9)	86 (28.2)
Mother's education: vocational	50 (15.8)	46 (15.0)
Mother's education: university	61 (19.5)	53 (17.4)
Proficient in English[a]	126 (40.1)	122 (40.0)
Grades, 4th-year (4.0 scale)	2.28 ± 0.38	2.36 ± 0.39
Taught by DLSP[b] trainees in basic medicine courses	No	Yes
Medical knowledge test		
Test date	26th Jan. 2013	8th Feb. 2014
Total test questions	260	260
Relevant test questions	208 (80.0)	223 (85.8)

Values are presented as mean ± standard deviation, number, or number (%), unless otherwise stated.
[a]Proficiency in English indicates that one evaluates his or her level as normal or above in a five-level scale of very poor, poor, normal, good, and very good. [b]Dr. LEE Jong-wook—Seoul Project.

There clearly exist differences in curriculum, learning environment, and the prevalence of diseases between the two countries. Therefore, a workshop was held where faculty members at UHS reviewed the test questions for relevance to their students' curriculum. As can be seen in Table 2, the share of relevant questions was 80% and 86% in the tests for the fourth-year students in 2012 and 2013, respectively. Although all of the questions were given to the students so that they were under the same conditions as when administered in Korea, only relevant questions were counted in the score, and the scores were standardized for the analysis described later.

The announcement on the test was made to students 2 months in advance. Although the test was not directly related to their official grade at school, professors at UHS made it clear that the test would help students prepare for school exams. Furthermore, prizes were awarded to the top performers in each test. Since some students from other provinces often go home during the break and tend to return to campus only after the school begins, the tests were held in the last week of their fourth academic year to ensure the presence of all of the students.

5. Statistical analysis

We conduct a descriptive statistical analysis in which the test scores of the two groups were compared at the medical subject level. Since the number of questions varied across the subjects covered in the study, the main analysis was based on the percent-correct score, which indicates the percentage of correct answers. A two-sided t-test was utilized for verifying the significance of any difference in scores between the two groups. The degree of difficulty is likely to vary from subject to subject and to adjust for this across subjects between the two tests, we converted the percent-correct scores into z-scores using the mean and standard deviation of the Korean students' scores. The reference group was the fourth-year students in medical colleges in Korea who participated in the same test by MEAC corresponding to each test. There were 778 students from 25 medical schools in Korea who participated in the test of basic medicine given to fourth-year students in 2012. The reference group for the fourth-year students in 2013 included 832 students from the same 25 medical schools. The mean and standard deviation of the test scores of the reference group is reported in Table 3. The mean percent-correct score of the reference group was 52.6 for the fourth-year students in 2012 and 54.0 for those in 2013. All of the analyses were implemented via STATA ver. 14.2 (Stata Corp., College Station, TX, USA).

Table 3. Descriptive statistics for Korean medical students' scores in the medical knowledge test

Subject	4th-year students in 2011 (N=778)[a]	4th-year students in 2012 (N=832)[b]
Metabolism	58.7±13.7	48.5±16.9
Genetics	35.0±48.0	53.8±21.5
Cell & tissue	49.5±13.6	46.3±21.3
Neoplasm	46.4±19.9	67.4±17.7
Infection	47.7±12.4	55.7±10.9
Immune system	43.8±17.9	48.3±19.5
Embryology	45.7±23.7	55.7±17.0
Human response	56.9±14.4	61.4±17.9
Musculoskeletal system	50.7±17.2	39.4±18.0
Blood & hematogenesis	47.1±20.4	53.3±25.8
Circulatory system	55.2±15.3	52.2±18.1
Respiratory system	56.4±20.3	60.6±19.9
Endocrine system	63.5±19.1	61.1±11.8
Kidney & urinary tract	63.7±17.5	54.9±20.9
Digestive system	53.4±21.6	50.4±15.1
Nervous system	50.6±13.6	52.1±14.3
Reproductive system	58.2±16.3	49.0±19.7
Total	52.6±10.5	54.0±10.4

Values are presented as mean±standard deviation. Both groups are from the same 25 medical colleges in Korea.
UHS, University of Health Sciences.
[a]This group serves as a reference for 4th-year students in 2012 at UHS. [b]This group serves as a reference for 4th-year students in 2013 at UHS.

RESULTS

1. Measuring treatment

At UHS, the curriculum of basic medicine is covered by the Faculty of Basic Science and that of clinical medicine by the Faculty of Medicine. Measuring the level of treatment requires categorization of the academic curriculum. UHS conducted a major reform of its curriculum from 2004 to 2009 through a partnership with the School of Medicine at the University of Calgary. As a result, UHS has an integrated curriculum where each topic of study (for example, the cardiovascular or respiratory systems rather than anatomy or pathology) is taught by a group of lecturers. The share of DLSP participants in total teaching hours for every topic was constructed based on the list of instructors for each course, the individual lecture descriptions, and teaching hours. The topic or subject of each lecture was categorized according to the classification of test items used by MEAC.

The education in the first 3 years at UHS consisted of 1,099 hours in total in 2012, 831 hours of which were allocated to the academic subjects covered by the test. The courses not covered by the test were mostly liberal arts. DLSP trainees delivered 84 hours (10.1%) of all lectures relevant to the test. However, the treatment was restricted to the lectures for third-year students because the fourth-year students in 2013 had had a chance to be taught by the trainees only in their third year. Table 4 indicates the lecture hours and the level of treatment by medical subject. The fourth-year students in 2013 received 23 hours of lectures from the DLSP trainees, which amounts to 2.8% of all lectures. The treated subjects of the study were the reproductive (22.2%), nervous (6.7%), and digestive (3.2%) systems.

2. Test scores

The results of the test of medical knowledge by subject are given in Table 5. The average percent-correct score of all questions was 23.1 for the fourth-year students in 2012 and 23.2 for those in 2013. The fourth-year students in 2012 had the highest performances in metabolism (28.8) and the endocrine system (28.5), and the lowest performances in genetics (14.8) and the kidney & urinary tract (19.1). On the other hand, the fourth-year students in 2013 had the highest scores in blood & hematogenesis (30.2) and the reproductive system (29.4), and the lowest in the respiratory system (18.1) and the kidney & urinary tract (18.8). While the fourth-year students in 2012 and

Table 4. Lecture hours by trainees of DLSP in 2012

Subject	Lecture hours	Contribution of DLSP trainees to the 3rd-year curriculum	
		Hours by trainees	Share (%)
Metabolism	77	-	-
Genetics	35	-	-
Cell & tissue	103	-	-
Neoplasm	-	-	-
Infection	35	-	-
Immune system	16	-	-
Embryology	23	-	-
Human response	22	-	-
Musculoskeletal system	32	-	-
Blood & hematogenesis	39	-	-
Circulatory system	45	-	-
Respiratory system	57	-	-
Endocrine system	66	-	-
Kidney & urinary tract	51	-	-
Digestive system	62	2	3.2
Nervous system	105	7	6.7
Reproductive system	63	14	22.2
Total	831	23	2.8

The courses directly related to the test of medical knowledge are presented. Total hours of teaching exclude 268 hours in other subjects including 'health care management' and 'health promotion and disease prevention.'
DLSP, Dr. LEE Jong-wook–Seoul Project.

Table 5. Results of medical knowledge tests among UHS students

Subject	No. of questions		Percent-correct score		Difference		Standardized score[a]	
	C2012	C2013	C2012 (A)	C2013 (B)	B-A	p-value	C2012	C2013
Metabolism	15	15	28.8±12.3	21.1±11.9	-7.7	0.00	-2.19±0.90	-1.62±0.71
Genetics	1	6	14.8±35.6	23.7±18.3	8.9	0.00	-0.42±0.74	-1.40±0.85
Cell & tissue	13	7	25.6±11.9	21.0±14.8	-4.7	0.00	-1.76±0.87	-1.19±0.69
Neoplasm	7	7	26.4±15.4	19.6±14.0	-6.8	0.00	-1.01±0.77	-2.70±0.79
Infection	39	43	20.6±6.3	22.9±7.6	2.3	0.00	-2.17±0.50	-3.01±0.70
Immune system	11	8	23.4±13.8	20.7±13.7	-2.7	0.01	-1.14±0.77	-1.41±0.70
Embryology	3	9	20.6±23.5	23.5±12.8	2.8	0.07	-1.06±0.99	-1.89±0.75
Human response	16	9	26.2±11.6	29.0±16.3	2.8	0.01	-2.13±0.80	-1.81±0.91
Musculoskeletal system	11	8	19.6±11.8	20.8±13.6	1.2	0.26	-1.81±0.68	-1.03±0.75
Blood & hematogenesis	7	4	19.3±14.5	30.2±19.3	10.9	0.00	-1.36±0.71	-0.90±0.75
Circulatory system	12	10	24.0±11.7	26.2±14.9	2.1	0.05	-2.03±0.76	-1.44±0.82
Respiratory system	8	8	22.1±14.5	18.1±14.2	-4.0	0.00	-1.69±0.72	-2.14±0.71
Endocrine system	11	24	28.5±14.9	26.6±11.5	-1.9	0.09	-1.84±0.78	-2.93±0.98
Kidney & urinary tract	10	10	19.1±12.8	18.8±11.6	-0.3	0.80	-2.55±0.73	-1.73±0.56
Digestive system	8	16	23.3±14.2	22.7±10.7	-0.6	0.56	-1.39±0.66	-1.84±0.71
Nervous system	26	33	20.2±8.8	22.5±7.4	2.3	0.00	-2.24±0.65	-2.07±0.52
Reproductive system	10	6	26.5±14.3	29.4±18.3	2.9	0.03	-1.95±0.88	-1.00±0.93
Total	208	223	23.1±3.9	23.2±4.6	0.1	0.81	-2.82±0.37	-2.97±0.44

Values are presented as mean±standard deviation.
UHS, University of Health Sciences; C2012, 4th-year students in 2012 at UHS; C2013, 4th-yearstudents in 2013 at UHS.
[a]The scores are standardized using the mean and the standard deviation of Korean students' scores in Table 3.

2013 had similar mean percent-correct scores for the questions, there were substantial differences in the scores by subject. Compared to the fourth-year students in 2012, those in 2013 had a higher percent-correct score in blood & hematogenesis and genetics but a lower score in metabolism and neoplasm. The p-values of the t-test imply that these differences were statistically significant, although the inference requires caution because the degree of difficulty of the test questions was not accounted for across the subjects in the two tests.

As an alternative measure, the standardized score reflects the performance relative to that of medical students in Korea. The mean standardized score of all subjects was -2.82 for the fourth-year students in 2012 and -2.97 for those in 2013. The highest performance by standardized score was observed in genetics (-0.42) for the fourth-year students in 2012 and for blood & hematogenesis (-0.90) for those in 2013. The lowest score was recorded for the kidney & urinary tract (-2.55) for the fourth-year students in 2012 and for infection (-3.01) for those in 2013. The fourth-year students in 2013 had higher standardized scores in the reproductive system and the kidney & urinary tract but lower ones in neoplasm and the endocrine system than their counterparts in 2012.

3. Estimating the treatment effect

Essentially, our analysis was to examine whether the test scores improved more in subjects with more trainees as instructors. We estimated the program impact by comparing the average scores of the treatment and control groups. The treatment group consisted of the digestive, nervous, and reproductive systems, while the other 12 subjects formed the control group. The subject of genetics was excluded because it had less than three questions in one of the tests. The subject of neoplasm was also excluded because it is not covered by the curriculum at UHS. The results are summarized in Table 6. The mean percent-correct score of the control subjects was 0.1 higher for the fourth-year students in 2013 than for those in 2012 whereas that of the treated subjects was 1.6 higher for the fourth-year students in 2013 than for those 2012. That is to say, the score difference for the treatment subjects was larger than for the control subjects.

Since the level of treatment was different for each treated subject, we did not expect the same effect for each subject. When only the digestive system was compared with the control group, the fourth-year students in 2013 had a lower score than their counterparts in 2012, implying that the treatment effect was negative. On the other hand, the percent-correct score for the

Table 6. Effects of DLSP Trainees on test scores of medical knowledge

Variable	Percent-correct score		Share of trainees in lecture hours (%)	
	4th-year students in 2012	4th-year students in 2013	4th-year students in 2012	4th-year students in 2013
Non-DLSP subjects	23.1±4.5	23.2±5.3	0.0	0.0
DLSP subjects	23.3±7.3	24.9±7.6	0.0	10.7
Digestive system	23.3±14.2	22.7±10.7	0.0	3.2
Nervous system	20.2±8.8	22.5±7.4	0.0	6.7
Reproductive system	26.5±14.3	29.4±18.3	0.0	22.2

Values are presented as mean±standard deviation.
DLSP, Dr. LEE Jong-wook-Seoul Project.

nervous system for the fourth-year students in 2013 was 2.3 higher than for those in 2012, as was that for the reproductive system (2.9 higher). It is puzzling to have a negative treatment effect for the digestive system, although the treatment effects for the individual subjects altogether suggests that a larger treatment is related to a larger impact.

The previous interpretation is limited in that the degree of difficulty of the tests varied across subjects and over two tests. Interestingly, these results are qualitatively consistent with those obtained when standardized scores were used, as is partially evident in Table 5. Hence, the comparisons between the fourth-year students in 2012 and 2013 and between the treated and control subjects may be viewed as estimating the effect of DLSP under the assumption that the reference groups for the two tests had the same level of academic achievement. Admittedly, this is a rather strong assumption even though the size of the reference groups was relatively large. The bottom line is that the findings are robust for the choice of outcome measures and the inference should be made with caution.

DISCUSSION

While the results suggest that DLSP had a positive impact on students' learning outcomes, there are a couple of issues threatening the credibility of the findings. First, the students' participation rates in the tests fluctuated over the years, which casts doubt on whether the samples in different years are comparable to each other. The issue seems less serious given that the participation rates were above 80% in both tests. One way to address this issue is to estimate the effects of the project separately for high-performing and low-performing students. In fact, the estimation results indicate that there was little difference in the effects for different groups of students (data not shown). Hence, the participation rate in the tests does not seem to have caused significant bias. Furthermore, one may be concerned that the fourth-year students in 2012 potentially had a different level of academic ability compared to those in 2013. The analysis carried out requires the assumption that the difference in the general academic ability between the two groups of students was the same for all medical subjects. That is to say, even if one group of students had a higher ability than the other, it did not cause bias in the estimate as long as the difference was the same for, for instance, the endocrine system and the reproductive system. This is because we compared the treated students with the control students and the treated subjects with the other subjects. This assumption does not seem to be a strong one.

Second, the test was neither compulsory nor reflected in the final grades. Therefore, the students may not have revealed their true academic performance because they were not strongly motivated in the test. Although it is true that the students' motivation could have been weak, the estimate of the treatment effect is unbiased as long as the fourth-year students in 2012 and 2013 had the same degree of motivation. Since the test was implemented in the same way for the 2 years, it is less likely that there was a significant change in the motivation among the students.

Third, one may be concerned that the magnitude of treatment was too small to be analyzed. While only the first year of DLSP was examined, we were measuring the difference over time and indeed found a significant impact. In fact, by comparing the non-DLSP and DLSP subjects, one could associate an increase in the trainees' share of the lecture hours with an increase in their test scores to estimate the size of the treatment effect. However, this inference could be overstretching in that in addition to the limitation of the study, the estimates should be interpreted

as a marginal effect at an early stage of the intervention. When measured against their teaching hours, the first group of trainees account for only 3% of the lecturers in the Faculty of Basic Science, thus the impact produced by additional trainees is likely to become smaller as the project is rolled over.

Fourth, the relevance of the test questions to the context of the Lao PDR needs to be further reviewed. Although the analysis was based on a set of test questions deemed by UHS professors as relevant, the composition of the questions in terms of medical subject and clinical practice might not have reflected the curriculum at UHS as well as it does at medical schools in Korea. Therefore, the fact that the percent-correct score of the overall test is 23% should not be interpreted as suggesting the absolute level of the academic performance of the students at UHS compared to those in Korea. We claim only that there is a relationship between the difference in the scores and the difference in the teaching hours delivered by the DLSP trainees.

Fifth, the study evaluates the project only with respect to educational outcomes. The analysis of the research outcome and the clinical practice performed by trainees would give us a more balanced view of the project's impact. One related study observed that the participants of the DLSP were more active both in teaching students and treating patients when comparing after and before the training [7]. On the other hand, the authors reported that the change in research activity is limited largely due to the lack of equipment and research funds. In our view, the quantitative analysis presented in this study complements this qualitative study.

Last but not least, it should be mentioned that the scope of the study was to measure the consequences of the program in the short term. However, a complete evaluation should be based on monitoring the trainees' teaching performance over a longer period of time. Specifically, a long-term evaluation system is required, including a review of the trainees' teaching plans upon their return to work, a follow-up of the university's support, and regular monitoring of performance.

The modeling of the training of health professionals in a leading institution as a way to improve the overall capacity of medical staff in a country has been in place since the 1960s. While our article contributes to the literature by evaluating such a program quantitatively, there are still a number of questions to be explored. One is whether a program like DLSP complements other projects to improve the educational environment. Another is whether the training-of-trainers model can be applied to a different level in the medical system. For example, the Continuing Professional Development Program (CPDP) developed by the Korea International Cooperation Agency, SNUCM, and UHS train medical associates/assistants (MAs) who are the main medical staff in charge of primary care in rural areas of the Lao PDR (a report based on a qualitative evaluation study has shown that the CPDP improved the clinical practice of MAs [8]). From the perspective of national healthcare policy, it is relevant to ask whether training-of-trainers models at different levels (like DLSP and CPDP) complement each other. These questions are ultimately directed toward the question of how to design an efficient package of programs to build up the medical education system in a developing country.

Methodologically, the study demonstrates a case where a test of medical knowledge can be utilized in evaluating programs targeting learning outcomes among students in the context of developing countries. The standardized test scores themselves can help to identify the strength and weakness of any academic program. Furthermore, to the extent that the test instruments can be readily employed in other contexts, the design of the research has the potential to compare official development aid programs across different countries.

AUTHORS' CONTRIBUTIONS

JK, TK, and JSS conceived the study. JK designed the test of medical knowledge and conducted analysis. TK and JSS interpreted the results. All authors contributed to writing and approved the final manuscript.

CONFLICT OF INTEREST: No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS: Authors would like to thank the former President Joon-Kyung Kim of KDI School of Public Policy and Management (KDI school) for his generous support. Authors benefited greatly from the discussion with Dr. Yong Choi and Dr. Jung Soo Kim who served as senior consultants for the DLSP. Authors are also thankful to Dr. Seong Jin Jeong at Children's Hospital for his cooperation in organizing the academic tests at UHS. Special thanks to Dr. Hyunbae Yoon at SNUCM who provided a professional opinion on the classification of academic curriculum. The authors appreciate the generous cooperation of

the institutions including KOFIH, SNUCM, UHS, the Lao Ministry of Health, the Prime Minister's Office and the Ministry of Health and Welfare of the Korean government and MEAC in Korea. The authors would also like to gratefully acknowledge the superb research assistance provided by Ms. Sunjin Kim, Mr. Boumny Inthakesone, Ms. Eunkyung Min, Mr. Sungoh Kwon, Ms. Ha Kyeong Lee, Mr. Daehong Kim and the excellent administrative support from the staff at the KDI School's Impact Evaluation Lab, Ms. Min Young Seo, Ms. Youngjoo Jung, Ms. Eunji Choi, Mr. Sungmook Kang, Ms. Joung-Hyun Kim, Ms. Hyomin Kwon.

FUNDING: The DLSP is funded by the KOFIH. The article is an output of the research project, "A real time impact evaluation of DLSP in Lao PDR" funded by the KDI school.

REFERENCES

1. Lee WJ. The influence of Minnesota Project on the Korean medical education [dissertation]. Seoul: Seoul National University; 2006.
2. Green ML. A train-the-trainer model for integrating evidence-based medicine training into podiatric medical education. J Am Podiatr Med Assoc. 2005;95(5):497-504.
3. Levine SA, Brett B, Robinson BE, Stratos GA, Lascher SM, Granville L, et al. Practicing physician education in geriatrics: lessons learned from a train-the-trainer model. J Am Geriatr Soc. 2007;55(8):1281-6.
4. Stratos GA, Katz S, Bergen MR, Hallenbeck J. Faculty development in end-of-life care: evaluation of a national train-the-trainer program. Acad Med. 2006;81(11):1000-7.
5. Hiner CA, Mandel BG, Weaver MR, Bruce D, McLaughlin R, Anderson J. Effectiveness of a training-of-trainers model in a HIV counseling and testing program in the Caribbean Region. Hum Resour Health. 2009;7:11.
6. Lee S, Shin JS, Kim DH, Kim EJ, Hwang J, Yoon HB. Evaluation of first year outcomes of Lee Jong-Wook Seoul Project using Kirkpatrick's four-level evaluation approach. Korean J Med Educ. 2014;26(1):41-51.
7. Shin JS, Lee S. The 4th year report on education program and its performance of the Dr. Lee Jong-wook—Seoul Project. Seoul: Seoul National University College of Medicine; 2015.
8. Yoon HB, Shin JS. Final report of the continuing professional development training project to strengthen the capacity of provincial and district hospitals in Lao PDR (2012-2014). Seoul: Seoul National University College of Medicine, Seoul National University Hospital; 2015.