

» Computer in der Medizinischen Ausbildung – Kontrollierte Evaluation eines computerbasierten Atlas der Histopathologie

Zusammenfassung: Für die Weiterentwicklung von Bildungsstandards ist es notwendig, die Qualität neuer Lehr- und Lernmittel zu prüfen. Im Rahmen einer kontrollierten Studie mit 72 Studenten des ersten Semesters des klinischen Studienabschnitts wurde MICROPAT, ein interaktiver Atlas der Histopathologie, mit Hilfe eines Crossover-Versuchsplans auf seinen Nutzen im Vergleich zu herkömmlichen Methoden des Wissenserwerbs mit Büchern hin überprüft. Das Modell konzentriert sich auf die Erhebung objektiver Daten, zusätzliche subjektive Benutzerangaben aus Fragebogen wurden ebenfalls berücksichtigt und in die Auswertung integriert. Die Erfolgskontrolle im Rahmen einer authentischen Prüfung am Mikroskop ergab, dass MICROPAT im Vergleich zu den konventionellen Medien ein gleichwertiges Medium ist. Eine ausschließliche Betrachtung richtiger und falscher Ergebnisse konnte nur einen nominalen, jedoch nicht signifikanten Vorteil für MICROPAT zeigen. Bezieht man in die Analyse einen zusätzlich vom Benutzer erfragten Sicherheit-Score ein, findet man eine signifikante Verbesserung der Lernleistung für das Computer-Lernen. Hinweise auf ein methodenspezifisch besseres Abschneiden von computererfahrenen Probanden fanden sich nicht. Die Teilnehmer der Evaluation beurteilten das Programm MICROPAT als gutes und geeignetes Lernmittel für die Histopathologie. Das Evaluationsmodell hat sich als praktikabel erwiesen und könnte Modellcharakter für weitere Untersuchungen im Rahmen der Qualitätskontrolle in der medizinischen Ausbildung haben.

Computers for Medical Schools – Controlled Evaluation of a Computer Based Atlas of Histopathology: To further the development of education standards, it is necessary to evaluate the quality of new teaching and learning tools. In order to assess the benefit of using an interactive atlas of histopathology we performed a randomized controlled trial with 72 3rd year medical students using a cross-over approach to compare computer-based learning with book-learning. A subjective evaluation was derived from questionnaires; the objective evaluation consisted of comparing scores on authentic examinations. Suitable correlations between both methods were demonstrated. The study outcome was assessed on diagnostic tasks in a learning environment using histological slides. Our results show that the pro-

T. C. Auhuber¹, H.-E. Schaefer¹, S. Schulz², R. Klar²

¹ Pathologisches Institut des Universitätsklinikums Freiburg (Direktor: Prof. Dr. med. H.-E. Schaefer)

² Institut für Medizinische Biometrie und Medizinische Informatik, Abteilung Medizinische Informatik des Universitätsklinikums Freiburg (Direktor: Prof. Dr. rer. nat. R. Klar)

gramme is equivalent to conventional means of acquiring knowledge, i.e. from textbooks and printed atlases. The computer group showed significantly better achievement only in the test results correlated with a self-rated diagnostic certainty-score. Students with previous computer experience generally performed better in both computer-based and book-guided learning, whereas the relative outcome of book-learning versus computer-learning revealed no correlation with the students' attitude towards computers. The students found MICROPAT a good and useful programme for learning histopathology. We suggest this evaluation scenario be used as a model for further standardized quality control studies in medical education.

Key words: Computer based training (CBT) – Evaluation – Cross-over trial – Histopathology

Einleitung

Das Potential von CBT (Computer Based Training) liegt in der Integration von verschiedensten Wegen der Wissensvermittlung mit multimedialen Techniken, in der interaktiven und nicht linearen Wissenspräsentation und vor allem in der Vermittlung von ständig aktualisierbarem Wissen über Netzwerke und kostengünstige Datenspeicher [30].

Das Ziel für die Entwicklung jedes CBT-Programms sollte neben allen ökonomischen und technologischen Überlegungen vor allem die Verbesserung, die Vereinfachung und die Beschleunigung der Wissensvermittlung mit einem solchen System sein [9]. Jede neue Methode sollte auch im Vergleich zu den bisher verwendeten traditionellen Medien beurteilt werden, um den Wert eines Programms einstuft zu können [3]. Als Maß dafür werden aus der Vielzahl von Evaluationsmöglichkeiten [37] häufig Benutzerfragebogen verwendet, um die Zufriedenheit der Benutzer, die Bedienbarkeit und den Unterhaltungswert des Programms zu ermitteln. Valide Rückschlüsse auf den postulierten Lernerfolg können so jedoch nicht gezogen werden. Bis jetzt gibt es nur wenige Studien, die den Nutzen computergestützter Lernmethoden im Vergleich zu herkömmlichen Lernmitteln bewerten und die Effektivität in der Wissensvermittlung sowie den Lernerfolg mit statistisch validen und reliablen Methoden beschreiben können. Standardisierungen, wie sie aus kontrollierten klinischen Studien be-

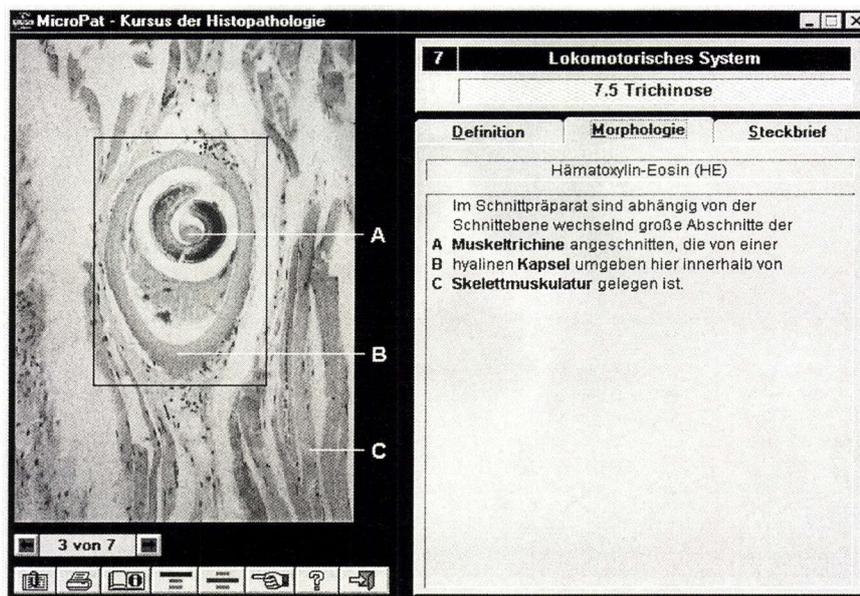


Abb. 1 Beispiel einer Atlasseite in MICRO-PAT aus dem Krankheitsbild „Trichinose“. Das mit Markierungen versehene Präparat auf der linken Seite des Bildschirms wird durch den Text auf der rechten Seite des Bildschirms näher erklärt. Unterhalb des Präparats kann man Navigationselemente erkennen, oberhalb des Textfeldes wird der Name des Krankheitsbildes und das zugehörige Organsystem angegeben.

kannt sind, fehlen aufgrund der Komplexität der Evaluation von Lehre und Lernen [3] vollständig.

Aus diesen Überlegungen heraus wurde an der Universität Freiburg mit Hilfe eines eigens dafür entwickelten Crossover-Designs eine kontrollierte Studie über den Lernerfolg des Programms „MICROPAT – Interaktiver Atlas der Histopathologie“ durchgeführt.

MICROPAT – Interaktiver Atlas der Histopathologie

Bei der Vermittlung von histopathologischen¹ Lerninhalten ist neben dem Verständnis wichtiger pathophysiologischer Mechanismen die Fähigkeit des „Muster-Erkennens“ (pattern recognition [25]) von grundlegender Bedeutung. Die Befähigung, am unbekanntem Präparat pathognomonische² Gewebe- und Krankheitsmuster zu erkennen, die kursorisch an Beispielen gezeigt und erlernt werden, ist eine wichtige Voraussetzung für diagnostische Entscheidungen in der Pathologie. Darum nimmt die Vermittlung der Histopathologie auch einen sehr breiten Raum im medizinischen Curriculum ein. Obwohl klassische Lehrformen wie Hauptvorlesung und Praktika weiterhin dominieren, erhofft man sich in bildorientierten Fächern wie der Pathologie zunehmend einen Nutzen von CBT für die Wissensvermittlung [1]. So kommen in den USA bereits in zwei Drittel der „medical schools“ Computerlernprogramme als Er-

gänzung zum klassischen Lehrmethodenspektrum zum Einsatz [20].

Auch am Universitätsklinikum Freiburg wurde auf der Basis eines Hypertextsystems³ [4,29] zur Unterstützung des Kurses „Allgemeine Pathologie“ für Medizinstudenten ein elektronischer Atlas für Histopathologie erstellt. Das Programm MICRO-PAT (**M**icroscopical **P**athology) basiert auf mehr als 1300 histologischen (neben etlichen makroskopischen und elektronenmikroskopischen) Fotografien, die durch erklärende Texte ausführlich und exakt beschriftet sind (s. Abb.1). Mehr als 600 weitere Texte werden in Form von Definitionen, Kurzcharakteristiken und Glossareinträgen zur Verfügung gestellt.

Der Zugriff auf die über 150 verschiedenen Krankheiten erfolgt über die topografische Zuordnung zu Organsystemen, über kausalpathogenetische Aspekte und über eine ausführliche Indexfunktion. Dem Atlas liegt ein Hyperlinksystem als Organisationsprinzip zugrunde. Wichtige Elemente der Navigation im Sinne des virtuellen Mikroskopierens sind verschiedene Bildauswahlfunktionen und eine Zoomfunktion, die die Vergrößerung definierter Bildausschnitte im Rahmen einer Bildsequenz ohne qualitative Einbußen ermöglicht. Die Bedienung erfolgt mit der Maus und weitgehend auch mit der Tastatur. Weitere Funktionen sind das Anbringen eigener Anmerkungen, das Ausdrucken bestimmter Programminhalte, eine Kurzhilfe und eine Übersicht über gebräuchliche Färbemethoden in der Histopathologie.

Evaluation

Im Rahmen der experimentellen Erstellung von Lehr- und Lernmitteln ist deren Qualitätsmessung von besonderer Bedeutung. Um eine Nutzen-Analyse mit verlässlichen Daten durchzuführen, bedarf es jedoch quantitativer Daten aus einer empirischen Erhebung [9]. Die Frage nach dem Nutzen neuer Medien lässt sich am besten in einem Vergleich mit den bisher benutzten, konventionellen Medien beantworten. Da kein geeignetes Modell zur Überprüfung des Nutzens von MICRO-PAT

¹ Die pathologische Anatomie, auch Histopathologie genannt (gr. ἵστος=Gewebe – gr. πάθος=Leiden – gr. λόγος=Lehre), erforscht, vor allem mit mikroskopischen Techniken, die krankheitsbedingten Zell- und Gewebsveränderungen an Untersuchungsmaterial, das durch Biopsien, operative Eingriffe oder Leicheneröffnungen gewonnen wird.

² Pathognomonisch = für eine Krankheit spezifisch.

³ Das Vernetzen von Texten in einem Multimedia-System führt zu einem Hypertext. Sind dabei mehrere Medientypen beteiligt, spricht man von Hypermedia. Die einzelnen Elemente des Hypermedia sind über Hyperlinks in Form markierter Wörter oder Symbole untereinander verknüpft.

zur Verfügung stand, entwickelten wir ein Modell zur Qualitätsprüfung und testeten es unter kontrollierten Bedingungen.

Studienziel

Ziel der Studie war es, den Lernerfolg des Computerprogramms MicroPat im Vergleich zu den bei der Vorbereitung auf die Histopathologie-Testate üblicherweise verwendeten Lehrbüchern zu messen. Zu diesem Zweck wurde in einer realen Prüfungssituation, nämlich dem Präparate-Erkennen am Mikroskop, ein Testat durchgeführt. Darüber hinaus sollten mit der Studie die individuelle Sicherheit der Probanden bei der Diagnosefindung, die Computererfahrung und zusätzlich ihre subjektive Einschätzung des Programms mit Hilfe eines Fragebogens erfasst werden.

Studiendesign

72 freiwillige Probanden aus dem ersten Semester des klinischen Studienabschnitts wurden randomisiert in zwei Gruppen mit jeweils 36 Probanden eingeteilt. Eine Gruppe sollte den Themenbereich A (Prostatakarzinom, invasiv lobuläres Mammakarzinom, invasiv duktales Mammakarzinom und Komedokarzinom) ausschließlich mit Büchern und den Themenbereich B (Adenomyomatose, Mastopathie, Fibroadenom und Gallertkarzinom) mit dem Computerprogramm bearbeiten. Die andere Gruppe sollte umgekehrt verfahren. In der Studie kam somit ein Crossover-Versuchsplan⁴ [21,33] zum Einsatz (s. Abb. 2).

Jeder Student hatte für das Studium mit dem Computer und mit den Büchern nacheinander je 50 Minuten zur Verfügung. Um Residualeffekte⁵ [22] zwischen der Reihenfolge der Bearbeitung mit Büchern und dem Computer zu vermeiden, wurden die Lernfolge in den beiden Gruppen variiert. Die Aufteilung erfolgte auch hier randomisiert zu gleichen Teilen mit jeweils 18 Probanden. Die Vorbereitung mit Büchern fand in einem gesonderten Raum statt. Nach der Vorbereitung mussten die Studenten aus den beiden Themenbereichen je drei unbekannte Präparate mit Hilfe des Mikroskops diagnostizieren. Es wurden vier verschiedene Prüfungsformen (vier Tableaus mit sechs Präparaten, bestehend aus jeweils drei Präparaten aus jedem Themenbereich) verwendet, die jeweils gleich häufig eingesetzt wurden; die Zuordnung der Probanden zu den einzelnen Prüfungsformen erfolgte zufällig. Die Probanden wurden nach der Diagnose und zusätzlich auf einer 3-stufigen Ska-

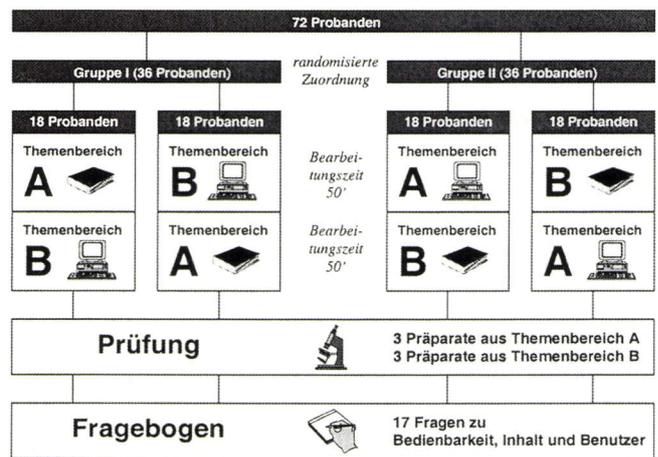


Abb. 2 Ablauf der Evaluationsstudie von MICROPAT als Crossover-Versuchsplan.

Tab. 1 Lineares Score-Modell zur Gewichtung der individuellen Sicherheitsangaben des Probanden.

Angabe	Gewichtung
„Ich bin bei meiner Diagnose ganz sicher.“	3 Punkte
„Ich bin mir bei meiner Diagnose nicht so ganz sicher.“	2 Punkte
„Ich habe bei meiner Diagnose geraten.“	1 Punkt

la (s. Tab. 1) nach der Sicherheit ihrer Aussage gefragt. Jede richtige Diagnose wurde mit einem Punkt bewertet. Die Sicherheitsangaben der Probanden wurden mit den in Tab. 1 aufgeführten Punkten je nach richtiger oder falscher Lösung mit positivem oder negativem Vorzeichen in die Auswertung mit einbezogen.

Weiterhin sollten die Studenten per Fragebogen 17 Fragen über die eigenen Computerkenntnisse, zur Beurteilung der Bilder, der Texte und des Programms allgemein beantworten. Die Bewertung erfolgte auf einer Ordinalskala von eins (= „sehr gut/absolut richtig“) bis fünf (= „sehr schlecht/absolut falsch“).

Zur Schaffung standardisierter Rahmenbedingungen wurden für alle Probanden die gleiche leistungsfähige Rechnerplattform sowie Standardliteratur der „Allgemeinen Pathologie“ und der Histopathologie bereitgestellt. Die Buchgruppe erhielt Kapitelangaben, die Computergruppe eine Kurzhilfe zum Programm. Während des gesamten Evaluationszeitraums war in jedem Raum eine Aufsichtsperson anwesend.

Ergebnisse

Zunächst wurden die individuellen Ergebnisse jedes einzelnen Probanden beim Computer- und Buch-Lernen miteinander verglichen. Es stellte sich heraus, dass bei der ausschließlichen Berücksichtigung von *Richtig- und Falschdiagnosen* das Computer-Lernen mit MICROPAT einen leichten Vorteil im Vergleich zur konventionellen Testatvorbereitung mit Büchern brachte: 36,1% der Studenten erzielten mit MICROPAT bessere Ergebnisse, 25,0% der Probanden schnitten mit Büchern besser ab und 38,9% der Studienteilnehmer waren in beiden Berei-

⁴ Ein Crossover-Versuchsplan, der auch als Changeover-Plan oder Überkreuzungsplan bezeichnet wird, hat in seiner einfachsten Anwendung zwei Perioden, d.h. zwei Abschnitte (hier: Themenbereich A und Themenbereich B), in denen zwei unterschiedliche Lernformen (hier: Lernen mit Computern und Lernen mit Büchern) getestet werden. Jeder Proband absolviert beide Lernszenarien nacheinander, wobei ein Teil der Probanden die Reihenfolge Computer/Buch und ein anderer Teil die umgekehrte Reihenfolge Buch/Computer einhält. Die Aufteilung der Probanden auf die beiden Gruppen erfolgt zufällig.

⁵ Residualeffekte werden auch als Carryover-, Nach-, Überhangs- oder Übertragungseffekte bezeichnet. Sie sind dann vorhanden, wenn der Ergebnisunterschied (hier: zwischen Computer und Buch) in der ersten Periode (hier: im Themenbereich A) vom Ergebnisunterschied in der zweiten Periode (hier: Themenbereich B) differiert.

Tab. 2 Ergebnisse des besseren Abschneidens nach Lernmethode.

Methode	Richtig-Falsch-Analyse	Score-System
besser mit Computer	0,361	0,583
besser mit Buch	0,250	0,319
kein Unterschied	0,389	0,097

chen gleich gut (s. Tab. 2). Diese Ergebnisse sind jedoch nicht signifikant.

Gewichtet man die Resultate mit den individuellen *Sicherheit-Scores*, so lässt sich zeigen, dass 58,3% der Probanden mit dem Computer besser abschnitten, 31,9% mit Büchern besser zu recht kamen und 9,7% der Studenten in beiden Prüfungsabschnitten gleich gut waren (s. Tab. 2). Nach dem Vorzeichen-test bzw. McNemar-Test sind die Ergebnisse auf einem Niveau von $\alpha = 0,05$ signifikant. Der Wilcoxon-Test konnte die Signifikanz ebenfalls bestätigen.

Betrachtet man die erreichten *Punktzahlen*, zeigt sich ein ähnliches Bild: Bei der Richtig-Falsch-Analyse zeigt sich ein nominaler, aber nicht signifikanter Vorteil für das Computer-Lernen mit MICROPAT im Vergleich zum Buch, der unter Berücksichtigung der Sicherheit-Scores deutlich signifikant wird. In diesem Zusammenhang hat sich auch herausgestellt, dass das neue Lernmedium „Computer“ die Diagnosesicherheit der Probanden signifikant beeinflusst (Wilcoxon, $\alpha = 0,05$) und sie zu einer höheren Sicherheit verleitet, ohne jedoch die Anzahl richtiger Diagnosen zu erhöhen (s. Tab. 3).

Hinweise auf ein besseres Abschneiden *computererfahrener Probanden* in Bezug auf das Lernen mit dem Computer fanden sich nicht (Spearman'scher Korrelationskoeffizient 0,04); Studenten mit Computererfahrung erzielten generell bessere Resultate (Score-System: 0,26 / Richtig-Falsch-Analyse: 0,33, jeweils Korrelationskoeffizient nach Spearman). Die Ergebnisse sind auf einem Niveau von $\alpha = 0,05$ signifikant. Wechselwirkungen mit den Variablen „Prüfungstag“, „Reihenfolge der Bearbeitung“ und „Auswahl der Prüfungspräparate“ traten nicht auf.

⁶ Man spricht dann von einem Periodeneffekt, wenn sich der Mittelwert der ersten Periode (hier: erreichte Punktzahl im Themenbereich A) vom Mittelwert der zweiten Periode (hier: erreichte Punktzahl im Themenbereich B) unterscheidet. Gründe für Periodeneffekte sind z.B. eine Ermüdung der Probanden, ein Nachlassen der Motivation oder eine Veränderung der Prüfungsbedingungen.

Gravierende *Periodeneffekte*⁶ [21] während der Tage der Durchführung der Studie, konnten nicht nachgewiesen werden. Da bei einem Crossover-Versuchsplan immer interindividuelle Vergleiche vorgenommen und Wertdifferenzen berechnet werden, können Periodeneffekte ohnehin weitgehend unberücksichtigt bleiben, denn der Wert der Differenz ist unabhängig vom absoluten Wertniveau. Im Gegensatz dazu müssen *Residualeffekte* [22] je nach Ausprägung besonders genau untersucht werden. In der MICROPAT-Studie traten weder signifikante negative noch signifikante positive Residualeffekte auf.

Die Auswertung des zusätzlichen *Fragebogens* ergab, dass die Bedienbarkeit und die Schnelligkeit des Programms durchschnittlich mit der Note 1,5 bewertet wurden; die Bilder wurden mit 1,8 Punkten, Umfang und Qualität der Texte mit 2,4 Punkten bewertet. Insgesamt fällt auf, dass für die einzelnen Elemente innerhalb der Bereiche Programmhandling, Bilder, Texte immer sehr ähnliche Einschätzungen gegeben wurden. Mit einem Wert von durchschnittlich 1,6 hinterließ das Programm einen durchwegs guten bis sehr guten Gesamteindruck.

Diskussion

Mit der vorliegenden Studie konnte ebenso wie in anderen Untersuchungen [27] demonstriert werden, dass neben der Auswertung von subjektiven Benutzerfragebogen objektive empirische Daten zur Lernerfolgskontrolle eine bessere Nutzenanalyse eines Computerlernprogramms erlauben.

Aus einer Gesamtheit von 237 Studenten nahmen 72 an der Studie teil, was unter den vorliegenden statistischen Bedingungen zwar eine geringe Zahl ist, im Vergleich zu anderen Studien aber eine relativ große Stichprobe darstellt. Auch wenn die Stichprobe nicht repräsentativ ist (die Probanden meldeten sich freiwillig), macht sie dennoch relevante Aussagen über eine bedeutende Untermenge von 30%.

Der Lernerfolg wurde in dieser Studie an einer Prüfungssituation gemessen, die der realen Prüfung am Ende des regulären Histopathologie-Kurses entspricht. Andere Studien, die den Lernerfolg an den Examensergebnissen gemessen haben [24], sind aufgrund großer Zeitdifferenzen zwischen Lern- und Prüfungsphase eher kritisch zu bewerten. Eine methodenspezifische Trennung des Lernerfolgs ist hierbei nicht mehr problemlos möglich und es können Interferenzen mit anderen Faktoren [2,26] auftreten. Dies ist wahrscheinlich auch ein wichtiger Grund, dass Längsschnittstudien schwer durchzuführen sind. Wenn verschiedene Methoden mit Hilfe realer Examina prospektiv getestet werden und dabei die Gruppentrennung strikt

Tab. 3 Arithmetische Mittel und Signifikanzen der Ergebnisse bei drei Prüfungspräparaten nach Lernmethode.

Methode	korrekte Diagnosen [0;3]	Score-System [- 9;9]	Sicherheit [- 9;9]
Computer	1,97	2,59	6,35
Buch	1,85	1,53	5,92
Buch und Computer	3,82	4,12	12,27
	nicht signifikant	signifikant	signifikant
	p = 0,37 (Wilcoxon)	p = 0,04 (Wilcoxon)	p = 0,03 (Wilcoxon)
	p = 0,29 (McNemar)	p = 0,02 (McNemar)	

eingehalten wird, ist die Chancengleichheit im Examen nicht mehr gewährleistet. Vorzuziehen sind daher authentische Szenarien wie in der MICROPAT-Studie, die einer wirklichen Prüfungsform entsprechen und die für die Teilnehmer der Studie keinen Nachteil bedeuten [6].

Auch wenn die Sicherheit von Aussagen offenbar durch das Medium „Computer“ beeinflusst wird, hat sich die Berücksichtigung von Sicherheit-Scores, die auch bei anderen Fragestellungen Verwendung finden [16], dennoch als zweckmäßig herausgestellt, da die Ergebnisse in ihrer Qualität differenzierter betrachtet werden können und Tendenzen in der Fragenbeantwortung klarer werden. Allerdings sollte die Score-Betrachtung der rein objektiven Lernerfolgskontrolle im Sinne einer Richtig-Falsch-Analyse in ihrer Wertigkeit untergeordnet werden, nicht zuletzt, da durch eine entsprechende Wahl der Punkteskala Signifikanzen künstlich erzeugt werden können.

Ein Vortest-/Nachtest-Design, wie in anderen Studien verwendet [5,35], erschien uns nicht notwendig, da das abgefragte Wissen als den Probanden weitgehend unbekannt vorausgesetzt werden konnte. Ein Eingangstest zur Ermittlung des Vorwissens ist jedoch dann anzuraten, wenn weitere Untersuchungen an der Gruppe dieser Probanden durchgeführt werden oder wenn ein Programm zu einem festen Bestandteil des Curriculums geworden ist [29]. Vollständigerweise sollte die Untersuchung dann zusätzlich die subjektive Bewertung des eigenen Vorwissens und des persönlichen Lernerfolgs enthalten [27].

Immer wieder findet man in Studien die Feststellung, dass die Verwendung des Computers zu einer Zeitersparnis beim Lernen führt [19]; es gibt aber auch gegenteilige Auffassungen [27]. Bei der MICROPAT-Studie lag das Ziel nicht im Nachweis einer Zeitökonomie beim Lernen mit Computern, sondern in der richtigen Diagnosestellung. Um individuellen Lernstrategien in den Lernphasen Raum zu geben, persönliche Nachteile bei der Bedienung von Computern auszugleichen und die Chancengleichheit zu erhalten, wurde den Probanden ausreichend und für ein gutes Ergebnis notwendige Zeit [2] gewährt. Es wurde allerdings auch hier beobachtet – wenn auch nicht statistisch nachweisbar –, dass die Computer-Lektionen schneller bewältigt wurden als die Buchlektionen. Andere Autoren konnten beim Vergleich von Evaluationsresultaten zeigen, dass die mit traditionellen Medien und CBT erzielten Ergebnisse generell gleich sind, wenn für das Studium der Themen jeweils dieselbe Zeit zur Verfügung gestellt wurde [17]. Diese Tendenz ist auch in der vorliegenden Studie deutlich erkennbar.

Die Einbeziehung eines zusätzlichen subjektiven Benutzerfragebogens hat sich bewährt, da hiermit zusätzliche Parameter, die die Qualität bzw. die Akzeptanz des Programms wesentlich beeinflussen oder die persönliche Merkmale des Probanden betreffen, erfragt werden und Fragebogenvariablen mit den Ergebnissen des Testats korreliert werden konnten. Die Validität und Reliabilität solcher Fragebogen ist dennoch wegen der Subjektivität der Angaben mit Zurückhaltung zu bewerten. Dass die durchaus differenzierten Ergebnisse bei der Bewertung des Programms so positiv ausgefallen sind, mag sicherlich am Wohlwollen der Probanden gelegen haben, ist aber auch Ergebnis der konsequenten Berücksichtigung von Quali-

tätskriterien [31,32] und der Einbindung von Medizinstudenten während der Programmentwicklung [24].

MICROPAT kann offensichtlich im Vergleich mit herkömmlichen Medien wie Büchern gut bestehen. Meta-Evaluationen über neue Medien im Vergleich zu konventionellen haben für andere CBT-Programme ähnliche Ergebnisse ermittelt [14]. Es wäre schließlich auch vermessen, eine deutliche Steigerung des Lernerfolgs mit Computern zu erwarten, da man damit bisherigen erfolgreichen Lernkonzepten ihre Berechtigung absprechen würde. Die Aneignung von Wissen kann auf verschiedene Weise erfolgen. Das Kennzeichen adaptiver Lernmedien ist die Beherrschung vieler Unterrichtsmethoden, die je nach Lerntyp verwendet werden sollten [7,29].

Für die Scheu vor einer vergleichenden Medienforschung gibt es viele Gründe [3]. Die Zunahme der Anzahl eingesetzter CBT-Programme [20] und die schlechte Vergleichbarkeit [11] der wenigen vorhandenen, kontrollierten Evaluationsstudien [27] zeigen jedoch die Notwendigkeit von kontrollierten, standardisierten und vergleichbaren Studien. Mittels Crossover-Design und Erfolgskontrolle in realitätsnahen Prüfungsszenarien wurde eine leicht durchführbare und dennoch aussagekräftige Validierungsmethode entwickelt. Die Vorteile des Modells liegen in der Reduzierung der Probandenzahlen um mindestens 50%, da jeder Proband gleichzeitig Test- und Kontrollperson ist; die Daten sind schnell zu erheben, und die Ergebnisse zeigen eine hohe Kontrollierbarkeit der Störgrößen [13], wie z. B. der Perioden- oder Residualeffekte. Die MICROPAT-Studie ist auch aufgrund der Gleichverteilung der Probanden auf die untersuchten Untergruppen sehr aussagekräftig. Insgesamt konnte ein großes Maß der Standardisierung beim Studiendesign erreicht werden, was anderen Studien nur teilweise gelingt [5].

Ausblick

Nicht alle Fragen konnten mit diesem Modell einer Evaluationsstudie beantwortet werden. Das Modell wurde nicht auf seine Stabilität in Bezug auf unterschiedliche Lernformen innerhalb des Computer-Lernens getestet. Es wäre deshalb gut, das Evaluationsdesign auch auf Lernsysteme anzuwenden, die weniger durch das selbständige explorative „browsing“ [8] durch das Programm, sondern durch so genannte „guided tours“ [36] mit einem geringeren Grad der Selbststeuerung gekennzeichnet sind. Der Vergleich mit weiteren Lernszenarien, wie z. B. mit Seminaren und Vorlesungen [28], oder mit einer Kombination unterschiedlicher Lehrmethoden erscheint sinnvoll [12]. So wurde u. a. in einer Studie gezeigt, dass Pathologen die visuelle Wissensvermittlung der auditiven bei der Ausbildung am Mikroskop vorziehen [15] oder die Kombination aus Multimedia-Programm und Vorlesung bessere Lernergebnisse liefert als das Medium „Vorlesung“ allein [10]. Unter Umständen ist aber der methodische Ansatz einer Evaluation auch problematisch, da zwei unterschiedliche Medientypen, wie in der MICROPAT-Studie Buch und Computer, miteinander verglichen werden. Eine Bestimmung moderierender Größen, wie z. B. habituelles Lernverhalten oder Examensangst, sollte dabei auch berücksichtigt werden. Neuere Untersuchungen haben gezeigt, dass der Einsatz von CBT u. a. auch das generelle Lernverhalten in Bezug auf soziale Aspekte verändert [23]. Darüber hinaus sollten auch Versuchsanordnungen diskutiert werden, die mehr als zwei Stichproben oder – anders als im

vorliegenden Fall – mehr als zwei Lernszenarien in einem Modell berücksichtigen. Da Lernleistung und Studierzeit den Kern von Evaluationen bilden, könnte die Variation der Zeitvorgaben zu unterschiedlichen Lerneffekten führen. Daneben ist die Stabilität des Systems bei weniger homogenen Datenbeständen in Frage zu stellen; bei der vorliegenden Studie konnte man von einem Idealfall ausgehen, da alle Daten ohne fehlende Werte vorliegen und fast alle anderen Einflussgrößen, wie z.B. Prüfungsform oder Prüfungsfolge, absolut gleichmäßig verteilt waren. Weitere Einschränkungen betreffen unter Umständen die Praktikabilität des Systems bei großen Probandenzahlen und die Gültigkeit des Modells für unterschiedliche Typen medizinischer CBT-Programme. Es stellt sich auch die Frage nach der Anwendbarkeit des Modells bei differenzierteren Zielgrößen mit mehr Merkmalsausprägungen als einer richtigen oder falschen Diagnose. Einer weiteren Entwicklung bedarf zudem die Abstimmung des Fragebogens mit dem Versuchsprinzip des Crossover-Designs. Mehr Kontrollfragen zur Verlässlichkeit der subjektiven Angaben des Probanden oder zusätzliche Fragen, die Korrelationen mit den objektiven Ergebnissen der Studie erlauben, könnten integriert werden.

Fazit

Der Weg zu *Evidence-Based Medical Education* ist schwierig, aber möglich. Es wird zu bewerten sein, ob der hier vorgestellte [3] und erfolgreich angewandte Ansatz zur Entwicklung und Evaluation von CBT-Programmen Modellcharakter haben kann und sich auf andere, vor allem bildorientierte Disziplinen der Medizin übertragen lässt. Eine Weiterentwicklung des verwendeten *Crossover-Evaluationsmodells* bezüglich seiner Komplexität, Reliabilität und Validität ist sicherlich anzustreben.

MICROPAT hat sich im Vergleich zu herkömmlichen Medien wie Büchern behaupten können und ist als *Supplementärmedium* [18,24] zum klassischen Lernmethodenspektrum gut geeignet. Die Erkenntnisse aus dem Freiburger Projekt stehen damit im Einklang mit Beobachtungen im Pathologie-Unterricht an amerikanischen „medical schools“ [20]. Eine *curriculare Einbettung* von MICROPAT erscheint nicht zuletzt wegen der hohen Akzeptanz unter den Studenten als gerechtfertigt [34].

Weitere Informationen zur vorgestellten Arbeit und zu MICROPAT sind unter dem Titel: Auhuber TC. Entwicklung und Evaluation eines computergestützten Lernsystems in der Medizin. MicroPat – ein interaktiver Atlas der Histopathologie mit adaptierbarem Tutor. Europäische Hochschulschriften VII/D/31. Frankfurt a.M.: Lang, 1998; erschienen oder unter der Adresse: <http://www.imbi.uni-freiburg.de/medinf/micropat.htm> im Internet abrufbar.

Literatur

- Andrew SM, Benbow EW. Conversion of a traditional image archive into an image resource on compact disc. *J Clin Pathol* 1997; 50, 7: 544–547
- Bock R, Herth G, Rubly M, Schonecke O. Computergestützter Anatomie-Unterricht – erste Schritte mit einem neuen Medium. *magazin forschung* 1996; 1: 19–21
- Campbell JK, Johnson C. Trend spotting: fashions in medical education. *Brit Med J* 1999; 318: 1272–1275
- Conklin J. Hypertext: An Introduction and Survey. *IEEE Computer* 1987; 20: 17–4
- Devitt P, Cehic D, Palmer D. Computers in medical education 2. Use of a computer package to supplement the clinical experience in surgical clerkship: an objective evaluation. 1998; 68, 6: 428–431
- Devitt P, Worthley S, Palmer E, Cehic D. Evaluation of a computer based package on electrocardiography. *Aust N Z J Surg* 1998; 28, 4: 432–435
- Devitt P, Palmer E. Computer-aided learning: an overvalued educational resource? *Med Educ* 1999; 33, 2: 136–139
- diSessa AA, Hoyle C, Noss R (eds). *Computers and Exploratory Learning*. NATO ASI Series. F: Computer and Systems Sciences. 146. Berlin, Heidelberg, New York: Springer; 1995
- Eitel F. Der Stellenwert Neuer Medien in der Aus- und Weiterbildung. *Viszeralchirurgie* 1998; 88: 139–145
- Elves AW, Ahmed M, Abrams P. Computer-assisted learning; experiences at the Bristol Urological Institute in the teaching of urology. *Brit J Urol* 1997; 80, suppl 3: 59–62
- Fieschi D, Fieschi M, Soula G, Degoulet P. Évaluation des méthodes d'enseignement assisté par ordinateur: À propos de vingt-six études comparatives publiées entre 1989 et 1992. *Pathol Biol* 1994; 42: 183–190
- Finley JP, Sharratt GP, Nanton MA, Chen RP, Roy DL, Paterson G. Auscultation of the heart: a trial of classroom teaching versus computer-based independent learning. *Med Educ* 1998; 32, 4: 357–361
- Friedman CP, Wyatt JC. *Evaluation Methods in Medical Informatics*. Berlin, Heidelberg, New York: Springer; 1997
- Glowalla U, Häfele G. Einsatz elektronischer Medien: Befunde, Probleme und Perspektiven. In: Issing L, Klimsa P (Hrsg): *Information und Lernen mit Multimedia*. Weinheim: Psychologie-Verlags-Union; 1995: 415–434
- Gray E, Duvall E, Sprey J, Bird CC. Pathologists dislike sound? Evaluation of a computerised training microscope. *J Clin Pathol* 1998; 51, 4: 330–333
- Hölker M, Breukelmann D. Vier Jahre Kursus „Ärztliche Basisfertigkeit“: Ein Erfahrungsbericht. *Med Ausbildung (Gesundheitswesen 60)* 1998; 15, 1: 18–22
- Jelovsek FR, Adebajo L. Learning Principles as Applied to Computer-assisted Instruction. *MD Computing* 1993; 10: 165–172
- Kallinowski F, Mehrabi A, Glückstein C, Benner A, Lindinger M, Hashemi B, Leven FJ, Herfarth C. Computer-based training – a new method in surgical education and continuing education. *Chirurg* 1997; 68, 4: 433–438
- Kallinowski F, Eitel F. New approaches to surgical education and continuing education. *Chirurg* 1998; 69, 12: 1323
- Kumar K, Daniel J, Doig K, Agamanolis D. Teaching of pathology in United States medical schools, 1996/1997 survey. *Hum Pathol* 1998; 29, 7: 750–755
- Lehmacher W. Zwei Stichproben von Verlaufskurven und Crossover-Versuche. *Medizinische Informatik und Statistik* 67. Heidelberg: Springer; 1986: 77–85
- Lehmacher W. Residual Effects in Cross-over Trials. In: Vollmar J, Hothorn LA (eds): *Cross-over clinical trials*. Biometrics in the pharmaceutical industry 7. Stuttgart, Jena, New York: Gustav Fischer; 1997: 41–65
- Lehmann HP, Freedman JA, Massad J, Dintzis RZ. An ethnographic, controlled study of the use of a computer-based histology atlas during a laboratory course. *J Am Med Inform Assoc* 1999; 6, 1: 38–52
- Lyon HC, Ueberla K, Batschkus M, Dugas M, Fischer MR, Hohnloser J, Eitel F, Weindl A, Shultz EK. Correlation between medical student use of Computer-Based Instruction (CBI) and high exam scores: Lessons learned from 10 years of CBI development and use. In: Adler M, Dietrich JW, Holzer MF, Fischer MR (Hrsg): *Computer Based Training in der Medizin. Technik – Evaluation – Implementation*. Aachen: Shaker; 1998: 105–112

- ²⁵ O'Brien MJ, Sotnikov AV. Digital Imaging in Anatomic Pathology. *Am J Clin Pathol* 1996; 106 S1: S25–S32
- ²⁶ Reigeluth CM. Instructional Design: What is it and why is it? In: Reigeluth CM (ed): *Instructional Theories and Models: An Overview of Their Current Status*. Hillsdale: Lawrence Erlbaum; 1983: 3–36
- ²⁷ Richardson D. Student perceptions and learning outcomes of computer-assisted versus traditional instruction in physiology. *Am J Physiol (Adv Physiol Educ)* 1997; 273: S55–S58
- ²⁸ Rogers DA, Regehr G, Yeh KA, Howdieshell TR. Computer-assisted learning versus a lecture and feedback seminar for teaching a basic technical skill. *Am J Surg* 1998; 175, 6: 508–510
- ²⁹ Schulmeister R. *Grundlagen hypermedialer Lernsysteme: Theorie – Didaktik – Design*. Bonn: Addison-Wesley; 1996
- ³⁰ Schulz S, Schrader U, Klar R. Computer-based Training and Electronic Publishing in the Health Sector: Tools and Trends. *Methods Inform Med* 1997; 36: 149–153
- ³¹ Schulz S, Auhuber T, Schrader U, Klar R. Quality Criteria for Electronic Publications in Medicine. In: Mantas J (ed): *Advances in Health Telematics Education: A Nightingale Perspective*. *Studies in Health Technology and Informatics* 51. IOS Press: Amsterdam; 1998: 217–226
- ³² Schulz S, Klar R, Auhuber T, Schrader U, Koop A, Kreutz R, Oppermann R, Simm H. Quality Criteria for Electronic Publications in Medicine. <http://www.imbi.uni-freiburg.de/medinf/gmdsqc/e.htm>. 1.2.1999
- ³³ Senn S. *Cross-Over Trials in Clinical Research*. Chichester, New York: Wiley; 1993
- ³⁴ Stensaas SS. Animating the curriculum: integrating multimedia into teaching. *Bull Med Libr Assoc* 1994; 82, 2: 133–139
- ³⁵ Teichman JM, Richards J. Multimedia to teach urology to medical students. *Urology* 1999; 53, 2: 267–270
- ³⁶ Trigg RH. Guided Tours and Tabletops: Tools for Communicating in a Hypertext Environment. *ACM Trans Off Inf Syst* 1988; 4: 398–414
- ³⁷ Wilkes M, Bligh J. Evaluating educational interventions. *Brit Med J* 1999; 318: 1269–1272

cand. med. Thomas C. Auhuber

Universitätsklinikum Freiburg
Pathologisches Institut
Albertstraße 19
79104 Freiburg i. Br.

E-mail: at@micropat.de

URL: <http://www.micropat.de/auhuber.htm>