Original Research · *Medical Education*

Meta-analysis & Review of Learner Performance & Preference: Virtual vs. Optical Microscopy

Adam B. Wilson PhD,[1*] Melissa A. Taylor MS,[2] Barbie A. Klein MS,[2] Megan K. Sugrue BS,[2] Elizabeth C. Whipple MLS,[3] & James J. Brokaw PhD[4]

[1]Department of Anatomy & Cell Biology, Rush University, Chicago, IL, USA
[2]Medical Sciences Program, Indiana University School of Medicine, Bloomington, IN, USA
[3]Ruth Lilly Medical Library, Indiana University School of Medicine, Indianapolis, IN, USA
[4]Department of Anatomy & Cell Biology, Indiana University School of Medicine, Indianapolis, IN, USA

Running Title: A Meta-analysis of Virtual Microscopy

*Correspondence to: Dr. Wilson, Department of Anatomy & Cell Biology, Rush University, 600 S. Paulina St., Armour Academic Center Suite 505A, Chicago, IL. USA. 60612. E-mail: Adam_Wilson@rush.edu. Phone: 312-942-5903.

This is the submitted manuscript of the article to be published in its final form as:

Wilson, A.B., M.A. Taylor, B.A. Klein, M.K. Sugrue, E.C. Whipple, and J.J. Brokaw. Meta-analysis and review of learner performance and preference: virtual vs. optical microscopy. Med. Educ. 2016. Preprint.

ABSTRACT

Background & Purpose For nearly two decades, a wealth of literature has been published describing the various capabilities, uses, and adaptations of virtual microscopy (VM). Many studies have investigated the effects and benefits of VM on student learning compared to optical microscopy (OM). As such, this study statistically aggregated the findings of multiple comparative studies through a meta-analysis to summarize and substantiate the pedagogical efficacy of teaching with VM.

Methods Using predefined eligibility criteria, teams of paired researchers screened the titles and abstracts of VM studies retrieved from seven different databases. After two rounds of screening, numerical and thematic data were extracted from the eligible studies for analysis. A summary effect size and estimate of heterogeneity were calculated to determine the effects of VM on learner performance and the amount of variance between studies, respectively. Trends in student perceptions were also analyzed and reported.

Results Of the 725 records screened, 72 studies underwent full-text review. In total, 12 studies were viable for meta-analysis and additional studies were reviewed to extract themes relating to learners' perceptions of VM. The meta-analysis detected a small yet significant positive effect on learner performance (SMD=0.28, [CI=0.09, 0.47], p=0.003), indicating that learners experience marked knowledge gains when exposed to VM over OM. Variation among studies was evident as high heterogeneity was reported. An analysis of trends in learner perceptions noted that respondents favored VM over OM by a large margin.

Conclusions Despite many individual studies reporting non-significant findings when comparing VM to OM, the enhanced power afforded by meta-analysis revealed that the pedagogical approach of VM is modestly superior to OM and is preferred by learners.

INTRODUCTION

The advent of virtual microscopy (VM) began in the 1990's following advances in computer processing and data storage that allowed high-resolution copies of whole glass slides to be digitized, manipulated, and eventually shared over the Internet.[1] Improvements in digital storage, resolution, image compression, and user interfaces continue to strengthen and reaffirm the diverse utility of VM for educational and clinical practice.[2]

VM is commonly reported as being a dynamic digital adaptation of traditional optical microscopy (OM) that mimics the experience of viewing glass slides at a wide range of magnifications.[3] VM provides a digital interface with the ability to focus (if equipped with z-scanning), pan, and magnify through multiple image layers in different focal planes.[4, 5] Through web-based VM platforms, multiple users can simultaneously access microscopic images from any location via an Internet enabled device.[4] For the purposes of this meta-analysis, static presentations of whole glass slides (e.g., fixed images on PowerPoint or recorded video) that could not be manipulated by users were not considered to meet the definition of virtual microscopy.

Throughout the literature, various academic settings have reported adopting VM technology as an educational, research, and clinical diagnostic training tool. As laboratory hours steadily decline at academic institutions,[3] rapid developments in computer-assisted instruction are supplementing traditional pedagogical approaches. As a result, dental schools,[6] veterinary schools,[1, 7] undergraduate human anatomy courses,[8, 9] medical schools,[3, 10-17] and various residency programs [18-20] have adopted VM into their curricula. In the context of research and clinical medicine, VM is increasingly being incorporated into applications such as the morphological screening of hematology slides,[4] automated image analysis of pathologic and

histologic specimens,[21] biorepositories and tissue banks,[21] as well as telepathology for diagnostic consultations.[22]

While the exploration of VM has been widely disseminated across the fields of histology and pathology, this work is the first to review and summarize the effects of VM through a meta-analysis. Meta-analysis, in general, is a powerful method for aggregating statistical data from a large collection of studies.[23] The findings of this work will allow for stronger summative and generalizable conclusions to be drawn about the usefulness of VM within educational settings.

The principal aim of this study was to evaluate the efficacy of VM, compared to optical microscopy (OM), concerning its effects on changes in learners' knowledge acquisition. We hypothesized that the use of VM would have no marked effect on learner performance scores across multiple studies and learner populations. Additionally, we anticipated that the benefits of VM would be acknowledged by learners as demonstrated by a dominance of positive perceptions.

METHODS

This meta-analysis was conducted in accordance with PRISMA guidelines for the reporting of systematic reviews and meta-analyses.[24] Published articles, dissertations, and meeting abstracts were searched between January 1995 and December 2014 on OVID, Web of Science, CINAHL, PsycINFO, EMBASE, ERIC, and Dissertations Proquest & Theses A&I. Key search terms included, for example, virtual/digital microscopy, virtual/digital slides, medical education, and paramedical education. Medical subject headings were also used and included headings such as microscopy, user-computer interface, and education.

Eligibility Criteria

Studies were included for preliminary review if they evaluated the educational effectiveness of VM compared to OM and either 1) reported empirical data on changes in learner performance (e.g., via pretests/posttests or via comparative treatment vs. control designs) or 2) reported on learner/user perceptions. No geographical restrictions were specified and only studies written in English were included. Studies that evaluated the diagnostic accuracy of VM compared to OM were not of interest in this analysis. Similarly, studies that compared VM to OM during test administrations only, did not meet the eligibility criteria. That is, an intervention in which learners were exposed to VM for some duration of time was required.

Study Selection and Data Extraction

Two teams of paired researchers conducted a preliminary screening of study titles and abstracts. Each team screened half of all articles retrieved from the electronic search. Using a crossover design, a secondary screening was performed in which each team of paired researchers evaluated and made decisions on the discrepancies that arose out of the opposite team's preliminary analysis. Decisions regarding inclusion/exclusion discrepancies often required a full article review and were settled by team consensus. This process further refined the number of studies for full-review and ensured agreement regarding the applicability of each study to the goals of the meta-analysis. Cohen's κ statistic and percent agreement were used to calculate inter-rater reliabilities for the dichotomous judgments made concerning the inclusion/exclusion of studies according to their abstract characteristics.[25] A Cohen's κ statistic of 0.61 or higher was considered to demonstrate substantial coding agreement between raters.[26] Additionally, the reference lists of articles marked for full-review were hand-searched to identify relevant studies

omitted by the electronic search. Articles identified for full-review were categorized as either learner 'performance articles', 'perceptions articles', or 'related articles' (e.g., reviews, editorials, commentaries, descriptive articles) based on the abstracts. Records classified as 'performance articles' that also included perceptions data were included in the thematic review. To avoid bias in data collection and to guard against variability in data interpretation, two teams of paired researchers extracted and coded data from articles selected for full-review. Lastly, published studies were excluded if they had incomplete datasets (i.e., were lacking sufficient raw data to calculate an effect size) and if attempts to acquire the data from the corresponding author(s) were unsuccessful.

Statistical and Thematic Analyses

Data were collected using a customized form generated in REDCap [27] and were exported to Microsoft Excel® for organization and cleaning. Data were then input into Review Manager (RevMan 5.3) to calculate standardized mean differences (using Hedges' adjusted g), a summary effect size, heterogeneity, and to generate forest and funnel plots. The summary effect size was calculated according to a random-effects model. Inverse variance was used to weight studies as a function of their sample size. On occasion, multiple standardized mean differences were computed for a single study if the study design examined effects across multiple measures (e.g., exams). It should also be noted that in the presence of exam level data, standardized mean differences were not calculated for subcomponents of a single exam. The magnitude of the summary effect size, reported as a standardized mean difference (SMD), was interpreted using Cohen's recommendations for small (0.20-0.49), medium (0.50-0.79), and large (≥0.80) effects.[28] Confidence intervals (CI) were also reported. For additional information on meta-

analytic procedures and computations, we refer you to the following resources.[29-31]

The presence of heterogeneity was detected using a Q statistic (distributed as a Chi-square statistic).[32] To complement the Q statistic, the extent of between-study variance was estimated with an I^2 statistic. The variance in effect estimates beyond chance was interpreted as being of nominal importance if the inconsistency in study results (I^2) was less than 25% and considerable heterogeneity was considered to exist if I^2 was greater than 75%.[33] To detect whether an overestimation of population effects was likely, publication bias (i.e., a disproportionate number of studies that present positive versus negative or inconclusive findings) was evaluated by exploring funnel plot symmetry.[34]

In a qualitative review of articles, themes were identified regarding the advantages and disadvantages of VM. To understand whether VM was predominantly preferred over OM, perceptions data were extracted and analyzed across studies from various representative survey questions. Because some studies polled students' opinions on the use of VM for studying versus its use for test taking, a Cochran-Mantel-Haenszel statistic[35, 36] was computed to test whether the proportion of learners preferring VM was the same regardless of whether VM was used for learning/studying or testing.

RESULTS

The electronic literature search returned 860 relevant records. Once duplicates were removed, 725 studies remained. After a preliminary and secondary screening that excluded a number of records, 72 studies underwent full-text review (33 performance articles and 39 perceptions articles). Of the 33 performance articles, 21 were excluded from the meta-analysis because they failed to meet the eligibility criteria or had insufficient data for calculating effect sizes. Data extracted from 12 studies were included in the meta-analysis to assess the effects of

VM on learner knowledge gains compared to OM and 37 studies were included in the thematic review (Figure 1). By happenstance, all meta-analysis data were extracted from published full-text articles and did not include data from abstracts or unpublished works.

After the preliminary screening, percent agreement and Cohen's κ for inter-rater agreement were calculated for each paired research team. The first 2-member team screened 362 studies and had a percent agreement to include/exclude studies of 90.3% with a Cohen's κ of 0.628. A total of 363 studies were screened by the second 2-member team who had a percent agreement of 87.9% and a Cohen's κ of 0.523. As outlined in the methods section, a secondary cross-over screening was necessary to resolve discrepancies.

Study Demographics

Table 1 descriptively summarizes the differences between the 12 studies included in the meta-analysis. Most studies occurred within the disciplines of histology or pathology with either medical students or undergraduate college students. While the designs of the studies varied, studies with larger sample sizes tended to rely on historical controls for comparisons while smaller studies frequently utilized randomized fully-crossed designs.

Meta-analysis of VM Effectiveness

As shown by the forest plot in Figure 2, the summary effect size was calculated in the context of 1,978 subjects exposed to VM and 3,950 subjects exposed to OM. By combining 18 different outcomes from 12 studies across two learner types, VM demonstrated a small positive effect (SMD=0.28, [CI=0.09, 0.47], p=0.003) on learning according to Cohen's convention for the magnitude of effect sizes. A significant Q statistic (p<0.001) indicated the presence of

heterogeneity. According to the total I^2 index, 89% of the total variation in study estimates was due to heterogeneity rather than sampling error. By convention, this represents considerable variation between studies. When segregating studies by learner type (i.e., medical students versus undergraduate students), studies conducted with undergraduate students were found to be homogenous (I^2=0%; p=0.46; Figure 2). Because total heterogeneity was greater than 50%, a random-effects model was used to calculate the summary effect size.[30] Using the formula

$$P=1-\Phi(1.96-(\frac{ES}{\sqrt{v}}))$$

proposed by Valentine et al.,[37] a two-tailed power analysis (P) for random-effects meta-analyses was computed using the observed values for summary effect size (ES=0.282), summary variance (v=0.009), and the standard normal cumulative distribution function ($\Phi(x)$). The computation revealed our meta-analysis had sufficient power (P=0.84) to detect what we considered to be the smallest important effect of .20 given the number of studies and within-study sample sizes. A funnel plot for detecting potential publication bias showed reasonably symmetry indicating that bias in favor of positive findings was not likely (Figure 3).

Analysis of Learners' Perceptions

In reviewing a number of articles that investigated learners' perceptions of VM, several themes emerged that placed VM at an advantage over OM (Table 2). Numerous studies noted that as an educational resource VM requires minimal maintenance and is more cost effective compared to glass slides (e.g., one Australian study projected a savings of over $1 million (AUD)).[38] Additionally, the ability of VM to disseminate identical slides to multiple users simultaneously was a frequently cited advantage. The most prevalent disadvantage to emerge

from the review was that VM does not afford students the experience of learning how to use an optical microscope. Additional themes are reported in Table 2.

It was also noted that studies on learners' perceptions collectively reported a general preference, or favorable attitude, toward VM over OM. For example, among the studies listed in Table 3, preference in favor of VM was reported on average by 70% of respondents. However, in studies that specifically compared VM to OM for studying versus testing, students preferred VM for studying and OM for test taking (Table 4). This observation was statistically supported by a Cochran-Mantel-Haenszel (CMH) test. Individuals who preferred VM indicated VM was useful for studying in significantly higher proportions ($\chi^2_{CMH}=70.37$, 1 df, p<0.001) than individuals preferring OM. Conversely, individuals favoring OM preferred OM for test taking in significantly higher proportions ($\chi^2_{CMH}=78.77$, 1 df, p<0.001) than individuals who favored VM for test taking.

DISCUSSION

This analysis revealed that across studies comprised of different learner types and content domains, learners who were exposed to microscopic content through virtual microscopy overall demonstrated a small yet statistically significant improvement in performance scores. Though the majority of individual studies independently reported non-significant findings, "By combining studies, a meta-analysis increases sample size and thus the power to study the effects of interest".[39]

While the major study finding was unexpected and refuted the study's hypothesis, the reported benefits of VM and learner perceptions data helped to elucidate and provide context for understanding this phenomenon. In general, the small reported effect may be attributable to both the 'ease of access' and 'ease of use' that VM affords over optical microscopy.

Recent years have witnessed an explosive growth in the use of educational technology in all facets of medical education,[40] and VM is yet another example of this trend. The provision of anytime/anywhere interactive learning has great appeal to the current generation of millennial learners who are already immersed in the prevailing digital culture. Having unrestricted access to view digital slides from anywhere and at any time may have contributed to small improvements in student performance. It is also conceivable that the pedagogical strategies used in some studies led to performance gains that were as much attributable to engaged student learning as to the use of VM per se. As medical educators embrace new teaching methods that facilitate active learning and greater student engagement, they often incorporate interactive technologies like VM to serve these ends. For example, VM has been effectively used in the context of student-centered activities such as team-based learning,[12] case-based learning,[14] peer teaching,[41] and collaborative education.[42] The true value of VM conflated with the benefits of novel teaching strategies may explain improved learning outcomes irrespective of the independent effects of VM. Further investigation is needed to more fully understand the influence of these potential interaction effects on learning outcomes.

Another explanation for this finding is that VM may inadvertently 'level the playing field' between those who are adept at using microscopes and those who struggle with the mechanics of optical microscopy (e.g., adjusting illumination and contrast, maintaining orientation, etc.).[43-48] In their survey of medical students studying histopathology, Kumar and co-authors[43] found that 81.6% of the respondents believed VM solved the problems they had experienced when using the optical microscope. A frequent student complaint about optical microscopes is that they are difficult to use and cause eye strain.[43-45] Perhaps the principal advantage of VM is that it diminishes the frustrations some students experience in using the optical microscope and

effectively eliminates this skill barrier as an impediment to effective learning. In support of this notion, several surveys reported strong majorities of students (80.7% - 93.8%) who believed that VM saved them time compared to using the optical microscopes.[46-48] This suggests that VM offers certain efficiencies that may benefit a subset of students who might otherwise be disadvantaged using optical microscopes.

As rationale for conducting a VM study with undergraduate students, Hussman et al. maintained that the lack of previously observed effects between traditional and virtual teaching methods might have been a result of the study subject (i.e., medical students) as opposed to the mode of instruction.[8] The reality of the nature of medical students is they "will take it upon themselves to learn the material irrespective of teaching modalities".[8] However, our findings suggest that the type of study subject (medical student vs. undergraduate student) did not moderate performance gains, as no heterogeneity was detected between subgroups (I^2=0%, p=0.82, Figure 2). That is, no genuine variation in performance across types of learners was found to exist.

While some studies reported students advocating for the elimination of OM,[6, 44-46] others indicated a preference for using VM and OM interchangeably. In general, authors expressed hesitation to completely eliminate OM citing A) students' need to develop proficient microscope skills for future coursework and clinical practice,[6, 43, 49-55] B) a preference for realistic slides rather than 'textbook' quality virtual slides,[45, 52, 54] and C) inadequate fine focus and illumination/contrast capabilities with VM systems.[43, 45, 48, 51, 54] Not only are some educators critics of this later point, but also clinical pathologists and regulatory agencies. In the United States, the Food and Drug Administration (FDA) regulates the manufacturing of digital pathology technologies, which are classified as medical devices. The FDA has recently drafted a

guidance document that outlines regulatory recommendations for the use of digital pathology whole slide imaging systems.[56] While several manufacturers have received clearance from the FDA to use their products for specific diagnostic and research purposes, VM has yet to be approved by the FDA as safe and effective for primary diagnosis.[57] As VM technologies continue to advance and become mainstays in clinical practice and education, it is only a matter of time before formal approval is granted by the FDA which is likely to foster broader acceptance of VM among medical educators. Nevertheless, OM has long proven itself to be an effective tool in clinical practice and education, and its wholesale elimination from the medical curriculum might be ill-advised. There are ample reasons to maintain at least a niche for OM in the curriculum, and some have argued for a hybrid approach using both OM and VM to retain the unique benefits of each.[58]

Limitations

The summary effect size was statistically significant, however, the amount of heterogeneity among medical student studies was considerable (I^2=93%, p<0.001). The limited number of medical student studies (k=7) and the absence of suitable information across studies precluded us from conducting post-hoc moderator analyses to discern the key factors contributing to the observed heterogeneity. As such, we suspect that the presence of heterogeneity may have been attributed to the variability in the sophistication and quality of VM platforms, the robustness of an institution's educational technology infrastructure, and/or the extent to which VM was accessed by users. In some studies, for example, students often reported technical difficulties and delays in downloading virtual slides.[11, 43, 59] Additionally, in one medical student study, access to VM slides was only available during laboratory sessions,[54]

whereas the other studies featured more advanced VM platforms with on-demand access via Internet connections. Variability in the quality of performance assessments (e.g., using the same versus different VM images for testing purposes) and study design (e.g., randomized vs. non-randomized) could have also led to increased heterogeneity, as these aspects relate directly to study validity. It is also reasonable to contend that confounding differences in the duration of the interventions, the timing of the administered assessments, and inherent differences between American and non-American medical education systems may have partly contributed to significant levels of heterogeneity. Additionally, results are limited because few studies included long-term follow-ups as part of their design to assess knowledge retention.

Although it is common for meta-analyses to report multiple effect estimates from the same study population, from a strict statistical standpoint this is a violation of the independence assumption that underlies the procedures for aggregating data.[29] This assumption was violated in this study to expand the number of effect estimates included in the analysis. In instances where studies reported outcomes on multiple exams yet failed to report global findings (e.g., course level outcomes), we chose to estimate the effects for each performance outcome rather than calculate a composite score under the presumption of equally weighted performance measures which could have grossly misrepresented the reported data. While the consequence of violating the independence assumption is potential bias in the summary effect estimate,[29] we chose to err on the side of overinclusion.

CONCLUSIONS

Despite many individual studies reporting non-significant effects, the enhanced power afforded by meta-analysis revealed that collectively learners who were exposed to the

pedagogical approach of VM performed at a slightly higher level than students who utilized traditional OM; as evidenced by the small yet statistically significant summary effect size. This finding in combination with the extracted themes and aggregated respondent preference data suggests the benefits of VM as an educational resource and pedagogical approach are appreciable. While editorials devoted to discussing the place and time to teach and optimally employ OM may still hold merit, we are hopeful the nature of this research and its ability to summarize the VM literature will encourage academicians to engage in concomitant research topics that move beyond the perpetual discussions of educational technology to focus on other contemporary medical education themes and issues.
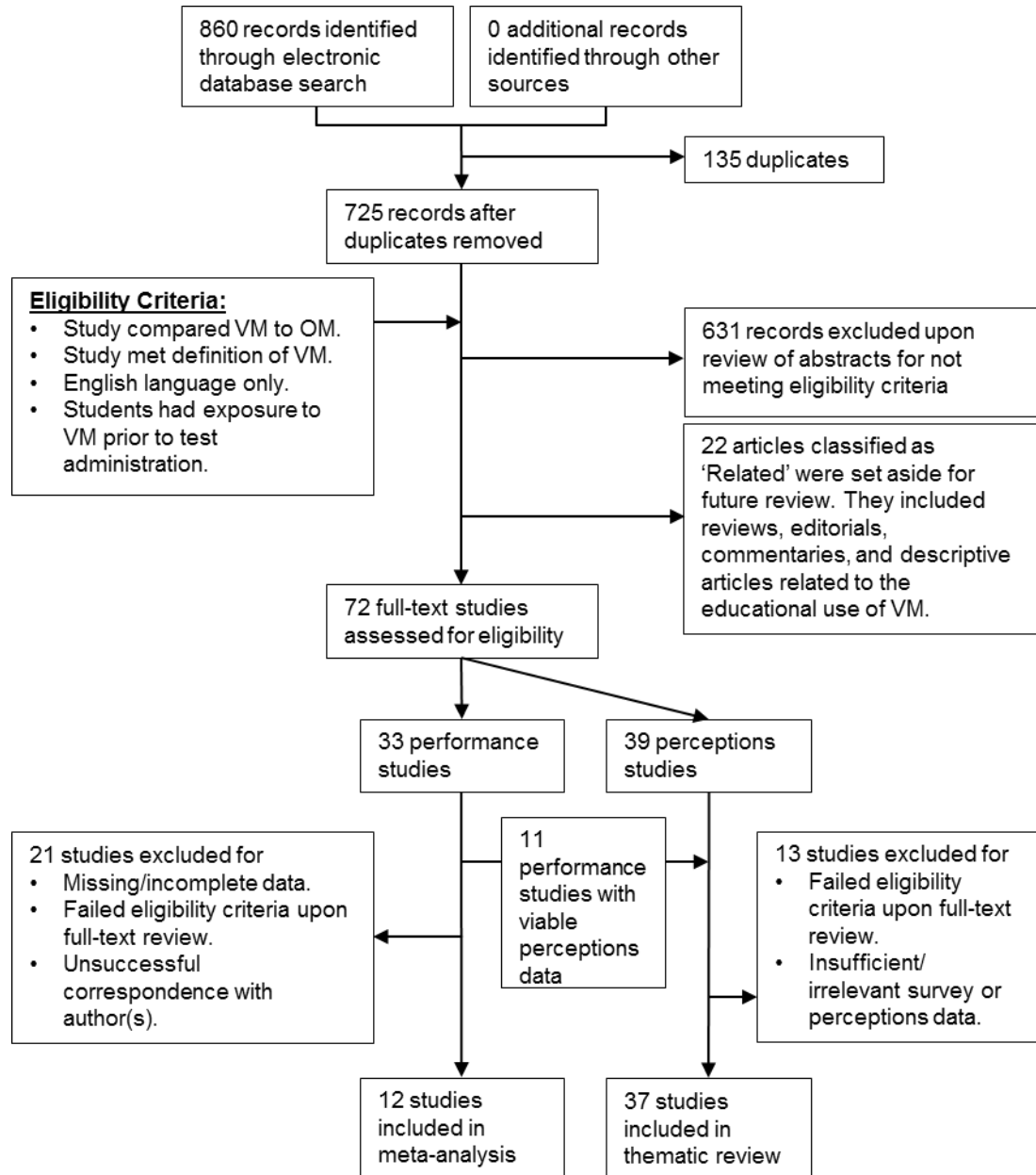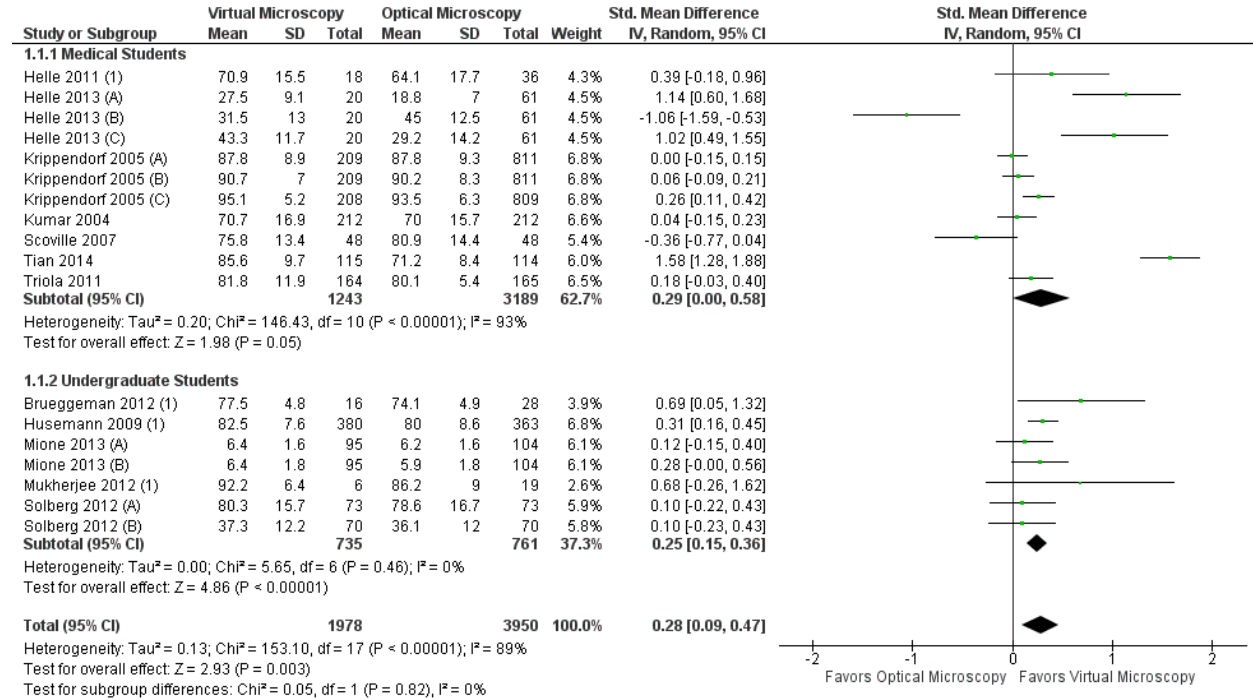
ACKNOWLEDGEMENTS

Figure 1: Flow chart of study information.

Table 1: Overview of studies included in meta-analysis

| Study	Country of Origin	Discipline	Learner Population	Study Design	Duration of Intervention	Comparability of Assessment Measures
Multi-Cohort Studies						
Krippendorf 2005 [52]	USA	Histology	Medical Students	Historical control	1 semester	Similar assessments
Scoville 2007 [11]	USA	Histology	Medical Students	Randomized cohorts	1 histology unit	Identical assessments
Husmann 2009 [8]	USA	Histology	Undergraduate Students	Historical control	1 semester	Similar assessments
Helle 2011 [60]	Finland	Pathology	Medical Students	Randomization not specified	1 week	Similar assessments
Triola 2011 [42]	USA	Histology	Medical Students	Historical control	1 semester	Identical assessments
Brueggeman 2012 [61]	USA	Hematology	Undergraduate Students	Randomized cohorts	1 semester	Identical assessments
Mukherjee 2012 [59]	USA	Cytotechnology	*Post-Baccalaureate (certificate program)	Historical control	Not specified	Similar assessments
Helle 2013 [16]	Finland	Pathology	Medical Students	Historical control	9 weeks	Not specified
Tian 2014 [62]	China	Histology	Medical Students	Non-random cohort assignment	1 semester	Identical assessments
Single Cohort Studies						
Kumar 2004 [43]	Australia	Pathology	Medical Students	VM vs OM cases compared	1 semester	Identical assessments
Solberg 2012 [54]	USA	Cytology	Undergraduate Students	Randomized cross-over design	1 cytology unit	Identical assessments
Mione 2013 [63]	Belgium	Histology	Undergraduate Students	Alphabetically randomized cross-over design	1 semester	Different pre/posttest, identical exams between comparative groups

*Classified with "undergraduate students" for analysis.

Figure 2: Random-effects model forest plot with summary and study specific effects of VM

Study or Subgroup	Virtual Microscopy Mean	SD	Total	Optical Microscopy Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Medical Students								
Helle 2011 (1)	70.9	15.5	18	64.1	17.7	36	4.3%	0.39 [-0.18, 0.96]
Helle 2013 (A)	27.5	9.1	20	18.8	7	61	4.5%	1.14 [0.60, 1.68]
Helle 2013 (B)	31.5	13	20	45	12.5	61	4.5%	-1.06 [-1.59, -0.53]
Helle 2013 (C)	43.3	11.7	20	29.2	14.2	61	4.5%	1.02 [0.49, 1.55]
Krippendorf 2005 (A)	87.8	8.9	209	87.8	9.3	811	6.8%	0.00 [-0.15, 0.15]
Krippendorf 2005 (B)	90.7	7	209	90.2	8.3	811	6.8%	0.06 [-0.09, 0.21]
Krippendorf 2005 (C)	95.1	5.2	208	93.5	6.3	809	6.8%	0.26 [0.11, 0.42]
Kumar 2004	70.7	16.9	212	70	15.7	212	6.6%	0.04 [-0.15, 0.23]
Scoville 2007	75.8	13.4	48	80.9	14.4	48	5.4%	-0.36 [-0.77, 0.04]
Tian 2014	85.6	9.7	115	71.2	8.4	114	6.0%	1.58 [1.28, 1.88]
Triola 2011	81.8	11.9	164	80.1	5.4	165	6.5%	0.18 [-0.03, 0.40]
Subtotal (95% CI)			1243			3189	62.7%	0.29 [0.00, 0.58]

Heterogeneity: Tau² = 0.20; Chi² = 146.43, df = 10 (P < 0.00001); I² = 93%
Test for overall effect: Z = 1.98 (P = 0.05)

1.1.2 Undergraduate Students								
Brueggeman 2012 (1)	77.5	4.8	16	74.1	4.9	28	3.9%	0.69 [0.05, 1.32]
Husemann 2009 (1)	82.5	7.6	380	80	8.6	363	6.8%	0.31 [0.16, 0.45]
Mione 2013 (A)	6.4	1.6	95	6.2	1.6	104	6.1%	0.12 [-0.15, 0.40]
Mione 2013 (B)	6.4	1.8	95	5.9	1.8	104	6.1%	0.28 [-0.00, 0.56]
Mukherjee 2012 (1)	92.2	6.4	6	86.2	9	19	2.6%	0.68 [-0.26, 1.62]
Solberg 2012 (A)	80.3	15.7	73	78.6	16.7	73	5.9%	0.10 [-0.22, 0.43]
Solberg 2012 (B)	37.3	12.2	70	36.1	12	70	5.8%	0.10 [-0.23, 0.43]
Subtotal (95% CI)			735			761	37.3%	0.25 [0.15, 0.36]

Heterogeneity: Tau² = 0.00; Chi² = 5.65, df = 6 (P = 0.46); I² = 0%
Test for overall effect: Z = 4.86 (P < 0.00001)

| **Total (95% CI)** | | | 1978 | | | 3950 | 100.0% | 0.28 [0.09, 0.47] |

Heterogeneity: Tau² = 0.13; Chi² = 153.10, df = 17 (P < 0.00001); I² = 89%
Test for overall effect: Z = 2.93 (P = 0.003)
Test for subgroup differences: Chi² = 0.05, df = 1 (P = 0.82), I² = 0%

Favors Optical Microscopy — Favors Virtual Microscopy

(1) Data provided by corresponding author
(A), (B), (C) Multiple reported findings from same student population

Figure 3: Funnel plot for assessing publication bias.
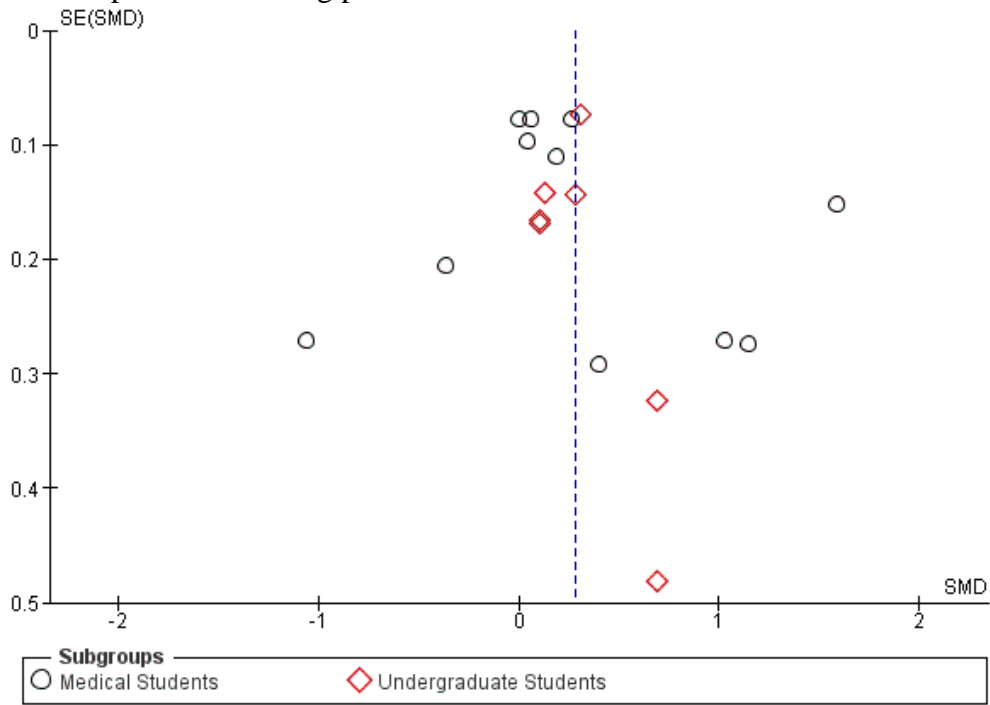
SMD: standardized mean difference; SE: standard error

Table 2: Themes regarding the advantages and disadvantages of VM

Themes	References
Advantages	
VM is an enduring educational resource that requires less maintenance (e.g., no microscope repairs, slides don't deteriorate, etc.) and is more cost effective than OM on a per student basis; (an added advantage for programs with expanding class sizes).	Bonser 2013;[64] Neel 2007;[7] Scoville 2007;[11] Braun 2008;[15] Sivamalai 2011;[38] Bowa 2014;[65] Kogan 2014;[51] Kumar 2004;[43] Krippendorf 2005;[52] Boutonnat 2006;[66] Glatz-Krieger 2006;[67] Kumar 2006;[68] Mills 2007;[53] Farah 2009;[6] Weaker 2009;[47] Maybury 2010;[44] Camparo 2012[5]
Slides selected for their educational quality can be viewed by multiple students simultaneously (including slides that were once too scarce or valuable to be used).	Bonser 2013;[64] Pinder 2008;[10] Husmann 2009;[8] Fontelo 2012;[69] Solberg 2012;[54] Kogan 2014;[51] Blake 2003;[46] Dee 2003;[70] Kumar 2004;[43] Krippendorf 2005;[52] Boutonnat 2006;[66] Glatz-Krieger 2006;[67] Mills 2007;[53] Farah 2009;[6] Weaker 2009;[47] Fonyad 2010[50]
Virtual slides can be accessed outside of the classroom (anytime/anywhere).	Tian 2014;[62] Pinder 2008;[10] Sivamalai 2011;[38] McCready 2013;[71] Kogan 2014;[51] Dee 2003;[70] Kumar 2004;[43] Mills 2007;[53] Farah 2009;[6] Maybury 2010[44]
Virtual slides can be annotated and used for large group demonstrations.	Tian 2014;[62] Pinder 2008;[10] Husmann 2009;[8] Glatz-Krieger 2006;[67] Fonyad 2010;[50] Maybury 2010[44]
Virtual slides can be viewed side by side for direct comparison or to facilitate slide orientation.	Husmann 2009;[8] Dee 2003;[70] Glatz-Krieger 2006;[67] Kumar 2006;[68] Fonyad 2010;[50] Maybury 2010[44]
VM purportedly reduces eye strain/fatigue.	Braun 2008;[15] Solberg 2012;[54] Becker 2006;[45] Mills 2007;[53] Farah 2009;[6] Maybury 2010[44]
VM eradicates the difficulty some students experience operating light microscopes.	Krippendorf 2005;[52] Becker 2006;[45] Boutonnat 2006;[66] Kumar 2006;[68] Farah 2009;[6] Maybury 2010[44]
Disadvantages	
Students may no longer be trained in how to use light microscopes.	Tian 2014;[62] Harris 2001;[13] Neel 2007;[7] Scoville 2007;[11] Husmann 2009;[8] Koch 2009;[19] Solberg 2012;[54] Kogan 2014;[51] Kumar 2004;[43] Krippendorf 2005;[52] Mills 2007;[53] Farah 2009;[6] Fonyad 2010;[50] Marchevsky 2003[55]
Technology related issues (e.g., servers, image quality/resolution, Wi-Fi connectivity, loss of focal planes, etc.).	Scoville 2007;[11] Braun 2008;[15] Koch 2009;[19] Sivamalai 2011;[38] Szymas 2011;[72] Fontelo 2012;[69] Solberg 2012;[54] McCready 2013;[71] Becker 2006;[45] Glatz-Krieger 2006;[67] Mills 2007;[53] Farah 2009;[6] Fonyad 2010;[50] Maybury 2010[44]
Startup costs associated with slide digitization and establishing a VM infrastructure.	Pinder 2008;[10] Kogan 2014;[51] Fonyad 2010;[50] Maybury 2010[44]
Loss of appreciation for normal variation between glass slides.	Solberg 2012;[54] Krippendorf 2005[52]

Table 3: Comparison of learners favoring VM over OM (studying versus testing not specified)

	N Favoring VM (%)	N Favoring OM (%)	Total N
Farah 2009[48]	30 (53)	17 (30)	57†
Farah 2009[6]	Oral biology: 41 (77)	9 (17)	53†
	Oral pathology: 33 (97)	4 (12)	34*
Anyanwu 2012[49]	180 (68)	67 (25)	265
Mukherjee 2012[59]	0 (0)	6 (100)	6
Bonser 2013[64]	74 (93)	6 (7)	80
Merk 2010[73]	167 (87)	25 (13)	192
Sivamalai 2011[38]	51 (96)	2 (4)	53
Rosas 2012[74]	31 (15)	152 (75)	204†
Solberg 2012[54]	46 (62)	25 (34)	74
McCready 2013[71]	103 (92)	2 (2)	112
Kogan 2014[51]	44 (34)	52 (40)	130
Braun 2008[15]	30 (43)	5 (7)	69†
Mills 2007[53]	62 (69)	57 (63)	90*
Brick 2014[75]	15 (83)	11 (61)	18*
Krippendorf 2005[52]	MS1: 199 (97)	-	206
	MS2: 107 (73)	-	147
Maybury 2010[44]	33 (43)	-	76
Pinder 2008[10]	MS1: 158 (91)	-	173
	MS2: 100 (99)	-	101
Becker 2006[45]	Path: 86 (85)	-	101
	Medical: 78 (80)	-	98
Average %	70%	33%	

†Neutral and missing responses explains discrepancies between the sum of VM and OM and total n
*Some respondents favored the use of both VM and OM
- No direct comparison to OM
MS1: first-year medical students; MS2: second-year medical students

Table 4: Comparison of learners preferring VM for studying versus test taking

		N preferring VM (%)	N preferring OM (%)	Total N
Scoville 2007[11]	Studying	11 (35)	13 (41)	32†
	Test Taking	3 (9)	15 (46)	33†
Neel 2007[7]	Studying	45 (72.6)	17 (27.4)	62
	Test Taking	10 (16.1)	44 (71)	62†
Koch 2009[19]	Studying	106 (80)	*26 (20)	132
	Test Taking	58 (44)	*74 (56)	132

*Missing data were computed based on VM data and total n.
†Neutral responses explain discrepancies between the sum of VM and OM and total n. The Cochran-Mantel-Haenszel test did not take into account neutral responses.

REFERENCES

[1]   Dee FR, Meyerholz DK. Teaching medical pathology in the twenty-first century: virtual microscopy applications. *J Vet Med Educ*. 2007;**34**:431-436.

[2]   Hedvat CV. Digital microscopy: past, present, and future. *Arch Pathol Lab Med*. 2010;**134**:1666-1670.

[3]   Bloodgood RA, Ogilvie RW. Trends in histology laboratory teaching in United States medical schools. *The Anatomical Record Part B: The New Anatomist*. 2006;**289**:169-175.

[4]   Hutchinson C, Brereton M, Burthem J. Digital imaging of haematological morphology. *Clin Lab Haematol*. 2005;**27**:357-362.

[5]   Camparo P, Egevad L, Algaba F, et al. Utility of whole slide imaging and virtual microscopy in prostate pathology. *APMIS*. 2012;**120**:298-304.

[6]   Farah CS, Maybury TS. The e-evolution of microscopy in dental education. *J Dent Educ*. 2009;**73**:942-949.

[7]   Neel JA, Grindem CB, Bristol DG. Introduction and evaluation of virtual microscopy in teaching veterinary cytopathology. *J Vet Med Educ*. 2007;**34**:437-444.

[8]   Husmann PR, O'Loughlin VD, Braun MW. Quantitative and qualitative changes in teaching histology by means of virtual microscopy in an introductory course in human anatomy. *Anat Sci Educ*. 2009;**2**:218-226.

[9]   Higazi TB. Use of interactive live digital imaging to enhance histology learning in introductory level anatomy and physiology classes. *Anat Sci Educ*. 2011;**4**:78-83.

[10]   Pinder KE, Ford JC, Ovalle WK. A new paradigm for teaching histology laboratories in Canada's first distributed medical school. *Anat Sci Educ*. 2008;**1**:95-101.

[11]   Scoville SA, Buskirk TD. Traditional and virtual microscopy compared experimentally in a classroom setting. *Clin Anat*. 2007;**20**:565-570.

[12]   Goldberg HR, Dintzis R. The positive impact of team-based virtual microscopy on student learning in physiology and histology. *Advances in physiology education*. 2007;**31**:261-265.

[13]   Harris T, Leaven T, Heidger P, Kreiter C, Duncan J, Dick F. Comparison of a virtual microscope laboratory to a regular microscope laboratory for teaching histology. *The anatomical record*. 2001;**265**:10-14.

[14]   McBride JM, Prayson RA. Development of a synergistic case‐based microanatomy curriculum. *Anat Sci Educ*. 2008;**1**:102-105.

[15]   Braun MW, Kearns KD. Improved learning efficiency and increased student collaboration through use of virtual microscopy in the teaching of human pathology. *Anat Sci Educ*. 2008;**1**:240-246.

[16]   Helle L, Nivala M, Kronqvist P. More technology, better learning resources, better learning? Lessons from adopting virtual microscopy in undergraduate medical education. *Anat Sci Educ*. 2013;**6**:73-80.

[17]   Carlson AM, McPhail ED, Rodriguez V, Schroeder G, Wolanskyj AP. A prospective, randomized crossover study comparing direct inspection by light microscopy versus projected images for teaching of hematopathology to medical students. *Anat Sci Educ*. 2014;**7**:130-134.

[18]   Bruch LA, De Young BR, Kreiter CD, Haugen TH, Leaven TC, Dee FR. Competency assessment of residents in surgical pathology using virtual microscopy. *Hum Pathol*. 2009;**40**:1122-1128.

[19]   Koch LH, Lampros JN, Delong LK, Chen SC, Woosley JT, Hood AF. Randomized comparison of virtual microscopy and traditional glass microscopy in diagnostic accuracy among dermatology and pathology residents. *Hum Pathol*. 2009;**40**:662-667.

[20]   Marsch AF, Espiritu B, Groth J, Hutchens KA. The effectiveness of annotated (vs. non‐annotated) digital pathology slides as a teaching tool during dermatology and pathology residencies. *J Cutan Pathol*. 2014;**41**:513-518.

[21]   Guzman M, Judkins AR. Digital pathology: a tool for 21st century neuropathology. *Brain Pathol*. 2009;**19**:305-316.

[22]   Rojo MG, Punys V, Slodkowska J, Schrader T, Daniel C, Blobel B. Digital pathology in Europe: coordinating patient care and research efforts. *Stud Health Technol Inform*. 2009;**150**:997-1001.

[23]   Greco T, Zangrillo A, Biondi-Zoccai G, Landoni G. Meta-analysis: pitfalls and hints. *Heart, lung and vessels*. 2013;**5**:219.

[24]   Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;**151**:264-269.

[25]   Cohen J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*. 1960;**20**:37-46.

[26]   Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977:159-174.

[27]   Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;**42**:377-381.

[28]   Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale NJ Erlbaum; 1988.

[29]   Petitti D. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. 2nd ed. New York Oxford University Press; 2000.

[30]   Field AP, Gillett R. How to do a meta‐analysis. *Br J Math Stat Psychol*. 2010;**63**:665-694.

[31]   Slavin RE. Meta-analysis in education: How has it been used? *Educational researcher*. 1984:6-15.

[32]   Huedo-Medina TB, Sánchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychol Methods*. 2006;**11**:193.

[33]   Higgins J, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;**327**:557-560.

[34]   Duval S, Tweedie R. Trim and fill: a simple funnel‐plot‐based method of testing and adjusting for publication bias in meta‐analysis. *Biometrics*. 2000;**56**:455-463.

[35]   Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies. *J Natl Cancer Inst*. 1959;**22**:719-748.

[36]   Yanagawa T, Fujii Y, Mastuoka J. Generalized Mantel-Haenszel procedures for 2× J tables. *Environ Health Perspect*. 1994;**102**:57.

[37]   Valentine JC, Pigott TD, Rothstein HR. How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*. 2010;**35**:215-247.

[38]   Sivamalai S, Murthy SV, Gupta TS, Woolley T. Teaching pathology via online digital microscopy: positive learning outcomes for rurally based medical students. *Aust J Rural Health*. 2011;**19**:45-51.

[39]   Walker E, Hernandez AV, Kattan MW. Meta-analysis: its strengths and limitations. *Cleve Clin J Med*. 2008;**75**:431-439.

[40]   Han H, Resch DS, Kovach RA. Educational technology in medical education. *Teach Learn Med*. 2013;**25**:S39-S43.

[41]   Bloodgood RA. Active learning: A small group histology laboratory exercise in a whole class setting utilizing virtual slides and peer education. *Anat Sci Educ*. 2012;**5**:367-373.

[42]   Triola MM, Holloway WJ. Enhanced virtual microscopy for collaborative education. *BMC Med Educ*. 2011;**11**:4.

[43]   Kumar RK, Velan GM, Korell SO, Kandara M, Dee FR, Wakefield D. Virtual microscopy for learning and assessment in pathology. *The Journal of pathology*. 2004;**204**:613-618.

[44]   Maybury TS, Farah CS. Electronic Blending in Virtual Microscopy. *Journal of Learning Design*. 2010;**4**:41-51.

[45]   Becker JH. Virtual microscopes in podiatric medical education. *J Am Podiatr Med Assoc*. 2006;**96**:518-524.

[46]   Blake CA, Lavoie HA, Millette CF. Teaching medical histology at the University of South Carolina School of Medicine: transition to virtual slides and virtual microscopes. *The Anatomical Record Part B: The New Anatomist*. 2003;**275**:196-206.

[47]   Weaker FJ, Herbert DC. Transition of a dental histology course from light to virtual microscopy. *J Dent Educ*. 2009;**73**:1213-1221.

[48]   Farah C, Maybury T. Implementing digital technology to enhance student learning of pathology. *Eur J Dent Educ*. 2009;**13**:172-178.

[49]   Anyanwu G, Agu A, Anyaechie U. Enhancing learning objectives by use of simple virtual microscopic slides in cellular physiology and histology: Impact and attitudes. *Advances in Physiology Education*. 2012;**36**:158-163.

[50]   Fónyad L, Gerely L, Cserneky M, Molnár B, Matolcsy A. Shifting gears higher-digital slides in graduate education-4 years experience at Semmelweis University. *Diagn Pathol*. 2010;**5**:73.

[51]   Kogan LR, Dowers KL, Cerda JR, Schoenfeld-Tacher RM, Stewart SM. Virtual Microscopy: A Useful Tool for Meeting Evolving Challenges in the Veterinary Medical Curriculum. *Journal of Science Education and Technology*. 2014;**23**:756-762.

[52]   Krippendorf BB, Lough J. Complete and rapid switch from light microscopy to virtual microscopy for teaching medical histology. *The Anatomical Record Part B: The New Anatomist*. 2005;**285**:19-25.

[53]   Mills PC, Bradley AP, Woodall PF, Wildermoth M. Teaching histology to first-year veterinary science students using virtual microscopy and traditional microscopy: a comparison of student responses. *J Vet Med Educ*. 2007;**34**:177-182.

[54]   Solberg BL. Digital and traditional slides for teaching cellular morphology: a comparative analysis of learning outcomes. *Clinical laboratory science: Journal of the American Society for Medical Technology*. 2012;**25**.

[55]   Marchevsky AM, Wan Y, Premi T, Krishnan L. Virtual microscopy as a tool for proficiency testing in cytopathology: a model using multiple digital images of Papanicolaou tests. *Arch Pathol Lab Med*. 2003;**127**:1320.

[56]   Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices: Draft Guidance for Industry and Food and Drug Administration Staff. In: Administration USFaD, ed., 2015.

[57]   Digital Pathology Association. 2015.

[58]   Pratt RL. Are we throwing histology out with the microscope? A look at histology from the physician's perspective. *Anat Sci Educ*. 2009;**2**:205-209.

[59]   Mukherjee MS. *An investigation of the incorporation of virtual microscopy in the cytotechnology educational curriculum* University of Nebraska Medical Center; 2012.

[60]   Helle L, Nivala M, Kronqvist P, Gegenfurtner A, Björk P, Säljö R. Traditional microscopy instruction versus process-oriented virtual microscopy instruction: a naturalistic experiment with control group. *Diagn Pathol*. 2011;**6**:S8.

[61]   Brueggeman MS, Swinehart C, Jane Yue M, Conway-Klaassen JM, Wiesner SM. Implementing virtual microscopy improves outcomes in a hematology morphology course. *Clin Lab Sci*. 2012;**25**:149.

[62]   Tian Y, Xiao W, Li C, et al. Virtual microscopy system at Chinese medical university: an assisted teaching platform for promoting active learning and problem-solving skills. *BMC Med Educ*. 2014;**14**:74.

[63]   Mione S, Valcke M, Cornelissen M. Evaluation of virtual microscopy in medical histology teaching. *Anat Sci Educ*. 2013;**6**:307-315.

[64]   Bonser SP, de Permentier P, Green J, Velan GM, Adam P, Kumar RK. Engaging students by emphasising botanical concepts over techniques: innovative practical exercises using virtual microscopy. *Journal of Biological Education*. 2013;**47**:123-127.

[65]   Bowa K, Anderson PG. Using Virtual Microscopy at Copperbelt University, Zambia. *Acad Med*. 2014;**89**:S107-S108.

[66]   Boutonnat J, Paulin C, Faure C, Colle P, Ronot X, Seigneurin D. A pilot study in two french medical schools for teaching histology using virtual microscopy 1. *Morphologie*. 2006;**90**:21-25.

[67]   Glatz-Krieger K, Spornitz U, Spatz A, Mihatsch MJ, Glatz D. Factors to keep in mind when introducing virtual microscopy. *Virchows Arch*. 2006;**448**:248-255.

[68]   Kumar RK, Freeman B, Velan GM, De Permentier PJ. Integrating histology and histopathology teaching in practical classes using virtual slides. *The Anatomical Record Part B: The New Anatomist*. 2006;**289**:128-133.

[69]   Fontelo P, Faustorilla J, Gavino A, Marcelo A. Digital pathology–implementation challenges in low-resource countries. *Anal Cell Pathol*. 2012;**35**:31-36.

[70]   Dee FR, Lehman JM, Consoer D, Leaven T, Cohen MB. Implementation of virtual microscope slides in the annual pathobiology of cancer workshop laboratory. *Hum Pathol*. 2003;**34**:430-436.

[71]   McCready ZR, Jham BC. Dental Students' Perceptions of the Use of Digital Microscopy as Part of an Oral Pathology Curriculum. *J Dent Educ*. 2013;**77**:1624-1628.

[72]   Szymas J, Lundin M. Five years of experience teaching pathology to dental students using the WebMicroscope. *Diagn Pathol*. 2011;**6**:S13.

[73]   Merk M, Knuechel R, Perez-Bouza A. Web-based virtual microscopy at the RWTH Aachen University: Didactic concept, methods and analysis of acceptance by the students. *Annals of Anatomy-Anatomischer Anzeiger*. 2010;**192**:383-387.

[74]   Rosas C, Rubí R, Donoso M, Uribe S. Dental students' evaluations of an interactive histology software. *J Dent Educ*. 2012;**76**:1491-1496.

[75]   Brick KE, Comfere NI, Broeren MD, Gibson LE, Wieland CN. The application of virtual microscopy in a dermatopathology educational setting: assessment of attitudes among dermatopathologists. *Int J Dermatol*. 2014;**53**:224-227.