**Characteristics of included original studies.**

|  **Component** |  |  |
| --- | --- | --- |
|

|  |  |
| --- | --- |
| Author | Thom Mitchell L., 2021 |
| Study title | Is remote near-peer anatomy teaching an effective teaching strategy? Lessons learned from the transition to online learning during the Covid-19 pandemic. DOI: 10.1002/ase.2122 |
| Study design | Quasi experimental |
| Objective | 1) Evaluate the effectiveness of the remote near-peer anatomy curriculum relative to the equivalent in-person learning environment using student assessment scores.2)Collect student perspectives on the online near-peer learning environment to identify its limitations and potential areas of improvement. |
| Setting | Cleveland, Ohio, USA |
| Location where the sample was obtained  | Case Western Reserve University School of Medicine. |
| Eligibility criteria  | All students enrolled in Anatomy classes of 2022 to 2024. \*Criteria are not shown clearly. |
| Was sample size calculation performed? | No |
| Number of participants recruited | 552 |
| Number of participants eliminated | 0 |
| Number of participants included in the study | 552 |
| Recruitment period | 1 Single period. Cluster study. 3 groups recruited at the same time. |
| Sampling type | Cluster sampling |
| **Methods** |
| Measurement quality of the variables |  |
| Variable definitions  | Effectivity of an online near-peer anatomy: Not defined conceptually.Student perception to the same: Satisfaction with the curriculum. |
| How were the variables measured?  | Effectivity of an online near-peer anatomy: Questionnaire titled "Gross Anatomy and Radiology Questions (GARQs)"Student perception to the same: Anonymous survey given at the end of the experiment. |
| Was bias control performed? | There was no bias control. |
| Was normality assessment performed? | Shapiro Wilk test was performed to evaluate normality.Levene's test was performed to evaluate Homogeneity of variances. |
| Statistical analysis performed  | Nonparametric tests: Kruskal Wallis to evaluate 3 cohorts. Chi-square analysis of subgroups. Likert scale evaluation survey was compared only in class 2023 vs 2024 using two-tailed independent samples t-tests. |
| Was confounding variables controlled? | There was no analysis of confounding variables  |
| **Results** |
| Demographic results of the sample  | Class 2020 n = 185, female 51.3%, male 48.7%Class 2023 n = 184, female 51.4%, male 48.6%Class 2024 n = 183, female 58.3%, male 41.7% |
| Results | Reliability of questionnaires, Cronbach’s Alpha 0.70-0.78 Performance of the cohorts by mean assessment scores: * 2022, Mean (SD) 93.64 (±5.86)
* 2023, Mean (SD) 93.75 (±4.09)
* 2024, Mean (SD) 92.04 (±4.83)
* Kruskall Wallis between 3 groups: p<0.001; post hoc 2022-2023 p=0.157; pot hoc 2022-2024 p <0.001; post hoc 2023-2024 p <0.001.
	+ Post hoc power analysis β=0.82.
	+ Effect size=η2=0.03 (small effect)
* Reliability of survey: α = 0.6 for the class of 2023; α = 0.8
* Difference between survey means:\*t-test 2023-2024, p=0.008, Estimators not shown.
 |
| Author conclusions | The online translation of the introductory anatomy curriculum showed a significant decrease in student assessment scores relative to prior in-person cohorts. However, the effect size was small. So, the authors conclude that near-person teaching method is comparable to the in-person format.  |
| Revisor commentaries | Well designed, however, there was several confounding factors that were not measured although those variables were difficult to control.  |

 |  |

|  **Component** |  |  |
| --- | --- | --- |
|

|  |  |
| --- | --- |
| Author | Antonopolous Ioannis, 2022 |
| Study title | Students’ perspective on the interactive online anatomy labsduring the COVID‑19 pandemic. doi.org/10.1007/s00276-022-02974-z |
| Study design | Transversal analytic study |
| Objective | Evaluate the students’ opinions on the efficacy and utility of the interactive online anatomy laboratory. |
| Setting | Athens, Greece |
| Location where the sample was obtained  | Not mentioned |
| Eligibility criteria  | Not mentioned |
| Was sample size calculation performed? | No |
| Number of participants recruited | 160 |
| Number of participants eliminated | 0 |
| Number of participants included in the study | 160 |
| Recruitment time | 02 June 2020 |
| Sampling type | Convenience sampling |
| **Methods** |
| Measurement quality of the variables | Self-assessment, Cronbach's alpha, knowledge 0.79, gain in knowledge 0.84, promotion of interest 0.63, intrinsic motivation 0.74, pandemic concerns 0.66. Internal consistency of knowledge test at T1 Cronbach's alpha 0.73 score image questions, 0.52 text based questions, 0.78 for all questions.  |
| Variable definitions  | Web questionnaire  |
| How were the variables measured?  | Student's opinion |
| Was bias control performed? | Questionnaire |
| Was normality assessment performed? | Not mentioned |
| Statistical analysis performed  | Outcome was categorical |
| Was confounding variables controlled? | ANOVA and Wilcoxon test |
| **Results** |
| Demographic results of the sample  | n=160; female n=97, male n=63. Group 1 n=58, Group 2 n= 61, Group 3 n=41 |
| Results | One way ANOVA between 3 groups: * Score\_1 was significant different between groups, p values not shown.
* Pairwise comparisons of means test show difference between group 1 and group 2 p val <0.05.
* Score\_2 shown significant difference between group 2 and group 4 p = 0.05.
* Score\_3 didn't differ among 3 groups.

Spearman correlation between scores:* Score\_1 - Score\_2, ρ = – 0.3, p value < 0.01.
* Score\_2 and Score\_3 (ρ = 0.6, p value < 0.01)
* No significant correlation was detected between Score\_1and Score\_3. Statistics not shown.
 |
| Author conclusions | The creation of dissection educational videos enabled the continuation of laboratory teaching using cadaveric material during the pandemic.Most of the total number of respondents found the ONALs attendance beneficial for their studying. |
| Revisor commentaries | n=160; female n=97, male n=63. Group 1 n=58, Group 2 n= 61, Group 3 n=41 |

 |  |

|  **Characteristic** |  |  |
| --- | --- | --- |
|

|  |  |
| --- | --- |
| Author | Schulte Henri, 2022 |
| Study title | Teaching gross anatomy during the Covid-19 pandemic: Effects on medical students’ gain of knowledge, confidence levels and pandemic related concerns. doi.org/10.1016/j.aanat.2022.151986 |
| Study design | Quasi experimental |
| Objective | Evaluate the effect of a six-hour dissection course in groups of two students to acquire knowledge and understanding of the head-neck region |
| Setting | Hannover Medical School |
| Location where the sample was obtained  | First-yearmedical students at Hannover Medical School |
| Eligibility criteria  | Not employed |
| Was sample size calculation performed? | No |
| Number of participants recruited | 281 |
| Number of participants eliminated | 166 |
| Number of participants included in the study | 115 |
| Recruitment period | June 2020 |
| Sampling type | Convenience Sampling |
| **Methods** |
| Measurement quality of the variables | Self-assessment, Cronbach's alpha, knowledge 0.79, gain in knowledge 0.84, promotion of interest 0.63, intrinsic motivation 0.74, pandemic concerns 0.66. Internal consistency of knowledge test at T1 Cronbach's alpha 0.73 score image questions, 0.52 text based questions, 0.78 for all questions.  |
| Variable definitions  | Knowledge on the anatomy neck section |
| How were the variables measured?  | Validated questionnaire, anonymous survey |
| Was bias control performed? | Not mentioned |
| Was normality assessment performed? | Not mentioned |
| Statistical analysis performed  | Kendall's Tau, Cronbach's alpha, Student's t test, Mann-Whitney Rank sum test. 2-way ANOVA.  |
| Was confounding variables controlled? | Not mentioned |
| **Results** |
| Demographic results of the sample  | Knowledge test difference: G1 - G2: (59.9 % ± 20.1 for G1 vs. 52.6 % ± 15.4 for G2, p = 0.232). Complete knowledge test. Text based questions: G1-G2: (53.3 % ± 19.9 vs. 42.9% ± 16.7, p = 0.09). G1 T1-T2 (1st vs 2nd measure): p=0.062G1 rated their confidence levels significantly higher (p = 0.049) than G2. Self-assessment not significative.  |
| Results | After a period of self-directed learning using a diversity of eLearning resources, a brief and intensive cadaver dissection course of the head-neck region in groups of two students did not result in better scores in the knowledge test compared to a control group.Although the knowledge test scores in the dissection group were not higher compared to the control group, students had a higher degree of self-confidence when answering knowledge questions correctly after the active cadaver dissection. |
| Author conclusions | Besides the multiple statistic test, no significant results were found.  |
| Revisor commentaries | None |

 |  |

|  **Characteristic** |  |  |
| --- | --- | --- |
|

|  |  |
| --- | --- |
| Author | Harrell Kelly M, 2021 |
| Study title |  |
| Study design | Crashing from Cadaver to Computer: Covid-driven Crisis-mode Pedagogy Spawns Active Online Substitute for Teaching Gross Anatomy.Doi: 10.1002/ASE.2121 |
| Objective | Quasi experimental |
| Setting | The aim of this study is to (1) describe the development and implementation of a novel online anatomy laboratory teaching and assessment package for reproductive anatomy using multi perspective videography and online learning principles so that other anatomy educators may consider implementation; and (2) assess the effectiveness of this online teaching and assessment package as it relates to student examination performance and overall perception. |
| Location where the sample was obtained  | Richmond Virginia |
| Eligibility criteria  | Virginia Commonwealth University |
| Was sample size calculation performed? | Not mentioned. Apparently, all students from a course in the period of January to May 2020 |
| Number of participants recruited | No |
| Number of participants eliminated | 184 |
| Number of participants included in the study | Not mentioned |
| Recruitment period | Not mentioned (Apparently n = 184) |
| Sampling type | January 2020 to May 2020 |
| **Methods** |
| Measurement quality of the variables | Kuder Richardson Formula 20 |
| Variable definitions  | Course engagement/satisfaction.Learning/knowledge. |
| How were the variables measured?  | Didactic examination consisted of knowledge-based, comprehension-based and application based questions. Practical examination used video sequences of dissected specimens derived from similar footage used in the laboratory dissection demonstration videos and video conference sessions.  |
| Was bias control performed? | Not mentioned.  |
| Was normality assessment performed? | No normality tests were performed. |
| Statistical analysis performed  | ANOVA, two sample t tests, Hedge's g effect size was measured |
| Was confounding variables controlled? | Not mentioned. |
| **Results** |
| Demographic results of the sample  | n = 154, female n = 98 (53%), male n = 86 (47%)Didactic test: students 2020 average 89.8/% (±12 SD); 2015 to 2019 mean 88% (± 2.5 SD).Practical test: 2020 94.9 (SD not mentioned); 2015 to 2019 average 82.5 (±2.4 SD) |
| Results | Differences in didactic scores: 2015 to 2019 and 2020 ANOVA p = 0.63; 2019 - 2020 t-test p = 0.25Differences in practical scores: 2015 to 2019 and 2020 ANOVA p=0.003; post hoc comparison 2020 - each year from 2015 to 2019 (p=0.006 or less, real values not shown)Effect size from each t-test of 2020:2015 g = 1.144; 2016 g = 0.794; g = 0.917; 2018 g = 0.937; 2019 g = 1.229 |
| Author conclusions | The results indicate that online teaching of anatomy using multi-perspective videography for laboratory demonstration videos, live/active video conference sessions and a video-based laboratory assessment can enhance measures of student performance and perception of their learning experience. |
| Revisor commentaries | The study lacks methodological quality in essential areas. The statistical assumptions were not considered, and the authors directly started to perform parametrical tests. Individual t-tests were performed even with multiple comparison without adjustment of p levels (Bonferroni's correction), which can lead to increase the risk of type 1 error. |

 |  |