

Computerdiagnose: Nutzen fragwürdig

Zusammengefasst und kommentiert von Etzel Gysling

a -- Semigran HL, Linder JA, Gidengil C et al. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ 2015 (8. Juli); 351: h3480

[\[LINK\]](#)

Zusammenfassung: Etzel Gysling

Es gibt heute eine ganze Reihe von Online-Programmen, die es ermöglichen sollen, als medizinischer Laie die eigenen Symptome so zu interpretieren, dass eine Vermutungsdiagnose oder wenigstens die Dringlichkeit einer ärztlichen Intervention («Triage») festgelegt werden kann. In der vorliegenden Studie wurden 23 solche Programme («symptom checkers», alle in englischer Sprache) auf ihre diesbezüglichen Qualitäten geprüft. Mit 11 dieser Programme konnten Diagnose und Dringlichkeit ermittelt werden, mit 8 nur die Diagnose und mit 4 nur die Dringlichkeit. 45 mögliche Krankengeschichten wurden nach drei verschiedenen Dringlichkeiten gruppiert: (1) Notfallintervention angezeigt – (2) Nicht-notfallmässige Konsultation sinnvoll – (3) Selbstbehandlung sinnvoll. Die 45 Szenarien wurden mit allen Programmen durchgespielt; dabei wurde überprüft, ob sich die richtige Diagnose zuerst oder mindestens innerhalb der 3 bzw. 20 wahrscheinlichsten Vermutungsdiagnosen ergab und ob eine korrekte Triage gemäss der Dringlichkeit erfolgte.

Gesamthaft identifizierten die Programme die korrekte Diagnose bei durchschnittlich 34% der überprüften Szenarien (95% CI 31-37%); in den ersten drei Vermutungsdiagnosen lag das richtige Resultat bei 51%, in den ersten zwanzig bei 58%. Die Diagnose war am häufigsten (im Mittel bei 40%) richtig, wenn eine nicht-dringliche Situation (Selbstbehandlung sinnvoll) vorlag – bei Notfällen war die Diagnose nur in 24% ganz richtig. Die Dringlichkeit wurde im Ganzen bei 57% aller Fälle korrekt beurteilt, wobei die erste Dringlichkeit (Notfälle) am besten triagierte wurde (durchschnittlich in 80% der Fälle richtig). Viel weniger zuverlässig war die Triage bei den anderen Szenarien (Konsultation sinnvoll: 55% richtig; Selbstbehandlung sinnvoll: 33% richtig). So kommt die Studie zum Schluss, die Computerprogramme hätten erhebliche Defizite, da sowohl die Diagnosen als auch die Dringlichkeits-Beurteilungen häufig falsch sind. Um Risiken zu vermeiden, tendieren die Programme dazu, auch weniger dringliche Situationen in eine höhere Dringlichkeitsstufe zu triagieren.

Wenn Computerprogramme infolge ihrer Risikoscheu tatsächlich vermehrt zu überflüssigen Notfallkonsultationen führen sollten, wäre ihre breite Anwendung keineswegs wünschenswert. Andererseits lassen sich die Programme ja nicht verbieten, so wenig, wie man es verbieten könnte, einfach via Google nach Symptomen zu suchen. Wie die Studienverantwortlichen richtig schreiben, interessiert insbesondere der Vergleich mit einer telefonischen Beratung – bisher gibt es aber nur wenige entsprechende Studien. Ich habe ergänzend ein paar Programme in deutscher Sprache kurz getestet; auch hier waren die Resultate eher enttäuschend. Die wichtigste Schlussfolgerung lautet wohl, diese Programme sollten verbessert und in der einen oder anderen Form validiert werden.