Jones LS, Paulman LE, Thadani MS, Terracio L. Medical student dissection of cadavers improves performance on on practical exams but not on NBME Anatomy subject exam.

Med Educ Online [serial online] 2001;6:2. Available from URL http://www.med-ed-online.org

# Medical Student Dissection of Cadavers Improves Performance on Practical Exams but not on the NBME Anatomy Subject Exam

Leslie Sargent Jones, PhD[*], Lance E. Paulman, PhD[*] Raj Thadani, MS[†], Louis Terracio, PhD[‡]

[*] Department of Developmental Biology and Anatomy, University of South Carolina School of Medicine
[†] The National Board of Medical Examiners
[‡] School of Dentistry, New York University

***Abstract*** - We have examined whether cadaver dissection by first year medical students (MIs) affected their performance in two test measures: the NBME Gross Anatomy and Embryology Subject Exam (dissection-relevant questions only), and practical exams given at the end of each major section within the course. The dissections for the entire course were divided into 18 regional dissection units and each student was assigned to dissect one third of the regional units; the other two-thirds of the material was learned from the partner-prosected cadavers. Performance for each student on the exams was then assessed as a function of the regions those students actually dissected. While the results indicated a small performance advantage for MIs answering questions on material they had dissected on the NBME Subject Exam questions relevant to dissection (78-88% of total exam), the results were not statistically significant. However, a similar, small performance advantage on the course practical exams was highly significant.

There has been much debate over the years on the value to the students of cadaver dissection in medical student Human Anatomy courses. Some studies have indicated that students learn anatomy as well by studying prosections as they do by traditional dissecting,[1,2,3,4] while at least one study reports some modest advantage for students performing traditional dissection over studying prosections as a means to learn anatomy.[5] The present study was designed to answer the question of whether cadaver dissection by medical students improves their performance on subsequent exams, either using a laboratory based practical exam, or relevant questions from the multiple choice, NBME Gross Anatomy and Embryology Shelf Exam.

Most of the existing studies provide evidence that dissection itself does not seem to provide an advantage to students in learning anatomy. In the study at the Medical College of Georgia, Bernard[2] compared the performance of medical students doing the traditional all-students-in-the-group dissection with that of an experimental group of students who took turns doing the dissections, with two dissectors teaching six other non-dissectors; the results illustrated that the experimental group had several advantages: they saved time, performed as well or slightly better on practical exams, and performed slightly better on the NBME Subject Exam, though the latter advantage was not significant.

In a study where half the students dissected the upper limb, while the other half dissected the lower limb simultaneously, and then each group taught the other half of the students on their already-dissected material, Peppler et al.[3] reported that over a four year period they found that there was generally little difference in the performance of each group on questions relating to either area, either on practicals, sectional exams, or the NBME Step 1 exam.

Finally, in an extremely interesting series of studies, Nnodim[6,7] and Nnodim et al.[4] compared prosection study to dissection as methods for teaching anatomy, and followed up with studies on retention of learned material five years after the students took the course. In a 1990 study, the students either studied using traditional dissection, or were assigned to an experimental group that studied only faculty-prosected material. In addition to the time saved for the prosection-only students, they also outperformed the traditional-dissection group on both practical and written tests. In a follow-up study, these authors managed to do a re-assessment of students five years after they had taken the anatomy course; half of the students had studied using traditional dissection, while half had been assigned to the experimental, non-dissecting group (studying prosections) five years before. These students were tested using a practical, an oral, and a written format, and the performance of both groups was equivalent overall, with the caveat that the non-dissecting students were somewhat superior in their recall as determined by "some qualitative considerations."[4] In an additional study in 1997 by the same group[7], the classes were divided up so that half continued with a traditional all-students-dissecting protocol, while an experimental group took turns dissecting with the dissection team in a design similar to our own in the present study and to that of Yeager.[5] In that case, the students participated in a rotating dissection schedule with peer-teaching of the non-dissectors, and those in the peer-teaching protocol performed better than the traditional, all-dissecting control students, both on practical exams and on a written exam.[7] This suggests a point echoed in the work of Yeager[8] that it may be the *teaching* of the dissection that imparts what advantage there is to the dissectors in this protocol, and this in turn gives the group as a whole a slight performance advantage. Our own results here support this observation as well.

For evidence on the other side of the debate, another study also from Yeager[5] at St. Louis University used a teaching design similar to our own: their students were divided into four groups, with one fourth dissecting at a time, while the other ¾ learned the anatomy from the material dissected by the other students. In this case, a small, but not significant, advantage on written tests for the dissectors in answering questions over the regions they dissected was reported.

Our own program of teaching anatomy is similar in design to that described by Yeager[5], in that we have the students each performing a third of the dissections and teaching them as prosecutions to their peers. In light of the data supporting the success of non-dissectors as anatomy students, it interested us to know whether there was a difference in the performance of our groups of students. Specifically, we have compared the performance of students dissecting each region with the performance of the two-thirds of the students who did not dissect that particular region; we compared their performance on both intramural practical exams, and on the extramural, NBME Gross Anatomy and Embryology Subject Exam (using the 78-88% of the exam that relates to dissection).

Preliminary reports of these findings have been presented previously.[9,10]

## Methods

**Course Logistics -** The Medical Embryology and Gross Anatomy course at the University of South Carolina School of Medicine is taught in the first semester, first year of the students' curriculum. The class contains 72-75 first year Medical Students (MIs). The students receive lectures as a group (average of five hours/week), and are then scheduled for nine hours of laboratory time. In a traditional curriculum, all 75 students would be in the laboratory for all nine hours, every week, for the full 17 weeks of the semester. However, in our curriculum, the students are assigned to dissect, and required to be in the laboratory, on a rotating schedule. This works as follows: all 75 students are assigned alphabetically (by last names, A-Z) to 12 dissection tables in the laboratory, thus, there are six (or seven) students at each table. The six students are divided into three dissection groups (A, B, and C), with two (or three) students per group; each table has 2-3 representatives from each of the three dissection groups. The pair assignments for the dissection teams are also alphabetical. The only exception to this is that family members are not paired as dissecting partners (e.g., siblings, spouses). One third of the class belongs to each of the three groups, and there are 36 teams of dissecting pairs: 12 A's, 12 B's, and 12 C's. These three teams take turns dissecting according to the rotating schedule (Fig. 1). Since there can be no certain way to predict the weak students, no attempt is made to do so, and no changes are ever made to pairings once the course is under way, unless a student drops out of school. In this case, a student from a three-member team at another table will be moved over to fill the slot and complete the pair.

The dissections for the entire course are divided into 18 topics, and each of the three groups of dissecting teams dissects one third, or six, of the topics. The dissection topics take about a week each, and the dissecting teams take turns dissecting on this rotating basis. After dissecting, each student, along with their team partner, is responsible for teaching the material in that dissection topic to the other four MIs at their table. The other 2/3 of the class not actively dissecting are left with free time either to work on their case study (*vide infra*) with their small group, or to study the material being dissected using the other learning tools available. They are also free to visit the dissection laboratory at any time, and may stay and watch the entire dissection if they chose. They are not encouraged to do this if they are disrupting dissection progress, however, and most students elect to visit the laboratory only at the end of each day's dissection, to see the result of that session's work,

**Ancillary Teaching Materials -** In our course there are several additional tools available to the students for the study of anatomy. These include: didactic lectures, textbooks and atlases, prosections (prepared by the fourth year medical student (MIV) teaching-assistants; 2 prosected cadavers were available for study and were used for exams), computer assisted learning programs (e.g., A.D.A.M.© & Human Anatomy© (Gold Standard), used at the students' discretion), biweekly radiology tutorials, small group case studies (two/course), and access to additional learning tools (bones, models, cross-sections, video dissections). In a previous study we reported how students ranked these tools.[1] The learning aids were ranked as follows (most to least useful): atlases, cadavers, lectures, textbooks, bones, teaching assistants, models/X-sections, case studies, computer programs, radiology tutorials, video tapes.

**Dissection Protocol -** The three, rotating dissection teams were assigned to the dissection topics as shown in Figure 1. The third of the class dissecting rotated after every dissection topic (e.g., "Superficial and Deep Back," or "Posterior Abdominal Wall," or "Orbit, Eye, & Ear"), and the next group took over for the next dissection topic. Every dissection team (two students) performed a total of six dissections, which they were then responsible for teaching to their group (four other students), and graded both for quality of dissection and effectiveness of teaching during a 30 minute Teaching Session (see below). On average, nine hours/week of assigned dissection time was scheduled, with both faculty (usually two) and teaching assistants (usually three) available in the lab (one for every two to three tables). The teaching assistants were fourth year Medical Students doing four-week rotations full-time in a course entitled "Clinical Anatomy." The course allows them to relearn anatomy by performing prosections and teaching, as well as to learn how to teach, examine, evaluate, and mentor. The system and value of the MIV training program has been reported previously.[11]

**Figure 1. Groups and Assigned Dissections**

| Group A | Group B | Group C |
|---|---|---|
| **Dissection 1**<br>pectoral region, breast, thoracic wall, superior mediastinum | **Dissection 2**<br>superior mediastinum, heart | **Dissection 3**<br>lungs, posterior mediastinum |
| **Dissection 4**<br>anterior abdominal wall, inguinal region | **Dissection 5**<br>abdominal cavity, peritoneum, abdominal viscera | **Dissection 6**<br>posterior abdominal wall |
| **Dissection 7**<br>anal triangle, urogenital triangle | **Dissection 8**<br>pelvic viscera | **Dissection 9**<br>face, parotid, scalp, craniotomy, cranial cavity |
| **Dissection 10**<br>orbit, eye, eyelids, ear (by prosection, but must teach) | **Dissection 11**<br>anterior & posterior triangles of the neck, suboccipital triangle | **Dissection 12**<br>temporal & infratemporal fossa, TMJ, prevertebral region, pharynx, larynx, nasal cavity |
| **Dissection 13**<br>pterygopalatine fossa, oral cavity, submental triangle | **Dissection 14**<br>superficial back, deep back, laminectomy, spinal cord (exterior) | **Dissection 15**<br>axilla, shoulder, brachial plexus |
| **Dissection 16**<br>arm, forearm, hand, upper limb joints | **Dissection 17**<br>thigh, gluteal region | **Dissection 18**<br>leg, foot, lower limb joints |

The three dissection groups (A, B, and C) with their assigned dissections are listed in order.

**Teaching Session protocol -** Teaching Sessions were formally scheduled, half hour sessions where the two dissecting partners demonstrated all structures from that dissection and 'taught' them to their four tablemates. In addition to the two dissectors teaching on the cadaver, the other four students also had assigned topics to teach, but were not graded on their contribution. These topics included the related osteology, radiology, clinical correlations, embryology, etc. These four students were also expected to be well prepared about the anatomy being demonstrated, to have visited the dissectors regularly to monitor the dissection as it progressed, and to be able to answer questions as the dissectors pointed out revealed structures.

A faculty member or TA was present during the entire 30 minute session, and graded both the quality of the teaching and the quality of the dissection at that time. A grade was assigned to each dissection partner individually for their teaching (out of 10 points), while the dissection received a grade shared by the dissection partners (out of 10 points). The graders and their tablemates had a check list of structures/topics to be demonstrated and discussed for the teaching session, and a similar list of structures to be checked for the dissection grading. The students were given a comprehensive list of all structures to be dissected and all topics to be covered in the teaching session generated by the faculty, however, the specific subset of structures to be graded each time was not available to the students. If for any reason a teaching session was clearly substandard, and the group did not recognize the problem, the grader was instructed to intervene, clarify the material, and instruct the group to perform the entire session again at a later time; this may have occurred once since the entire program was started.

**Exam protocols -** The NBME Gross Anatomy and Embryology Subject Exam was administered at the end of the semester as the students' only final exam in Gross Anatomy; it counted as 13% of their final grade. The exam consisted of 114 graded questions that were coded as A, B, or C, indicating which group dissected the material tested by that item. Questions not relating to a dissection were removed from analysis (e.g., early embryology). This analysis was performed for two years of first year classes (first year students of 1997, 1998).

The practical exams were timed (75 seconds per item), 50-question identification exams with half of the questions being two-part questions. Examined material included tagged cadaver items, in addition to radiology, cross-section, and osteology questions. The exam items were keyed according to which dissection they were covered in and which of the three groups (A, B, or C) did the dissection. Four practical exams were administered during the one semester Medical Embryology and Gross Anatomy course, one after each of the four major course sections: Back & Thorax; Abdomen, Pelvis, & Perineum; Head & Neck; and Limbs. The practicals together represented 40% of the student's total grade in the course. The coding of the questions in both test formats allowed for a comparison between MI performance on questions relating to areas they **had** dissected and MI performance on the questions over material they had **not** dissected. Questions on each exam that were not strictly from dissected material (e.g., radiographs, cross sections, osteology questions) were eliminated from the data collected. The data were collected for the first year students of 1998.

**Statistical analysis**

**Practical Exams -** For any individual student, the comparison is unbalanced, as they only dissected 1/3 of the material and thus the analysis could not be split evenly among the total of 200 questions (50 questions/exam X 4 exams), but was a comparison of approximately 66 questions on dissected topics vs. 134 undissected topics. Data was collected as percent correct of the total available questions on dissected or non-dissected material (N=75 students). A paired sample t-statistic was used to test for statistically significant differences in performance on dissected versus undissected items within each of the three groups.

**NBME Subject Exams -** In 1997, there were 31, 41, and 28 items relating to body regions dissected by groups A, B and C respectively (88% of the total number of NBME Subject Exams questions). Fourteen items were not related to a dissected region (e.g., embryology) and were removed from analysis. In 1998, there were 28, 31, and 38 items dissected by groups A, B, and C respectively (78% of the total number of NBME Subject Exams questions). There were 28 items not related to a dissected region and they were removed from analysis. Thus, each student (1997, N=73; 1998, N=75) dissected roughly one third of the items included in the analysis. Performance on these items was compared with performance on the remaining undissected items using a paired sample t-statistic to test for statistically significant differences in performance on dissected versus undissected items.

**Results**

**Practical Exam Results -** There was a statistically significant difference between the performance of the Dissectors (performing better) and the Non-dissectors in the intramural practical exams, with the Dissector group earning a higher average grade (Total for all exams, Mean$\pm$S.D.: Dissectors 88.64$\pm$

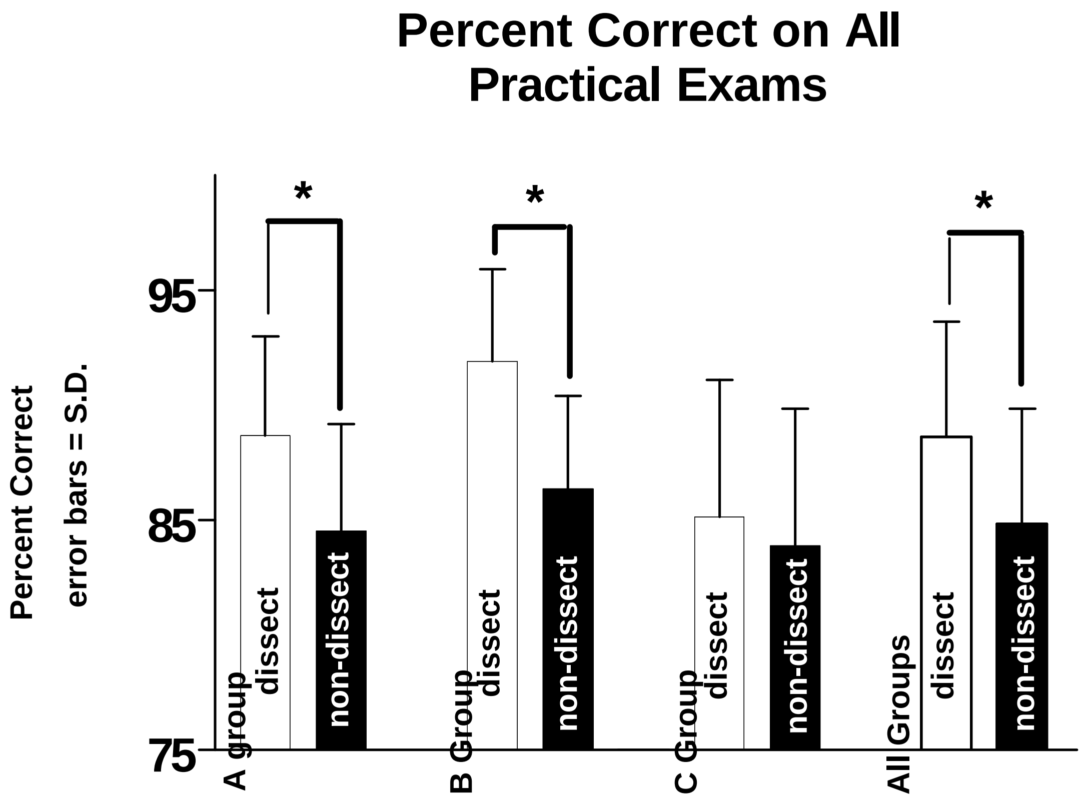


**Figure 2. Results for Practical Exam Performance**

Unfilled bars represent percent of correct answers given by students over material they had dissected, while black bars represent percent correct answers given by the same students over material they had not dissected. The data is shown broken down into the three groups of dissection teams, and as the grouped, total data for all the groups combined (the whole class). Significant differences ($P \leq .01$) are indicated by asterisks. Note that these results were obtained only for the class taught in 1998, and thus the "All Groups" from this graph can be compared to the "1998" bars in Figure 3; they are data collected from the same students.

5.75; Non-dissectors, 84.85±5.77, $p \leq .01$; see Fig. 2). The individual data for two of the three dissection groups (A and B) were also significantly different (Group A: Dissectors 88.68±8.6; Non-dissectors 84.53±9.3; $p \leq 0.01$; Group B: Dissectors 91.3±7.98; Non-dissectors 86.36±8.1; $p \leq 0.01$), but an effect of dissection was not seen in the Group C students (Dissectors 85.1±11.9; Non-dissectors 83.9±11.9; see Fig. 2).

Note that there was some difference between the three groups as a whole (Dissectors + Non-dissectors), with the B group performing on average the best (88.66±6.6) and significantly better than either Group A (85.87±7.5; $p \leq .01$) or Group C (85.44±9.3; $p \leq .01$), while those two groups were not different from each other.

**NBME Subject Exam Results for Dissection-Relevant Questions** - For this section of the study, the major finding was that there was not a statistically significant advantage for Dissectors vs. Non-dissectors in their performance on the questions relevant to dissection in the NBME Gross Anatomy and Embryology Subject Exam either by individual year (1997: Mean±S.D Dissectors, 67.7±12.4, Non-dissectors, 65.3±10.6; 1998: Dissectors, 70.3±11.1, Non-dissectors, 69.8±9.7), or combined as a total (Both years: Dissectors, 69.1±11.8; Non-Dissectors, 67.6±10.4; see Fig. 3).

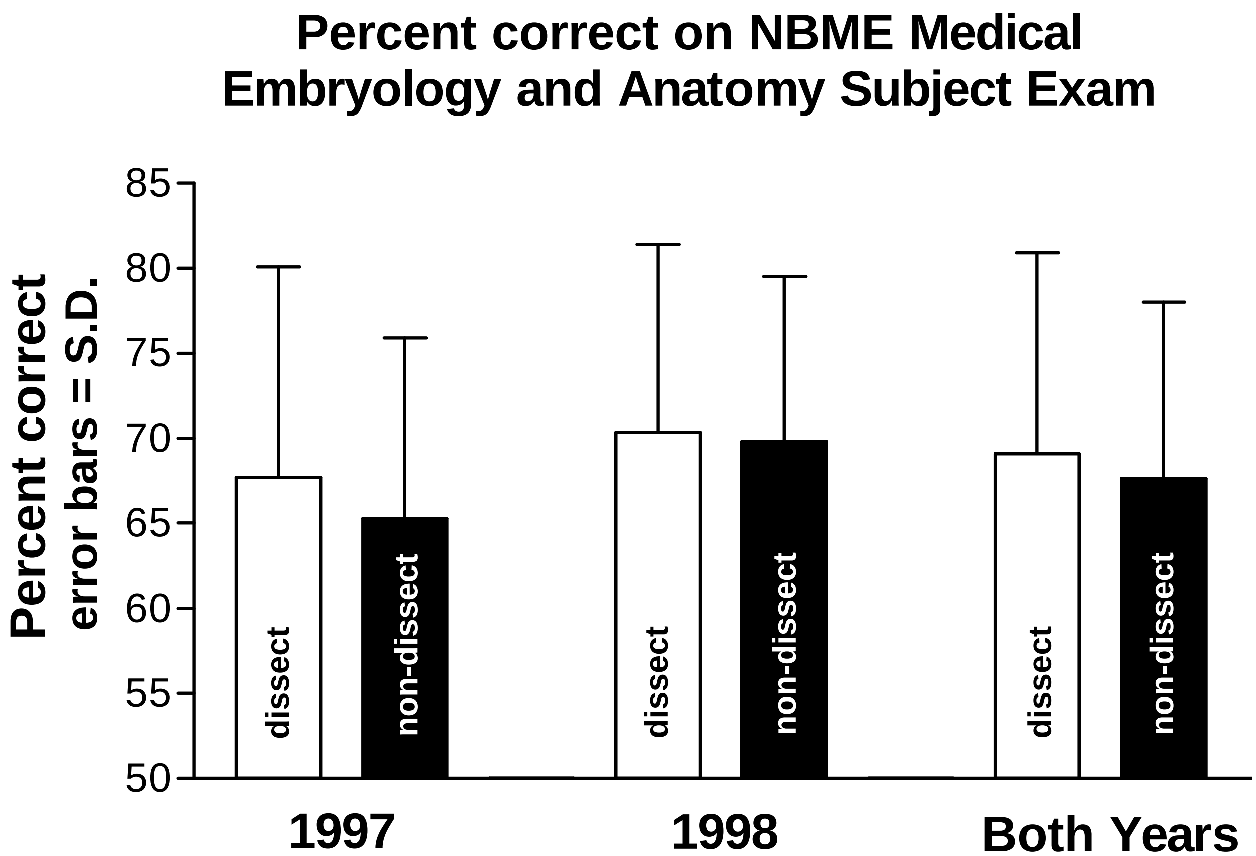


**Figure 3. Results for Dissection-relevant Questions on the NBME Medical Embryology and Anatomy Subject Exam Performance.**

Results are illustrated for both years examined, and as grouped data representing totals for both years. The unfilled bars represent percent of correct answers given by students over material they had dissected out of the 88% (1997) or 78% (1998) of the total questions on the exams; black bars represent percent correct answers given by students over material they had not dissected. Error bars represent standard deviation. Note that the "All Groups" bar from Figure 2 represents the results from the same class illustrated in this figure as the "1998" data.

When the performance of all students in each of the groups were compared (Dissectors + Non-dissectors, A vs. B vs. C), there were no differences between any of the three groups for either of the two years studied.

## Discussion

The major finding of this two-pronged study was that dissection provided a small advantage to students being tested *on the cadavers.* However, this difference did not translate into a performance advantage on the NBME Gross Anatomy and Embryology Subject Exam questions relevant to dissection (78-88% of the total exam). The implications for the debate on the need for cadaver dissection in teaching medical gross anatomy are that the value of the tests themselves requires assessment. If it is the case that the NBME test is the best reflection of student learning and subsequent performance in later, clinical rotations, then there is no need for the substantial investment in running a cadaver laboratory; the present results indicate that dissection itself does not give the students a performance advantage on what is learned for and then tested by that exam. However, if the cadaver-based practical exams are a better measure of student understanding and learning of the three-dimensional concepts taught in medical gross anatomy, then the data illustrate a small but consistent advantage for students performing the dissections and then teaching them.

These results only fuel the debate on the importance of dissection, and focuses attention on what test is used to assess the students' knowledge: the test is critical, and not all tests will address the pedagogical question asked. As reported by Stanford et al.[12] previously, the impact of dissection on learning depends on how you measure the learning. In that study, the students were studying cardiac anatomy using cadaver dissection augmented by a computer-based program (with control groups getting either no additional instruction, dissection only, or the computer-based program only). The students were tested using two tests, one on cadaver specimens (analogous to our practical exam) and the other asking questions on computer tomographic images. Their results indicated that the combination of both teaching aids (dissection and computerized images) resulted in students performing the best overall, but the various permutations of teaching approaches did not produce a consistent, predictable difference in performance on the two types of tests. They found that the computer-based instruction (which utilized tomographic images) did enhance performance on the test using such images, but still concluded that "computer instruction should not replace dissection for teaching cardiac anatomy."

While it may seem evident that the type of test will determine much about the results, it is not so evident which testing instrument is <u>really</u> assessing anatomical knowledge. Without a consensus on the best method of testing, we are left with no way of determining what aspects of our teaching program <u>are</u> best for student learning. If the goal of the course is to produce students who perform well on the NBME-style tests, then dissection is of little value. However, if one believes that only a practical exam really tests student knowledge of anatomy in a manner relevant to subsequent, clinical needs, then dissection does have a positive, though modest impact, on the learning of most students.

Taken together these results provide support for the case to reduce the role of cadaver dissection in the teaching of Gross Anatomy, particularly if the institution places great weight on NBME performance. Even with using the practical exams as the standard of performance, however, the act of dissection confers only a modest advantage for the students dissecting, and the teaching of the material by the students may be of as much value as the actual, hands-on dissection. In an era of course-size reduction in medical schools, and the need to contain costs, the conclusions here should encourage curriculum design that explores additional mechanisms of study that may reduce the students' dissection time while increasing their time spent in actual study.

**Acknowledgements**

The authors thank Dr. Jeffrey Patton for his comments on the manuscript. Supported by the Department of Developmental Biology and Anatomy, University of South Carolina School of Medicine, and the National Board of Medical Examiners.

**References**

1. Jones LS, Welsh MG, Terracio L. First year medical students' views on computer programs:" give us our teaching assistants." FASEB Journal 1998;12:5635

2. Bernard GR. Prosection demonstrations as substitutes for the conventional human gross anatomy laboratory. J Med Educ. 1972;47:724-8.

3. Peppler RD, Kwasigroch, TE Hougland, MW. Evaluation of simultaneous teaching of extremities in a gross anatomy program. J Med Educ.1985;60:635-9.

4. Nnodim JO, Ohnaka EC, Osuji CU. A follow-up comparative study of two modes of learning human anatomy: by dissection and from prosections. Clin Anat. 1996;9:258-62.

5. Yeager VL. Peer teaching in gross anatomy. J Med Educ 1981;56:922.

6. Nnodim JO. Learning human anatomy: by dissection or from prosections? J Med Educ. 1990;24:389-95.

7. Nnodim JO. A controlled trial of peer-teaching in practical gross anatomy. Clin Anat. 1997;10:112-7.

8. Yeager VL Learning gross anatomy: dissection and prosection. Clin Anat. 1996;9:57-9.

9. Jones LS, Paulman LE, Terracio L. Does cadaver dissecting by students affect their gross anatomy practical exam performance? FASEB Journal 1999;13:770.4.

10. Terracio L, Thadani R, and Jones LS. Does the process of dissecting by students affect their performance on the NBME Gross Anatomy and Embryology final? FASEB Journal 1999;13:770.5.

11. Terracio L, Welsh MG, and Jones, LS. Simultaneous teaching of first year and fourth year gross anatomy: a win-win situation. FASEB Journal 1995; 9:6162.

12. Stanford W, Erkonen WE, Cassell MD, Moran BD, Easley G, Carris RL, Albanese MA Evaluation of a computer-based program for teaching cardiac anatomy. Invest Radiol. 1994;29:248-52.

**Corresponding Author:**

Leslie Sargent Jones, Ph.D.
Department of Developmental Biology and Anatomy
University of South Carolina
School of Medicine
Columbia, SC, 29208
Phone: 803-777-8005
FAX: 803-777-2214
Email: leslie.jones@schc.sc.edu