

Ein Modell zur Evaluation medizinischer CBT-Programme

T. Auhuber, S. Schulz, U. Schrader, R. Klar¹

Zusammenfassung

Computergestützte Lernhilfen halten in Deutschland erst langsam Einzug in die universitäre Ärzteausbildung. Zur Förderung der Verbreitung und Verbesserung dieser neuen Medien, aber auch zur besseren Auswahl von Lernsystemen ist es für die Sicherung und die Weiterentwicklung von Bildungsstandards notwendig, die Qualität dieser Lehr- und Lernmittel zu prüfen. Im Rahmen einer kontrollierten Studie an einem an der Universität Freiburg entwickelten Lernsystem für Histopathologie wurde ein Crossover-Versuchsplan erarbeitet, der den Lernerfolg bei unterschiedlicher Vorbereitung zweier Themenbereiche durch das Lernsystem oder durch Bücher untersucht. Durch eine Erfolgskontrolle am Mikroskop wurde die Lernleistung gemessen und auf einer dreistufigen Skala die subjektive Sicherheit bei der Diagnosenstellung erfaßt. Dieser Sicherheitsscore war bei Vorbereitung mit dem Computer signifikant höher, was sich jedoch nicht in einer besseren Lernleistung niederschlug. Hinweise für einen Vorteil computererfahrener Probanden bezüglich des Abschneidens mit dem Lernsystem fanden sich nicht.

Einleitung

Auch wenn computerunterstützte Lernsysteme in Deutschland erst allmählich größere Verbreitung finden, ist es notwendig, sich bereits jetzt über die Qualitätsmessung und -prüfung solcher Systeme Gedanken zu machen. Gerade im Bereich der Medizin ist diese Frage durch die Verantwortung der Lehre für das zukünftige ärztliche Handeln besonders wichtig [5]. Der Markt für Lehrmedien ist durch eine Publikationsflut [3] mittlerweile sehr unübersichtlich geworden, was die Selektion guter Lehrmedien erheblich erschwert. Mit dieser Studie wird ein Modell zu Quali-

tätsprüfung für medizinische CBT-Programme vorgestellt, das am Beispiel des elektronischen Histopathologie-Atlanten MicroPat getestet wurde.

MicroPat ist ein histopathologisches Lernsystem zur Unterstützung des Kurses „Allgemeine Pathologie“ für Medizinstudenten der Universität Freiburg. Das System ist in zwei verschiedene Module aufgeteilt: ein Atlas für Histopathologie (Abb. 1) und ein Lernprogramm mit Quizmöglichkeit [3].

Enthalten sind über 1000 digitalisierte Diapositive aus dem Bildarchiv des Pathologischen Instituts der Universität Freiburg, die die wesentliche Grundlage des Systems bilden.

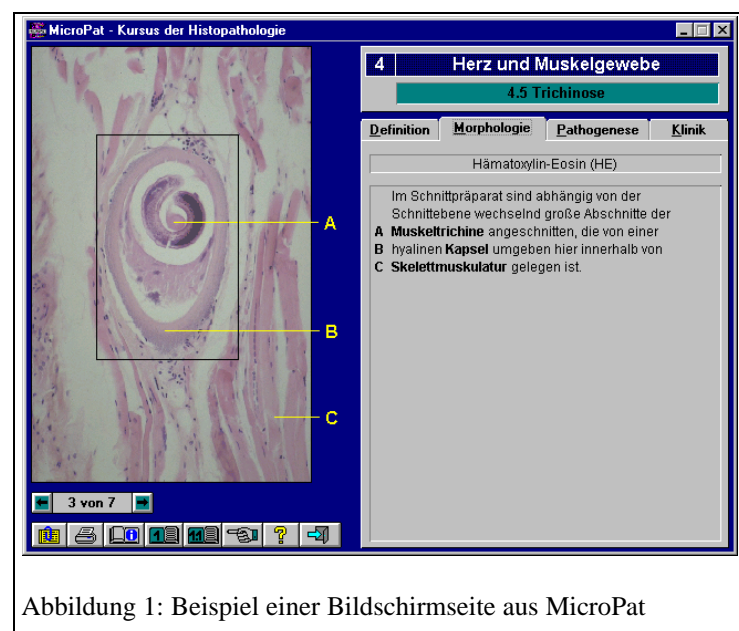


Abbildung 1: Beispiel einer Bildschirmseite aus MicroPat

Zwischen den einzelnen Vergrößerungsstufen läßt sich über eine Zoomfunktion navigieren. Die Texte sind thematisch untergliedert in allgemeine Abschnitte wie Definition, Pathogenese

¹ Abteilung Medizinische Informatik, Universitätsklinikum Freiburg i. Br., Stefan-Meier-Str. 26, D-79104 Freiburg i. Br., http://www.imbi.uni-freiburg.de/medinf/mi_home.htm

und Klinik, die unabhängig vom gerade angezeigten Bild angesehen werden können. Die morphologischen Bildbeschreibungen sind für jede Abbildung spezifisch erstellt worden. Anmerkungen, Bildverweise und Erklärungen zu Färbungen sind ebenso wie Querverweise als Hyperlinks realisiert. Eine Indexfunktion erlaubt eine Schlagwortsuche. Gemäß dem Aufbau des Mikroskopierkurses wurde das Text- und Bildmaterial topographisch angeordnet. Die Benutzerführung orientiert sich an gängigen GUI (graphical user interface)-Standards.

Studiendesign

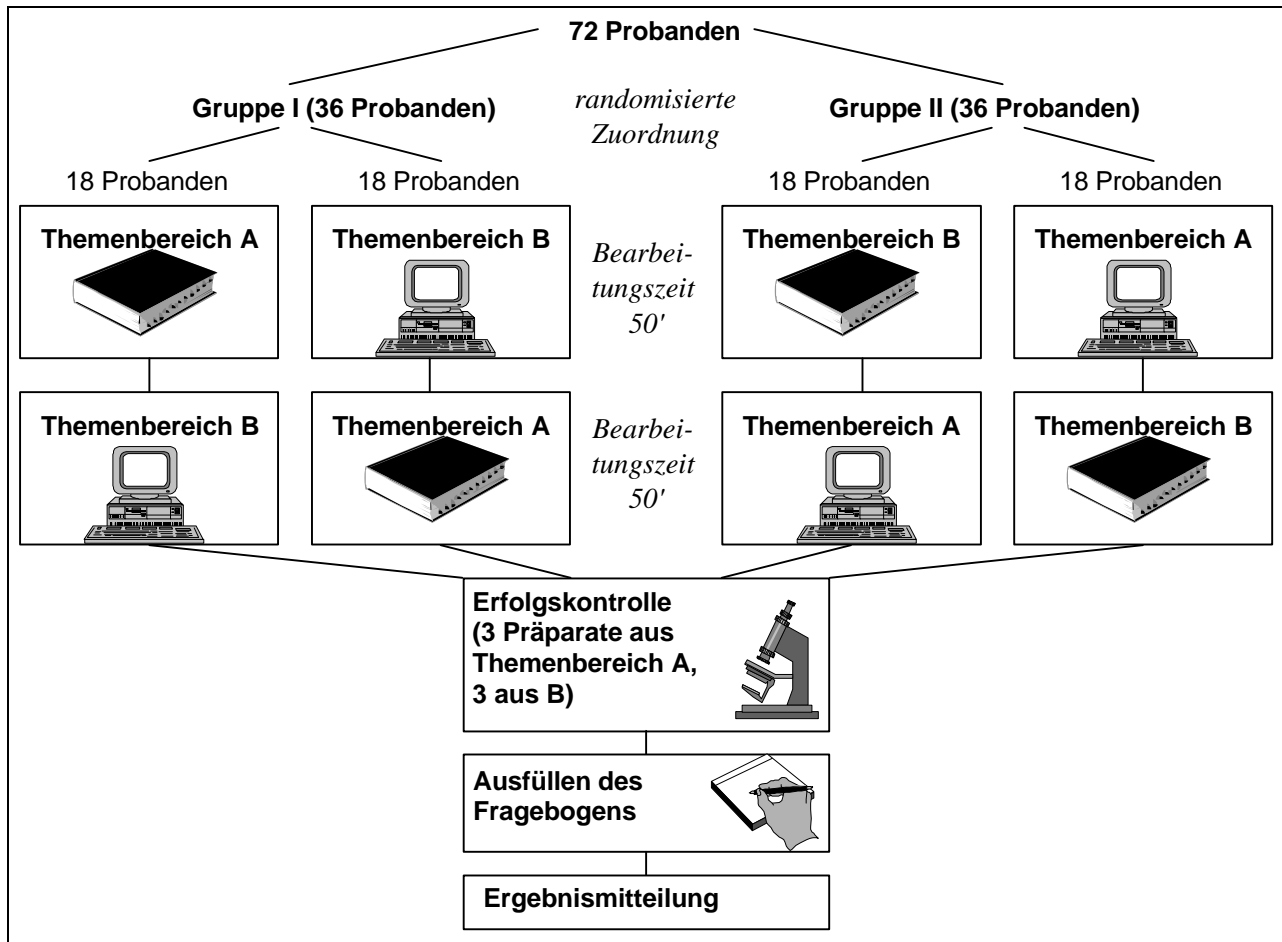


Abbildung 2: Ablauf der Evaluationsstudie des Histopathologie-Programms MicroPat

Die Evaluation wurde mit 72 freiwilligen Probanden aus dem ersten klinischen Studienabschnitt durchgeführt. Während der Studie war nur der Atlasteil des Systems aktiviert, und die Navigation war nur innerhalb desjenigen Themenbereichs möglich, welcher im Sinne dieser Evaluationsstudie bearbeitet werden sollte.

Die Studenten wurden randomisiert zwei Gruppen zugeordnet (Abb. 2). Eine Gruppe sollte den Themenbereich A (Prostatakarzinom, invasiv lobuläres Mammakarzinom, invasiv duktales Mammakarzinom und Komedokarzinom) ohne Computer ausschließlich mit Büchern und den Themenbereich B (Adenomyomatose, Mastopathie, Fibroadenom und Gallertkarzinom) mit Computer und nur fakultativ mit Büchern bearbeiten. Die andere Gruppe verfuhr umgekehrt (Crossover-Versuchsplan [4]). Jeder Student hatte für das Studium mit dem Computer und mit den Büchern nacheinander je 50 Minuten zur Verfügung. 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. Die Diagnose wurde schriftlich festgehalten, wobei jeder Student auf einer dreistufigen Skala (von „Ich bin bei meiner Diagnose ganz sicher.“ = 3 Punkte, bis „Ich habe bei meiner Diagnose geraten.“ = 1

Punkt) die Sicherheit seiner Aussage bewerten sollte. Anschließend wurden die Studenten aufgefordert, einen Bogen mit 17 Fragen über die eigenen Computerkenntnisse, die Beurteilung der Bilder, der Texte und des Programms allgemein auszufüllen. Die Bewertung erfolgte hier auf einer Ordinalskala [6] von eins (= „sehr gut / stimme ich völlig zu“) bis fünf (= „sehr schlecht / stimme ich überhaupt nicht zu“).

Ergebnisse

Eine erste Auswertung zeigte, daß 58.3% der Probanden mit dem Computer bessere Resultate erzielen konnten, 31.9% kamen mit den Büchern besser zurecht und 9.7% der Studenten schnitten in beiden Prüfungsabschnitten gleich gut ab. Nach dem Vorzeichentest bzw. McNemar-Test sind die Ergebnisse auf einem Niveau von $\alpha=0.05$ signifikant. Dieses Ergebnis muß allerdings durch den Befund relativiert werden, daß eine eindeutige Tendenz bestand, den mit dem Computer bearbeiteten Themenbereich überzubewerten. Das heißt, die Studenten fühlten sich „sicherer“ in ihrer Diagnose, der Score – sowohl im Positiven als auch im Negativen – war höher. Wurden bei der Analyse nur Richtig- und Falschdiagnosen berücksichtigt, konnte die Signifikanz trotz eines leichten Vorteils für MicroPat (36.1% gegenüber 25.0%, kein Unterschied 38.9%) nicht aufrecht erhalten werden. Weiterhin stellte sich heraus, daß die Probanden im Themenbereich B wesentlich besser abschnitten als im Themenbereich A (Median +4 vs. 0 Punkte). Der Vorteil für das Lernen mit MicroPat kommt ausschließlich im „leichten“ Themenbereich B zum Tragen, während das Lernen des „schweren“ Themenbereichs A mit MicroPat oder mit Büchern gleiche Erfolge liefert. Eine Korrelation der „Einschätzung der eigenen Computerkenntnisse“ (aus dem Fragebogen) mit der Differenz des Lernerfolgs („Computer-Lernen“ minus „Buch-Lernen“) fand sich nicht (Spearman'scher Korrelationskoeffizient 0.04). Die Auswertung des Fragebogens ergab, daß die Bedienbarkeit und die Schnelligkeit des Programmes durchschnittlich mit der Note 1.5 bewertet wurde. Ebenfalls gut fiel die Benotung der Bilder mit 1.8 aus; mit 2.4 Punkten schnitten die Texte deutlich besser als befriedigend ab. Mit Werten von durchschnittlich 1.6 hinterließ das Programm einen durchweg positiven Gesamteindruck.

Diskussion

Da kontrollierte Evaluationsstudien oft fehlen oder schlecht vergleichbar sind [2], sollte mit dieser Crossover-Versuchsplanung mit anschließender, den realen Verhältnissen entsprechenden Prüfung, eine leicht durchführbare und dennoch aussagekräftige Validierungsmethode entwickelt werden, die als Modell für ein standardisiertes Verfahren dienen kann. Der Crossover-Plan hat einige Vorteile. Jeder Proband ist gleichzeitig Test- und Kontrollperson. So kann die Anzahl der benötigten Teilnehmer verringert werden. Außerdem können einige Effekte, wie z. B. ein Motivationsverlust bei der Kontrollgruppe, vermieden werden, da jeder Proband an allen Elementen der Studie teilnimmt. Der Nachteil eines solchen Verfahrens liegt in Perioden-, Residual- und Behandlungseffekten [4].

Es konnten aus einer Gesamtheit von 237 Studenten 72 gewonnen werden, was unter den statistischen Bedingungen eines nicht normalverteilten Kollektivs bezüglich der erreichten Punktzahl eine eher geringe Zahl ist. Da die Teilnahme freiwillig war, kann die Stichprobe nicht als repräsentativ betrachtet und eine mögliche Überrepräsentation PC-erfahrener Studenten nicht ausgeschlossen werden. Allerdings muß man auch die zeitliche Terminierung (letzte Semesterwoche) als wesentlichen Grund der Nicht-Teilnahme und die Verlosung von Lehrbüchern unter den Probanden als zusätzlichen Anreiz für die Teilnahme in Betracht ziehen. Zudem ist eine Gruppe von 72 Personen aus einer Gesamtheit von 237 so groß, daß damit charakteristische Aussagen über eine bedeutende Untermenge von 30% getroffen werden können.

Die mikroskopischen Präparate waren den Teilnehmern nicht bekannt. Ziel der Studie war deshalb auch, festzustellen, ob mit Hilfe dieses Mediums die Studenten die Fähigkeit des „Muster Erkennens“ am unbekanntem Präparat besser erlernen. Erschwerend war für die noch unerfahrenen Probanden, daß sich Präparate aus den Organsystemen Mamma und Prostata infolge ihrer starken Ähnlichkeit differentialdiagnostisch besonders schwer abgrenzen lassen.

Die Diagnosen wurden nach ihrer Richtigkeit und ihrer Sicherheit der Entscheidung bewertet. Dieses Vorgehen erwies sich als sinnvoll, um Unterschiede in der Qualität der korrekten oder falschen Diagnosen herauszufiltern. Wie sich zeigte, müssen Aussagen auf der Grundlage dieses Scores zurückhaltend bewertet werden, wenn das Abschneiden der beiden bearbeiteten Themenbereiche – wie in unserer Studie – beträchtliche Unterschiede aufweist und die Sicherheit von Aussagen offenbar durch das neue Lernmedium „Computer“ beeinflusst werden. Daß bei der Konzeption von MicroPat weitestgehend Standards berücksichtigt wurden, die in lernergesteuerten Systemen [1] verbreitet sind, so z. B. eine einfache Menüstruktur, die Symbolik der Steuerelemente oder die Navigationsmöglichkeiten, stimmt mit dem Ergebnis überein, daß die unterschiedliche technische Versiertheit der Probanden das Abschneiden beim Lernen mit MicroPat nicht beeinflusst. Allerdings kann durch die Angst vor dem neuen System oder durch die Begeisterung beim ersten Kontakt eine Beeinflussung der Lernleistung nicht ausgeschlossen werden [9].

Eine weitergehende Analyse der Daten [7] sowie zusätzliche Studien zur Messung des Lernerfolgs von CBT-Systemen sind notwendig. Das hier vorgestellte Evaluationsmodell mag hierbei als Anregung dienen. Unabhängig von der didaktischen Beurteilung von Lernsystemen ist jedoch die Tendenz zur CD-ROM-basierten Publikation unverkennbar, können doch mit diesem Medium große Bilddatenmengen in guter Farbqualität und mit einem, im Gegensatz zum Buch, geringen finanziellen Aufwand dargestellt werden [5, 8].

Literatur

1. Bodendorf, F.: Computer in der fachlichen und universitären Ausbildung. Handbuch der Informatik. Bd. 15.1. München: Oldenbourg 1990.
2. Fieschi, D.; Fieschi, M.; Soula, G.; Degoulet, P.: Evaluation of computer-assisted instruction methods. Apropos of 26 comparative studies published between 1989 and 1992. *Pathologie Biologie* 42: 2, 1994, 183-190.
3. Klar, R.; Auhuber, T.; Schulz, S.: MicroPat – An Example for Integrating an Atlas of Digital Images into a Computer Based Training System. In: *Health Telematics Education*. 1997 (in print).
4. Lehmacher, W.: Zwei Stichproben von Verlaufskurven und Crossover-Versuche. *Medizinische Informatik und Statistik* 67. Heidelberg: Springer 1986. 77-85.
5. Leven, F. J.; Schulz, S.; Alle, W.; Klar, R.: Rechnergestützte Lehr- und Lernsysteme in den Klinika: Stand und zukünftige Entwicklungen. In: Buchholz, W.; Haux, R. (Hrsg.). *Informationsverarbeitung in den Universitätsklinika Baden-Württembergs*. Heidelberg 1995, 187-193.
6. Polit-O'Hara, D.: *Nursing research: principles and methods*. 4th ed. Philadelphia; New York; London: J. B. Lippincott 1991, 282-293.
7. Schulz, S.; Auhuber, T.; Klar, R.: Kontrollierte Evaluationsstudie von MicroPat – Interaktiver Atlas der Histopathologie. In: Conradi, H.; Kreutz, R.; Spitzer, K. (Hrsg.): *CBT in der Medizin – Methoden, Techniken, Anwendungen*. Aachen: Verlag der Augustinus Buchhandlung 1997, 117-123.
8. Schulz, S.; Schrader, U.; Klar, R.: Computer-based Training and Electronic Publishing in the Health Sector: Tools and Trends. In: *Methods of Information in Medicine* (in print).
9. Selbmann, H.-K.: Validierung von Software und Teachware. In: Baur, M.; Michaelis, J. (Hrsg.). *Computer in der Ärzteausbildung*. München: Oldenbourg 1990, 165-173.