Achieving physical examination competence through optimizing hands-on practice cycles: a prospective cohort comparative study of medical students

Zinan Zhang[a,b*], Zhenwei Tang[a,c*], Fang Wang[d], Jingjia Yue[e], Youzhou Tang[f], Boyue Jiang[a], Yue Gou[a], Ben Lv[g], Anliu Tang[b], Xiaohong Tang[e,h]

[a] Xiangya School of Medicine, Central South University, Changsha, Hunan, China

[b] Department of Gastroenterology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.

[c] Department of Dermatology, Xiangya Hospital of Central South University, Changsha, Hunan, China.

[d] Department of Endocrinology and Metabolism, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.

[e] Department of Cardiovascular Medicine, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.

[f] Department of Nephrology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.

[g] Department of Haematology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.

[h] The Clinical Skills Training Center, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.

*These two authors contributed equally to this article.

Corresponding Author:
Anliu Tang, E-mail: tanganliuxy3@csu.edu.cn TEL/FAX: (86) 88618457
Xiaohong Tang, E-mail: tangxh007007@163.com TEL/FAX: (86) 88618274
The Third Xiangya Hospital of Central South University, 138 Tongzipo Road, Changsha, Hunan 410013, China.

Abstract

Background: Deliberate practice (DP) was proposed for effective clinical skill training, which highlights focused, repetitive practice and feedback as the key points for practice. Although previous studies have investigated the effect of feedback in DP, little is known about the proper repetitive cycles of clinical skills training especially in physical examination (PE) training.

Methods: We drew learning curves and designed a comparative study to find out the optimal number of hands-on practice cycles, an important aspect of DP, in abdominal PE training for medical students. A comparative study was conducted to validate the optimal number of hands-on practice by dividing students into two cohorts including Cohort A (high-frequency hand-on training) and B (low-frequency hand-on training).

Results: The learning curve study of 16 students exhibited a threshold of four repetitive practices when 81.25% students reached the competence score. A total of 74 students' final exam scores were collected for analysis. Students in Cohort A (4-5 PEs) scored significantly higher than those in Cohort B (≤3 PEs) [84.41±11.78 vs 76.83±17.51] in the final exam (P = 0.030)].

45   **Conclusion:** High-frequency practice can improve students' competence of abdominal PE
46   skill. We recommend four cycles of hands-on practice for each student in a training course like
47   PE training.
48
Introduction
50   Clinical skill performance is considered to be a core competency for medical students and is
51   crucial to professionalism in medical practice. Most importantly, it contributes to successful
52   outcomes in patient care (Zhang et al. 2015). For instance, physicians' underperformance in
53   physical examination (PE), a basic clinical skill and an important part of a physician's daily
54   activity (Grune 2016), has become a frequent cause of medical errors that cause nearly 100,000
55   deaths per year (Institute of Medicine Committee on Quality of Health Care in 2000; Verghese et
56   al. 2015). It has been repeatedly observed that medical students and clinical interns should
57   improve their clinical skills (Krautter et al. 2015; Ramani et al. 2010; Xiaoqing 2015). In our
58   institution, there are more than 100 pre-clinical students who annually start their clinical skills
59   training at their third year. Usually, 10 to 15 students take a three-hour PE training course
60   together. The mean scores of PE final exams are between 70 to 80 out of 100, far from 90, the
61   predefined level of competency determined by our institution experts. In order to find the
62   reasons for such unsatisfactory results, we distributed questionnaires to 10 medical schools. We
63   found that the average number of hands-on practice cycles among the students was 2.69 based on
64   the student questionnaires. 67.65% of the students thought that the practice cycles in class were
65   insufficient or seriously inadequate (supplementary materials). Practice cycles seem to be the
66   crux of the problem.
67   Achieving competence in skills usually requires the involvement of proper practice.
68   According to behaviorism, practice serves as an important part of the psychomotor skill
69   education, and can not only improve the learner's clinical ability, but also enhance the cognitive
70   enrichment, especially memory formation (McGaghie & Harris 2018). In recent years, deliberate
71   practice (DP) has been recommended for effective skills training, emphasizing focused,
72   repetitive practice and feedback (Motola et al. 2013). The positive effect of DP in improving PE
73   training has been supported by systematic reviews. However, most studies about DP in PE
74   training focused on the effect of feedback in DP (Criley et al. 2008; Houck et al. 2002;
75   Mookherjee et al. 2013; Smith et al. 2006). There were few studies exploring the appropriate
76   number of hands-on practice cycles, which is another key element of DP (Byrne et al. 2008) in
77   PE training that can help medical students reach the best learning outcomes within limited class
78   hours.
79   To address this issue, we carried out the research with two aims: **1.** to find the optimal
80   number of practice cycles in a PE training course; and **2.** to validate the threshold (cut-off cycle)
81   defined by the learning curve, as the abdominal PE an example.
82
Materials & Methods
84   The ethical approval for this study was provided by the Ethical Committee of The Third
85   Xiangya Hospital of Central South University, China(2018-S406). The written informed consent
86   was obtained from every subject before the enrollment. All selected subjects volunteered to
87   participate in the study.
88   ***Standardization of instructors and examiners***
89   Four instructors and assessors involved in this study were carefully selected by the
90   Department of Internal Medicine of the Third Xiangya hospital. They had similar educational

91 and clinical backgrounds (M.D. and senior physician), which ensured their consistency in
92 abdominal PE experience. Before they taught PE class, they received standardized training to
93 ensure the consistency of their teaching method and scoring. After a pre-class training, five
94 student volunteers performed the abdominal PE and each of them was scored independently by
95 all the assessors using the same assessment instrument to help evaluate the consistency of the
96 assessors' scores (k=0.908).
97 ***Learning curve study***
98 The flow chart of the learning curve study and comparative study is shown in Figure 1.
99 Angoff method (Shulruf et al. 2016) was conducted to determine the competence score for the
100 students who took part in the learning curve. A total of five experts were enrolled and
101 determined that the competence score was 90 percent. Procedures of the learning curve study are
102 as follows: First, 16 third-year medical students were recruited voluntarily and instructed how to
103 perform abdominal PE. Then the students performed abdominal PE on their classmates and were
104 assessed by the same assessment instrument for at least five times. A brief instructor-led
105 feedback was given after each assessment. The scores of participants at each turn were collected
106 for analysis.
107 ***Comparative Study***
108 In order to validate the threshold of the learning curve study, a comparative study was
109 conducted. During the 2019 academic year (From March 2019 to July 2019), all inexperienced
110 third-year medical students who started to learn physical examination in Xiangya School of
111 Medicine of Central South University were eligible for inclusion. None of the students had
112 previous experience in PE. Students who did not sign the consent form of the research program
113 were excluded.
114 Participants were randomly divided into cohort A and B. The instructor-student ratio of
115 each cohort was 1:12 to 1:14. All participants attended the abdominal PE training course. In the
116 class, the instructors first demonstrated and showed every detail of abdominal PE step by step to
117 make sure that students could independently perform the skill later, and then the students
118 practiced abdominal PE on their classmates. During the practice session, instructors provided
119 feedback based on the students' performance. Instructors would demonstrate the correct
120 procedure if students made any mistake. Each participant in cohort A tried four or five PEs in
121 class (high-frequency hand-on training), while each participant in cohort B did no more than
122 three PEs (low-frequency hand-on training). The fourth cycle was the cut-off cycle of PE
123 training according to the learning curve study.
124 75 days later, participants who were still in the program and did not receive the learning
125 curve study were randomly selected to have a final exam on abdominal PE. The scores of the
126 final exam were used as an indicator to evaluate the performance of the students in the two
127 cohorts. Due to the limitation of test time and the number of examiners, some students were
128 randomly selected to participate in the test. Students who didn't reach the competence score
129 would accept the remediation, retraining and retest later until they reach the competence score.
130 ***Assessment instruments***
131 We developed our own assessment tool based on the scale that our institute used for years in
132 abdominal PE evaluation. As recommended by the Ministry of Health of the People's Republic
133 of China on Diagnostic Syllabus, and the five domains described by Messick (Messick & Series
134 1990), the reliability and validity of the assessment tool are demonstrated as follows: The content
135 was developed by experienced experts based on the diagnostic textbook designated by the
136 Ministry of education of China (Wan XH). The response process was ensured by the same

137 training and exam time for students. The internal structure was demonstrated by unifying the
138 training assessors and by measuring the Kendall's W value of the assessment instruments.
139   ***Statistical analysis***
140   The recorded data were analyzed by IBM SPSS version 23.0 (IBM Corporation, Armonk,
141 NY, USA). The cartogram was drawn by Origin Pro 2019 version 9.6.0.172 (OriginLab
142 Corporation, Northampton, MA, USA). The Kendall's Coefficient of Concordance was
143 conducted based on the scores of the five student volunteers' PE performance assessed by all the
144 assessors. The students' t tests were used for quantitative variables. Chi-square test was used to
145 detect the difference in the attainment rate of competence score in each repetition in the learning
146 curve study. Statistical significance was set at a P value of <0.05.

147

148 **Results**
149   ***Learning curve***
150   In the learning curve study, 81.25% of the students reached the competence score when they
151 performed the procedure at the fourth cycle of practice. Although more students could achieve
152 competence score after practicing five times during the classes, the effect of the improvement of
153 scores at the fifth cycle compared with that at the fourth cycle was not obvious (P > 0.05. **Figure**
154 **2**). Therefore, the fourth cycle was considered to be the cut-off cycle of abdominal PE teaching
155 and was used for the following comparative study.

156

157   ***Comparative study***
158   A total of 118 students were enrolled in the comparative study, including 59 students in
159 cohort A and 59 students in cohort B. Of the 118 students, 25 students withdrew from the
160 training course and 74 students were randomly selected to participate in the test. Of the 74
161 students whose data was included in the final exam, 34 (45.90%) were from Cohort A and
162 received high-frequency hands-on training while 40 (54.10%) were from Cohort B and received
163 low-frequency hands-on training. The former scored significantly higher than the latter in the
164 final exam (84.41±11.78 vs 76.83±17.51; P = 0.030; Figure 3).

165   **Discussion**
166   In this study, we focused on the cycles of hands-on practice, an important aspect of DP, in
167 abdominal PE training through learning curve drawing and validated comparative study. Our
168 results demonstrated that the number of students' encounters with hands-on practice would
169 influence the outcome of DP in abdominal PE training, and a threshold of four repetitive
170 practices should be recommended for DP designing in abdominal PE training.
171   It is well known that DP can be used in medical education, especially in simulation-based
172 medical education. A meta-analysis proved that DP in simulation-based courses could improve
173 trainees' skill acquisition more efficiently compared to traditional training methods (McGaghie
174 et al. 2011). Generally, the core of DP locates at practice and feedback. In this method of
175 training, the learner repetitively practices procedures and undergoes assessment with feedback,
176 resulting in observed improvement of skill performance (Duvivier et al. 2011; Ericsson 2004;
177 McGaghie et al. 2011). In the long run, a constant skill improvement and fair maintenance can
178 also be observed following this method (McGaghie et al. 2011). Therefore, DP seems to be an
179 ideal means of PE training (Mookherjee et al. 2013). Previous studies focused on the application
180 of repetitive hands-on practice and DP in cardiac auscultation skills, advanced cardiac life
181 support and temporary hemodialysis catheter insertion (Barsuk et al. 2009; Butter et al. 2010;

182 Wayne et al. 2005). However, as for pre-clinical PE training courses, little information could be
183 provided from previous studies. In contrast, our study has explored the potential of repetitive
184 hands-on practice in PE course, a basic skill training applying DP theory. Still, more detailed
185 information should be investigated with practical application of DP in medical training. For
186 instance, Davidson et al proved a beneficial role of providing first-year cardiology fellows up to
187 1 hour of DP on a right heart catheterization simulator. However, the author also declaimed that
188 the design of DP was not 'extensive' and no specific parameters such as numbers of cycle were
189 provided to describe the quality of DP (Davidson et al. 2021). Similarly, Barsuk et al
190 demonstrated the beneficial role of 2-hour DP designing in central venous catheter (CVC)
191 simulator courses, but no specific information about the numbers of repetition was
192 involved(Barsuk et al. 2009). In this case, we systematically investigate this component of DP
193 designing and we suspected that future studies related to DP should take the numbers of cycle
194 into consideration.
195   PE training program with DP is to make most students achieve the competence scores
196 required by the curriculum objectives. Although it is widely accepted that practice makes perfect,
197 the training time for every student is limited in reality and the students cannot practice over and
198 over without limitation. Therefore, it is important to determine an optimal number of practice
199 cycles in a time-limited PE class. Learning curve, one of the best ways to evaluate how quickly it
200 takes to acquire new skills or knowledge, has been widely used in researches regarding medical
201 skill learning such as surgical operation (Kockerling 2018; Lodhi 2009), patient condition
202 assessment (Cheung et al. 2014), and basic clinical skill (Jiang et al. 2011), et al, but rarely
203 reported in PE study. In general, these studies all showed that a learning plateau would appear
204 after successive and repetitive practice. We assumed that describing the cut-off between the rise
205 and the plateau of learning curve could help optimize course design to balance training efficiency
206 and training time. In this study, we drew a learning curve of DP based PE courses and found that
207 most students (>80%) could reach the competence score after performing four PEs. As a result,
208 we assumed that the optimum number of hands-on practice cycles in a PE training class was
209 four. Our comparative study further validated the assumption, in which two cohorts were
210 established according to the cut-off value. Students from the high-frequency hands-on training
211 cohort practiced four or five times and acquired better scores than their counterparts. The impact
212 of DP on learning was largely determined by the quality of practice. We hypothesized that, in the
213 higher frequency group, students were assigned to practice four or five times, which set a
214 quantitative requirement of DP, and this further facilitated the impact via repetitive practice of
215 DP. Also, since individual feedback was involved in DP, students from high-frequency groups
216 might receive more feedback from instructors, which therefore may help them acquire better
217 competence. Future attention can be paid on detailed requirements of the feedback in DP design.
218   We believe that the repetitive cycle of DP plays a significant role in the learning efficiency.
219 For better outcomes, the curriculum designers could increase the instructor/student ratio and
220 enlarge the practice portion in a time-limited PE training to ensure that every student can get at
221 least four opportunities to perform PE on standard patients or their classmates.
222   The study has several limitations: 1. We did not conduct baseline test because participants
223 who had any previous PE training or exposure were excluded at the beginning of the research
224 project. As a result, we assumed that all participants had no experience in PE and their different
225 knowledge or mastery of PE was not considered as a variable. 2. Due to limited time in the
226 practice session and the shortage of teaching resources, we did not conduct a cross-design to
227 ensure that every participant in the project could receive both high-frequency hands-on training

228 and low-frequency hands-on training. Given limitations mentioned above, our team plan to
229 further our study with lager student population that focusing more on the relationship between
230 the numbers of cycle and feedback. Besides, we are interested in examining whether the PE of
231 other regions systems would follow the same pattern.

Conclusions

233 The number of practice cycles for students in abdominal PE training course should be
234 stipulated to ensure that students have adequate performance chances in class. High-frequency
235 hands-on training strategy can enable students to acquire and maintain PE skill better and help
236 improve their final exam scores. We recommend four hands-on practices for each student in a
237 time-limited PE training course.

238

Acknowledgements

240 We extend sincere gratitude to all the students and volunteers who have participated in our
241 research program.

242

References

244 Barsuk JH, Ahya SN, Cohen ER, McGaghie WC, and Wayne DB. 2009. Mastery learning of temporary hemodialysis
245 catheter insertion by nephrology fellows using simulation technology and deliberate practice. *Am J Kidney*
246 *Dis* 54:70-76. 10.1053/j.ajkd.2008.12.041
247 Butter J, McGaghie WC, Cohen ER, Kaye M, and Wayne DB. 2010. Simulation-based mastery learning improves
248 cardiac auscultation skills in medical students. *J Gen Intern Med* 25:780-785. 10.1007/s11606-010-1309-x
249 Byrne AJ, Pugsley L, and Hashem MA. 2008. Review of comparative studies of clinical skills training. *Med Teach*
250 30:764-767. 10.1080/01421590802279587
251 Cheung PP, Dougados M, Andre V, Balandraud N, Chales G, Chary-Valckenaere I, Dernis E, Gill G, Gilson M, Guis
252 S, Mouterde G, Pavy S, Pouyol F, Marhadour T, Richette P, Ruyssen-Witrand A, Soubrier M, Nguyen M,
253 and Gossec L. 2014. The learning curve of nurses for the assessment of swollen and tender joints in
254 rheumatoid arthritis. *Joint Bone Spine* 81:154-159. 10.1016/j.jbspin.2013.06.006
255 Criley JM, Keiner J, Boker JR, Criley SR, and Warde CM. 2008. Innovative web-based multimedia curriculum
256 improves cardiac examination competency of residents. *J Hosp Med* 3:124-133. 10.1002/jhm.287
257 Davidson LJ, Chow KY, Jivan A, Prenner SB, Cohen ER, Schimmel DR, McGaghie WC, Barsuk JH, Wayne DB, and
258 Sweis RN. 2021. Improving cardiology fellow education of right heart catheterization using a simulation
259 based curriculum. *Catheter Cardiovasc Interv* 97:503-508. 10.1002/ccd.29128
260 Duvivier RJ, van Dalen J, Muijtjens AM, Moulaert VR, van der Vleuten CP, and Scherpbier AJ. 2011. The role of
261 deliberate practice in the acquisition of clinical skills. *BMC Med Educ* 11:101. 10.1186/1472-6920-11-101
262 Ericsson KA. 2004. Deliberate practice and the acquisition and maintenance of expert performance in medicine and
263 related domains. *Acad Med* 79:S70-81. 10.1097/00001888-200410001-00022
264 Grune S. 2016. [Anamnesis and clinical examination]. *Dtsch Med Wochenschr* 141:24-27. 10.1055/s-0041-106337
265 Houck WA, Soares-Welch CV, Montori VM, and Li JT. 2002. Learning the thyroid examination--a multimodality
266 intervention for internal medicine residents. *Teach Learn Med* 14:24-28. 10.1207/S15328015TLM1401_7
267 Institute of Medicine Committee on Quality of Health Care in A. 2000. In: Kohn LT, Corrigan JM, and Donaldson
268 MS, eds. *To Err is Human: Building a Safer Health System*. Washington (DC): National Academies Press
269 (US)
270 Copyright 2000 by the National Academy of Sciences. All rights reserved.
271 Jiang G, Chen H, Wang S, Zhou Q, Li X, Chen K, and Sui X. 2011. Learning curves and long-term outcome of
272 simulation-based thoracentesis training for medical students. *BMC Med Educ* 11:39. 10.1186/1472-6920-11-
273 39
274 Kockerling F. 2018. What Is the Influence of Simulation-Based Training Courses, the Learning Curve, Supervision,
275 and Surgeon Volume on the Outcome in Hernia Repair?-A Systematic Review. *Front Surg* 5:57.
276 10.3389/fsurg.2018.00057
277 Krautter M, Diefenbacher K, Koehl-Hackert N, Buss B, Nagelmann L, Herzog W, Junger J, and Nikendei C. 2015.
278 Short communication: final year students' deficits in physical examination skills performance in Germany. *Z*

279   *Evid Fortbild Qual Gesundhwes* 109:59-61. 10.1016/j.zefq.2015.01.003
280 Lodhi MFBJPMJ. 2009. Learning curve for Lichtenstein hernioplasty.224-227.
281 McGaghie WC, and Harris IB. 2018. Learning Theory Foundations of Simulation-Based Mastery Learning. *Simul*
282   *Healthc* 13:S15-S20. 10.1097/SIH.0000000000000279
283 McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, and Wayne DB. 2011. Does simulation-based medical education
284   with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative
285   review of the evidence. *Acad Med* 86:706-711. 10.1097/ACM.0b013e318217e119
286 Messick, and Series SJERR. 1990. VALIDITY OF TEST INTERPRETATION AND USE. 1990:1487-1495.
287 Mookherjee S, Pheatt L, Ranji SR, and Chou CL. 2013. Physical examination education in graduate medical
288   education--a systematic review of the literature. *J Gen Intern Med* 28:1090-1099. 10.1007/s11606-013-2380-
289   x
290 Motola I, Devine LA, Chung HS, Sullivan JE, and Issenberg SB. 2013. Simulation in healthcare education: a best
291   evidence practical guide. AMEE Guide No. 82. *Med Teach* 35:e1511-1530. 10.3109/0142159X.2013.818632
292 Ramani S, Ring BN, Lowe R, and Hunter D. 2010. A pilot study assessing knowledge of clinical signs and physical
293   examination skills in incoming medicine residents. *J Grad Med Educ* 2:232-235. 10.4300/JGME-D-09-
294   00107.1
295 Shulruf B, Wilkinson T, Weller J, Jones P, and Poole P. 2016. Insights into the Angoff method: results from a
296   simulation study. *BMC Med Educ* 16:134. 10.1186/s12909-016-0656-7
297 Smith CA, Hart AS, Sadowski LS, Riddle J, Evans AT, Clarke PM, Ganschow PS, Mason E, Sequeira W, and Wang
298   Y. 2006. Teaching cardiac examination skills. A controlled trial of two methods. *J Gen Intern Med* 21:7-12.
299   10.1111/j.1525-1497.2005.00254.x
300 Verghese A, Charlton B, Kassirer JP, Ramsey M, and Ioannidis JP. 2015. Inadequacies of Physical Examination as a
301   Cause of Medical Errors and Adverse Events: A Collection of Vignettes. *Am J Med* 128:1322-1324 e1323.
302   10.1016/j.amjmed.2015.06.004
303 Wan XH LX. Diagnostics (The Ninth Round Of Planning Teaching Materials For The Five-Year Undergraduate
304   Clinical Medicine Major In National Colleges And Universities). *Beijing: People's Medical Publishing*
305   *House of China* 9th ed.
306 Wayne DB, Butter J, Siddall VJ, Fudala MJ, Linquist LA, Feinglass J, Wade LD, and McGaghie WC. 2005.
307   Simulation-based training of internal medicine residents in advanced cardiac life support protocols: a
308   randomized trial. *Teach Learn Med* 17:210-216. 10.1207/s15328015tlm1703_3
309 Xiaoqing X. 2015. The present situation and thinking of clinical basic skill teaching in medical education. *Education*
310   *(Higher Education)* 000:94-94.
311 Zhang MY, Cheng X, Xu AD, Luo LP, and Yang X. 2015. Clinical simulation training improves the clinical
312   performance of Chinese medical students. *Med Educ Online* 20:28796. 10.3402/meo.v20.28796
313

Figure 1

Figure 1

The flow chart of enrollment process for students who took part in learning curve study and comparative study.
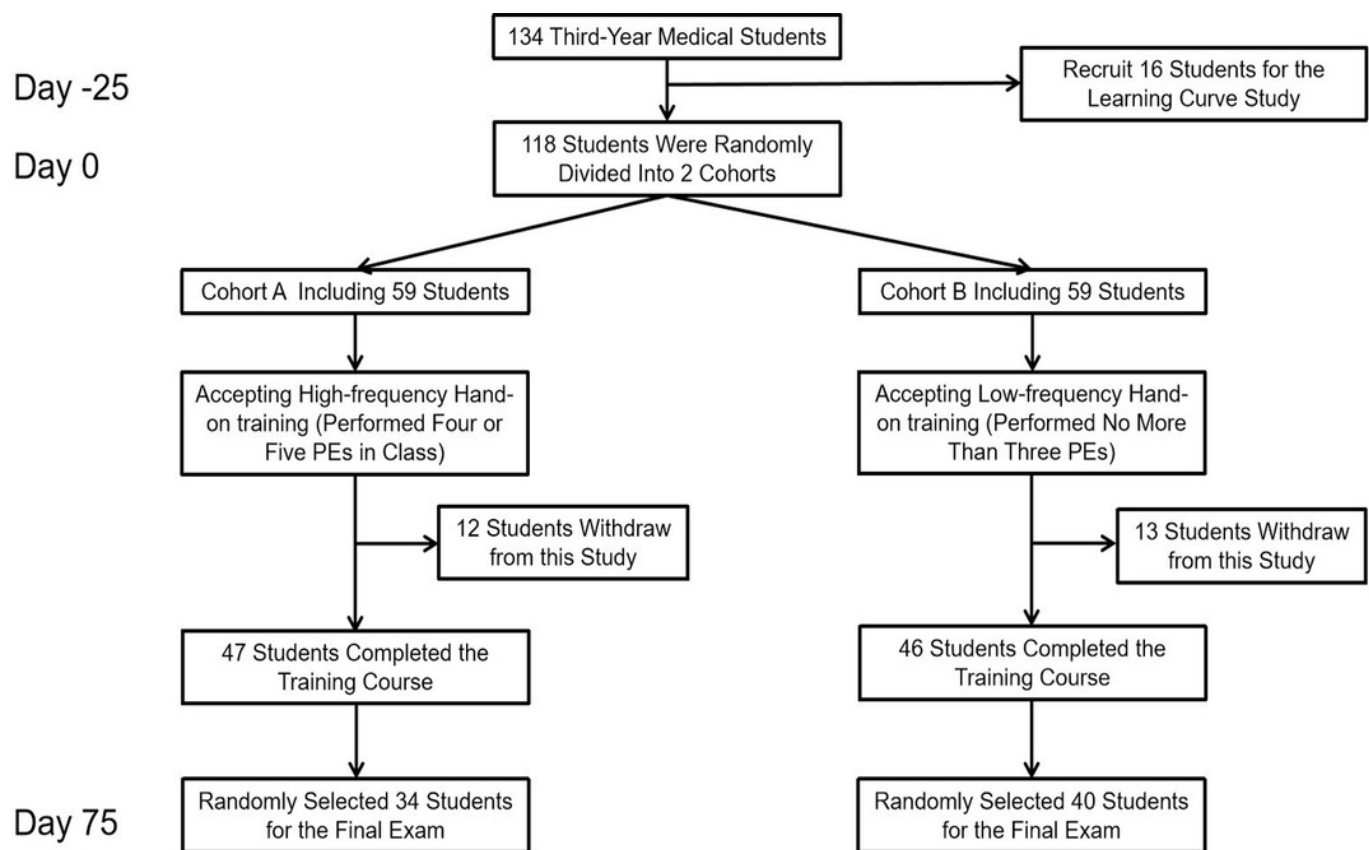
Figure 2

Figure 2

The relationship between the practice cycles and the percentage of students who reached the minimum passing score. The X-axis represents the number of cycles, and the Y-axis represents the proportion of students who reached the passing score to all enrolled learning curve study students. (* P < 0.05 when compared to the previous cycle.)
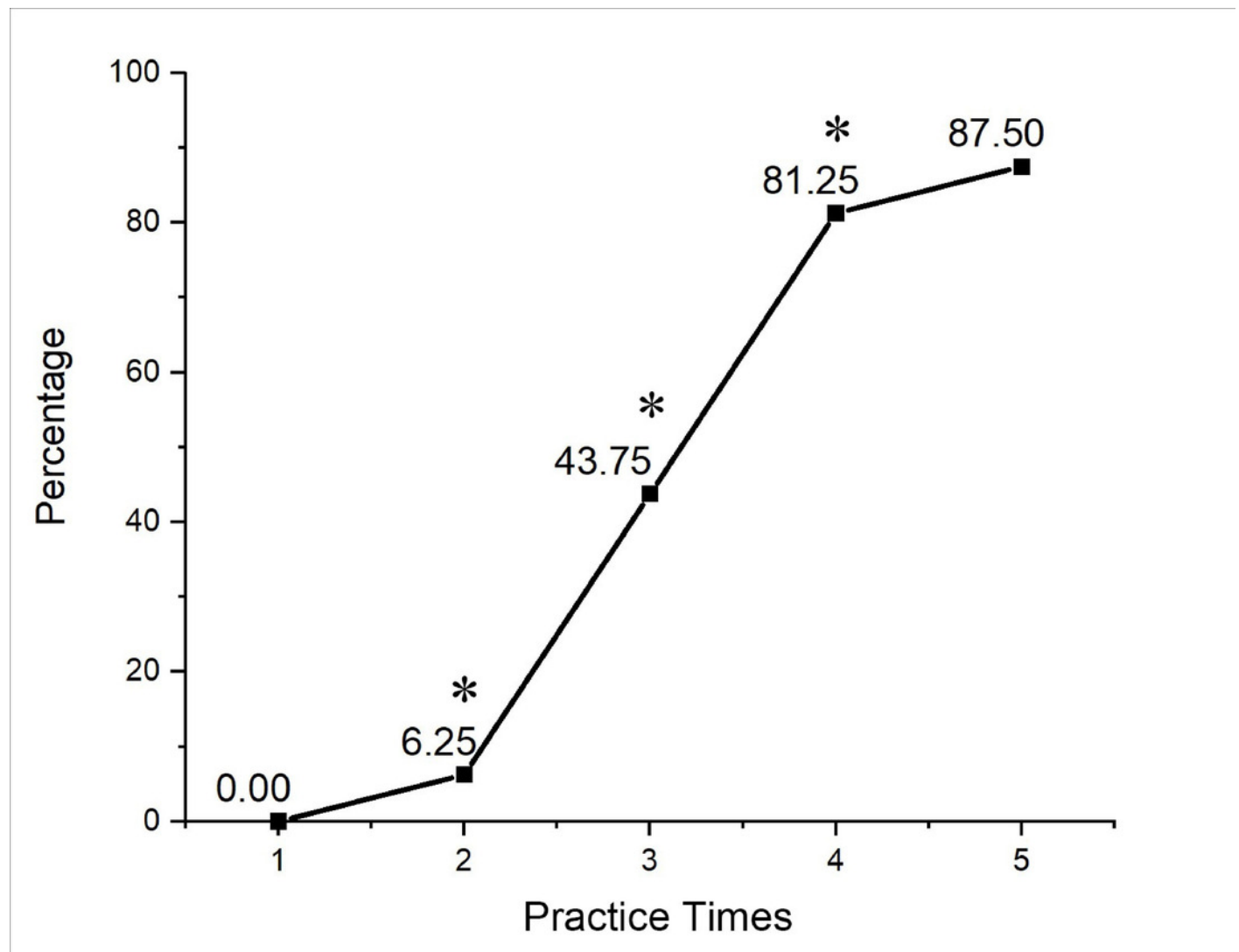
Figure 3

Figure 3

Results of the comparative study of low-frequency hands-on training and high frequency hands-on training. The Y-axis represents the scores of the students' final exam. Students in high-frequency cohort scored higher than that of low-frequency cohort (84.41±11.78 vs 76.83±17.51; P = 0.030).