

Aus dem Zentrum für Operative Medizin der Universität zu Köln  
Klinik und Poliklinik für Allgemein-, Viszeral- und Tumorchirurgie und  
Transplantationschirurgie  
Direktorin: Universitätsprofessorin Dr. med. C. Bruns

# **„Symptomchecker“ und studentische „Checks“**

**Ergebnisevaluation von diagnostischen und Triage-  
Assessments medizinischer Diagnoseprogramme und  
Medizinstudierender**

Inaugural-Dissertation zur Erlangung der Doktorwürde  
der Medizinischen Fakultät  
der Universität zu Köln

vorgelegt von  
Ole Humbracht  
aus Münster

promoviert am 22. Februar 2023

Gedruckt mit Genehmigung der Medizinischen Fakultät der Universität zu Köln  
2022

Dekan:                   Universitätsprofessor Prof. Dr. G. R. Fink

1. Gutachter:       Privatdozent Dr. med. R Kleinert

2. Gutachter:       Professor Dr. med. V. R. Burst

#### Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich Unterstützungsleistungen von folgenden Personen erhalten

Herr PD Dr. med. Robert Kleinert

Herr Dr. med. Rabi Raj Datta

Herr Dr. phil. Thomas Dratsch

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

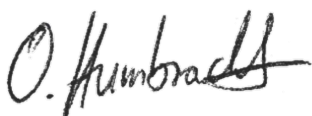
Die dieser Arbeit zugrunde liegenden Experimente sind von mir mit Unterstützung von Herrn Felix Friedemann Brohl durchgeführt worden.

#### Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Köln, den 05.07.2022

Unterschrift:



## Danksagung

Herrn Priv.-Doz. Dr. Robert Kleinert danke ich für die Überlassung des Themas und für die Unterstützung während der gesamten Arbeit.

Herrn Dr. Rabi Raj Datta danke ich für die hervorragende Betreuung und Unterstützung bei der Durchführung der gesamten Arbeit.

Mein Dank gilt allen Probandinnen und Probanden, die durch ihre jeweilige Teilnahme diese Studie ermöglicht haben.

Auch möchte ich allen Mitarbeitenden der Allgemein-, Viszeral-, Tumor- und Transplantationschirurgie der Uniklinik Köln danken, die dazu beigetragen haben, dass ich dieses Thema mit Freude bearbeiten konnte.

Dem KISS danke ich für die zur Verfügung gestellten Räumlichkeiten und für die Unterstützung während der Datenerhebung.

Mein besonderer Dank gilt Felix Friedemann Brohl, der mir mit Rat und Tat während des gesamten Projekts zur Seite stand.

Meinen Eltern, meinen Geschwistern und Dorothea möchte ich für ihre bedingungslose Unterstützung und Ermutigung sowohl während des Studiums als auch darüber hinaus danken.

# Inhaltsverzeichnis

<b>ABKÜRZUNGSVERZEICHNIS</b>	<b>7</b>
<b>1. ZUSAMMENFASSUNG</b>	<b>9</b>
<b>2. EINLEITUNG</b>	<b>11</b>
2.1. Historische Aspekte medizinischer Diagnostik	11
2.2. Der diagnostische Prozess	12
2.3. Bedeutung medizinischer Diagnosen	14
2.4. Bedeutung der Triage	15
2.5. Arzt-Patienten-Beziehung	16
2.6. Der „neue“ Patient	18
2.7. Symptomchecker – moderne diagnostische Tools	19
2.8. Fragestellungen und Ziel der Arbeit	21
<b>3. MATERIAL UND METHODEN</b>	<b>23</b>
3.1. Studiendesign	23
3.2. Stichprobengewinnung	23
3.2.1. Aufnahme der Probanden	23
3.2.2. Einschlusskriterien	23
3.2.3. Ausschlusskriterien	23
3.3. Material	24
3.3.1. Suchstrategie für die Symptomchecker	24
3.3.2. Einschlusskriterien für die Symptomchecker	24
3.3.3. Ausschlusskriterien für die Symptomchecker	24
3.3.4. Verwendete Symptomchecker	25
3.3.5. Provenienz der klinischen Fallvignetten	27
3.3.6. Verwendete klinische Fallvignetten	28
3.4. Methoden	32
3.4.1. Operationalisierung der Zieldimensionen	32

3.4.2.	Versuchsaufbau der Studierendenuntersuchung	33
3.4.3.	Versuchsaufbau der Symptomcheckeruntersuchung	36
3.4.4.	Dokumentations- und Auswertungsmethoden	36
3.4.5.	Statistische Methoden	37
<b>4.</b>	<b>ERGEBNISSE</b>	<b>38</b>
4.1.	Symptomchecker-Assessments	38
4.1.1.	Diagnosequalität der Symptomchecker-Assessments	38
4.1.2.	Triagequalität der Symptomchecker-Assessments	39
4.1.3.	Differenzielle Ergebnisse der Symptomchecker-Assessments	39
4.2.	Studierenden-Assessments	42
4.2.1.	Demographische und Fachsemesterdaten der Studierenden	42
4.2.2.	Diagnosequalität der Studierenden-Assessments	43
4.2.3.	Triagequalität der Studierenden-Assessments	44
4.2.4.	Studierenden-Assessments und Studienfortschritt	44
4.2.5.	Studierenden-Assessments und Symptompräsentation	46
4.2.6.	Studierenden-Assessments und subjektive Bewertungsdimensionen	48
4.3.	Vergleich Symptomchecker- und Studierenden-Assessments	50
4.3.1.	Deutschsprachige Symptomchecker- und Studierenden-Assessments	51
4.3.2.	Englischsprachige Symptomchecker- und Studierenden-Assessments	52
4.4.	Zusammenfassung der Ergebnisse	54
<b>5.</b>	<b>DISKUSSION</b>	<b>56</b>
5.1.	Wissenschaftskontext der Studie	56
5.2.	Diskussion der Ergebnisse	57
5.2.1.	Die diagnostische Leistungsfähigkeit der Symptomchecker	57
5.2.2.	Die präklinische Triagefähigkeit der Symptomchecker	60
5.2.3.	Leistungsfähigkeit der Studierenden	64
5.2.4.	Leistungsvergleich der Symptomchecker und der Studierenden	66
5.2.5.	Methodische Limitationen und Aussagekraft der Studie	67
5.2.6.	Resümee und Ausblick	68
<b>6.</b>	<b>LITERATURVERZEICHNIS</b>	<b>71</b>
<b>7.</b>	<b>ANHANG</b>	<b>79</b>

<b>7.1.</b>	<b>Abbildungsverzeichnis</b>	<b>79</b>
<b>7.2.</b>	<b>Tabellenverzeichnis</b>	<b>79</b>
<b>7.3.</b>	<b>Standardisierte klinische Fallvignetten auf Deutsch in ihrer vollständigen Form</b>	<b>81</b>
<b>7.4.</b>	<b>Einverständniserklärung</b>	<b>90</b>

## Abkürzungsverzeichnis

ÄApprO	Approbationsordnung für Ärzte
AG	Arbeitsgemeinschaft
ALT	Alanin-Aminotransferase
App	Applikation
ATS	<i>Australasian Triage Scale</i>
BMI	<i>Body Mass Index</i>
BP	Blutdruck
COPD	Chronisch obstruktive Lungenerkrankung
CRP	C-reaktives Protein
CTAS	<i>Canadian Triage and Acuity Scale</i>
EC	<i>Emergent care</i>
eGK	Elektronische Gesundheitskarte
et al.	<i>et alii/et aliae</i>
FDA	<i>Food and Drug Administration</i>
ft3	freies Trijodthyronin
ft4	freies Thyroxin
GAS	Gruppe A Streptokokken
G-BA	Gemeinsamer Bundesausschuss
GEE	<i>Generalized estimating equations</i>
HINTS	<i>Health Information National Trends Survey</i>
ICD	<i>International Statistical Classification of Diseases and Related Health Problems</i>
INR	<i>International normalized ratio</i>
iOS	Betriebssystem für Smartphones des Herstellers Apple Inc.
KI	Konfidenzintervall
m	männlich
M	Mittelwert
MTS	<i>Manchester Triage Scale</i>
NE	<i>Non-emergent care</i>
py	<i>pack years</i>
SC	<i>Self care</i>
SD	Standardabweichung/ <i>standard deviation</i>
Sek.	Sekunde
T	Temperatur
TSH	Thyroidea-stimulierendes Hormon



USA	Vereinigte Staaten von Amerika
w	weiblich
WHO	Weltgesundheitsorganisation
ZNA	Zentrale Notaufnahme

# 1. Zusammenfassung

**Hintergrund:** Im digitalen Zeitalter suchen immer mehr Patienten<sup>a</sup> medizinische Informationen im Internet. Neu sind in diesem Rahmen so genannte Symptomchecker, die nach der Angabe von Symptomen Verdachtsdiagnosen stellen und Dringlichkeitseinschätzungen abgeben. Sie richten sich an Patienten und übernehmen somit ärztliche Aufgaben. Durch ihre Funktionen ergeben sich mögliche Vorteile, wie das Einsparen von Ressourcen bei nicht behandlungsbedürftigen Beschwerden oder eine telemedizinische Versorgung in strukturschwachen Regionen. Fehleinschätzungen und Fehldiagnosen durch die Symptomchecker könnten jedoch auch zu einer Überbelastung von Notaufnahmen führen und bergen Risiken, wie das Nichterkennen ernster Krankheitszustände und Notfälle.

**Zielsetzung:** Ziel der vorliegenden Studie war die Evaluation der Leistungsfähigkeit der Symptomchecker bezüglich Diagnose und Triage, um in der Folge ihren potenziellen Nutzen einschätzen und ihn mit der Fähigkeit von menschlichem Personal vergleichen zu können.

**Methode:** Diese monozentrisch prospektive Studie wurde im Rahmen der AG Lehre der Abteilung Allgemein-, Viszeral- und Tumorchirurgie des Universitätsklinikums Köln durchgeführt. Im Rahmen dieser Untersuchung wurden 143 Medizinstudierende des klinischen Studienabschnitts sowie 16 Symptomchecker getestet. Deren Leistungsfähigkeit in Diagnosestellung und Triage wurde anhand standardisierter allgemeinmedizinischer Fallvignetten (n=49) unterschiedlicher Dringlichkeit ermittelt. Dazu wurden diese in die Benutzeroberflächen der Symptomchecker eingegeben. Im Falle der Studierenden wurde die Bearbeitung der Vignetten in Textform und mit Hilfe eines Chat-gestützten Anamnesegesprächs via der Nachrichtenfunktion des Messengers Skype durchgeführt.

**Ergebnisse:** Die Symptomchecker nannten in durchschnittlich 38.02% der Fallbearbeitungen die richtige Diagnose an erster Stelle und in 56.44% innerhalb der ersten drei Verdachtsdiagnosen. Im Gegensatz dazu waren es bei den Studierenden (in Textfallbearbeitung) 64.10% präzise Diagnosen und 80.19% noch korrekte Differentialdiagnosen. Neben dieser signifikant besseren Leistung in der Diagnosestellung übertrafen die Studierenden die Symptomchecker auch in der Triage, wobei hier der Unterschied nur bei den Fällen höchster Dringlichkeit signifikant war. In der Untersuchung

---

<sup>a</sup> *Disclaimer:* Aus Gründen der besseren Lesbarkeit und eines höheren Textverständnisses wird an dieser Stelle und in der gesamten Dissertationsschrift auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet. Sämtliche Formulierungen und Bezeichnungen beziehen sich dabei gleichermaßen auf alle Geschlechter.

zeigte sich eine hohe Varianz der verschiedenen Symptomchecker, sodass einige Angebote in ihrer Leistung den Studierenden nahekamen.

**Schlussfolgerung:** Symptomchecker erreichen zu diesem Zeitpunkt noch nicht die Leistung von Menschen mit medizinischem Basiswissen in Bezug auf Diagnosestellung und Triage, sodass nur einzelne Angebote potenzielle Vorteile in der Gesundheitsversorgung bringen könnten. Die Entwicklung der Symptomchecker sollte in der Zukunft durch interdisziplinäre Forschung kritisch begleitet werden.

## 2. Einleitung

Diagnosen sind in der Medizin allgegenwärtig, sie nehmen unterschiedliche Funktionen wahr und existieren aus multiplen Gründen. Sie implizieren ätiologische Vorstellungen und eröffnen präventives Handeln. Auch ermöglichen sie sowohl die Identifikation von Behandlungsoptionen als auch die Prognosestellung und erschaffen einen Kommunikationsrahmen.<sup>1</sup> Mittels Diagnosen können Krankheiten systematisiert werden (Nosologie), zugleich geben sie die Voraussetzung für die Triage ab, also die Zuordnung einer Dringlichkeit und eines Behandlungssettings.

### 2.1. Historische Aspekte medizinischer Diagnostik

In der Medizin bezeichnet die Diagnose die Feststellung einer Erkrankung, wodurch sie einen integralen Bestandteil des ärztlichen Handelns bildet. Das Wort leitet sich ursprünglich vom altgriechischen διάγνωσις, *diagnosis* ab. Dieses bedeutet „Unterscheidung“ oder „Entscheidung“ und setzt sich zusammen aus *diá-* („durch“) und *-gnósis* („Erkenntnis“ oder „Urteil“).<sup>2</sup> Als begriffliche Identifizierung eines Krankheitszustands setzt eine Diagnose die sorgfältige Untersuchung der Symptomatik und die Kenntnis der Vorgeschichte voraus.<sup>3</sup> Im medizinischen Kontext stellt die Diagnose damit die Entscheidung dar, mit der der Arzt das weitere Vorgehen bestimmt. Sie bildet das Bindeglied zwischen der Untersuchung und der Therapie.

Das Konzept der Diagnose ist heute in der Medizin selbstverständlich, hat sich jedoch über eine lange Zeit entwickelt. In der Zeit der griechischen Antike und damit auch der Zeit des Hippokrates war das medizinische Wissen und seine Terminologie noch sehr begrenzt. So fand ein erzählender, beschreibender Stil Anwendung, um ein Bild von Krankheit zu umreißen, für den heute ein einzelner Begriff ausreichend ist.<sup>4</sup>

Der Impuls, die Medizin im Sinne der heutigen Klassifikation zu ordnen, ging Mitte des 17ten Jahrhunderts von zwei Engländern aus, zum einen von John Graunt (1620–1674), zum anderen von Thomas Sydenham (1624–1689), der von vielen als Vater der klassifikatorischen Medizin bezeichnet wird. Graunt beschäftigte sich mit Eigenschaften von Todesstatistiken und forschte zudem an einer Klassifikation von Krankheiten.<sup>5</sup> Thomas Sydenhams Ansatz hingegen beruhte auf seinen Beobachtungen am Krankenbett, anhand derer er diagnostisch-kurativ forschte.<sup>6</sup> Seine empirische Arbeitsweise lehnte er an die Methodik der Botanik an und sah einen Schlüsselbaustein des ärztlichen Handelns in der Differenzierung von Krankheitsbildern.<sup>7</sup>

Aus den Arbeiten von Graunt und Sydenham resultieren die zwei zentralen Bedeutungen diagnostischen Handelns in der modernen Medizin: Zum einen sollen Diagnosen die Grundlage bilden für die statistische Dokumentation, um epidemiologische und sozialmedizinische Ansätze zu ermöglichen. Zum anderen bilden sie im klinischen Feld die

Voraussetzung für die systematische Beschreibung und Einteilung der Krankheiten (Klassifikation und Nosologie)<sup>5</sup>. In der ICD, der *international classification of diseases*, werden diese beiden Ambitionen diagnostischen Handelns realisiert.

Die ICD ist die am weitesten verbreitete medizinische Klassifikation und findet inzwischen im gesamten Gesundheitswesen Anwendung, das heißt nicht nur im klinischen Alltag, sondern auch in Bereichen wie etwa der Kostenerstattung, der Verwaltung, der Epidemiologie und Forschung.<sup>8</sup> Die Existenz einer internationalen diagnostischen Klassifikation belegt die zentrale Bedeutung der Diagnosestellung in der modernen Medizin, die aufgrund ihrer weitreichenden sozialrechtlichen Relevanz exklusiv approbierten Ärzten und Ärztinnen vorbehalten ist, welche durch ihre lange Ausbildung zur Diagnosestellung befähigt werden sollen<sup>9-11</sup>.

## 2.2. Der diagnostische Prozess

Der Diagnoseprozess oder auch die Diagnosefindung beschreibt den Weg hin zur Diagnose, also jene Methode, welche dem Arzt zur Beurteilung der Beschwerden und Symptome seines Patienten dient.<sup>1</sup> Zu diesem Zweck erfolgt bei initialer Vorstellung des Patienten eine primäre Sammlung an Informationen. Dazu helfen unter anderem das Anamnesegespräch und die klinische Untersuchung, durch die eine Reihe von Symptomen und Hinweisen zusammengetragen werden können.<sup>12</sup>

Auf Basis dieser ersten Statuserhebung werden sodann Hypothesen generiert. Es werden vorläufige Verdachtsdiagnosen formuliert, die im Folgenden durch weitere Informationssammlung, wie zum Beispiel mittels apparativer und Labordiagnostik, auf ihre Plausibilität hin geprüft werden müssen. Abschließend werden die hypothetisch formulierten Verdachtsdiagnosen evaluiert und der Prozess gegebenenfalls neu durchlaufen, bevor es zur endgültigen Formulierung einer Diagnose kommt.<sup>12</sup>

Dieser scheinbar fast lineare Diagnoseprozess, wie die *National Academies of Sciences* ihn beschreiben, stellt aber nur eine grobe Annäherung dar. Denn für das diagnostische Handeln bedarf es neben der puren Informationsgewinnung des „*clinical reasoning*“, also des klinischen Vorwissens und klinischen Denkens<sup>12,13</sup>, um Arbeitshypothesen zu entwickeln, die den Blick für subtile Krankheitsphänomene öffnen.

Eine allgemeingültige Theorie des komplexen diagnostischen Prozesses existiert nicht, sodass auch die Literatur zu diesem Thema geradezu fragmentiert ist.<sup>13</sup> Allerdings hat das Modell des *Dual Processings* in der Medizin breite Anerkennung gefunden<sup>14,15</sup>. Als anwendbare Annäherung an den Denkprozess des Arztes verbindet es Erkenntnisse aus der Grundlagenforschung zur menschlichen Entscheidungsfindung mit erfahrungswissenschaftlichen Befunden zum ärztlichen diagnostischen Handeln. Die Grundannahme dieses *Dual Processings* geht vom Zusammenspiel zweier Denksysteme

aus, bezeichnet als System 1 und System 2. Davon beinhaltet das System 1 intuitive und das System 2 analytische Denkweisen beziehungsweise kognitive Prozeduren.<sup>15</sup>

Wenn sich etwa ein Patient mit einer einfach ersichtlichen Krankheit mit eindeutiger Anamnese und Klinik präsentiert, finden die Denkweisen des System 1 Anwendung. So bereiten beispielsweise Erkrankungen mit typischen Symptomen, wie die Hauterscheinungen beim Herpes Zoster oder die typischen Beschwerden einer Sinusitis oder einer Harnwegsinfektion, den meisten Ärzten keine Schwierigkeiten bei der Diagnosestellung. Es handelt sich um „Blickdiagnosen“, auf die sich der Arzt mühelos festlegen kann. Die Erkennung pathognomonischer Symptome und typischer Symptomkonstellationen erfolgt dabei intuitiv, durch das Erkennen eines Musters oder durch eine einfache Verknüpfung.<sup>16</sup> Das System 1 benötigt wenig Aufmerksamkeit und Anstrengung, läuft schnell ab und ist stark automatisiert. Seine Denkweisen beruhen auch auf Erfahrung, Stereotypen, Redundanzermittlung und anderen mentalen Hilfestellungen.<sup>14</sup>

Deshalb wird auf dieses System-1-Denken besonders auch unter widrigen Umständen, wie Zeitdruck, Überforderung oder defizitärer Informationslage, zurückgegriffen. So kann eine den Umständen entsprechend beste Entscheidung getroffen werden.<sup>16</sup> Das System 1 ist jedoch anfällig für Fehler und falsche Schlussfolgerungen<sup>14</sup>, intuitive Urteilsbildungen stehen unter starkem Einfluss äußerer und innerer Faktoren, (z.B. Arbeitsumfeld, Charakter des Patienten, affektiver Zustand des Arztes u.a.).<sup>16,17</sup>

Darum kommt dem System 2 eine wichtige Aufgabe zu. Es ist die Denkweise, die angewendet wird, sobald die Symptome und Befunde des Patienten nicht unmittelbar zu einer offensichtlichen Entscheidung führen oder kognitive Dissonanzen auftreten. Dies kann zum Beispiel der Fall sein, wenn die Informationen keine Blickdiagnose zulassen, die Symptomkonstellationen unvollständig sind oder sich mehrere Differentialdiagnosen anbieten. Das System 2 umfasst analytische Denkweisen, wie die hypothetisch-deduktive Methode, Ausschlussverfahren oder die Verwendung von Entscheidungsbäumen.<sup>16,18</sup> Das Denken im System 2 dauert damit länger, erfordert eine hohe Aufmerksamkeit und wird bewusst vom Arzt gesteuert.<sup>14</sup>

Nach Croskerry laufen die beiden Systeme, das intuitive und das analytische, nicht parallel zueinander ab, sondern es wird in der Entscheidungsfindung ein schneller Wechsel zwischen ihnen vollzogen. So können Intuitionen und rationale Gegenargumente am Ende zu einer wohl austarierten Entscheidung führen (siehe Abbildung 1). Der Großteil des Denkprozesses entfällt dabei insgesamt auf das intuitive System 1, da es für den Arzt weniger kognitiven Aufwand bedeutet und damit effizienter ist.<sup>14</sup>

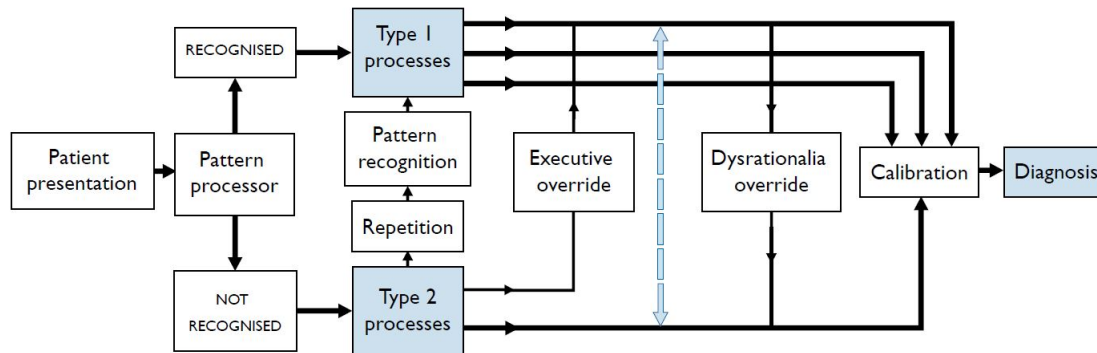


Abbildung 1. Schema des Dual Processing nach P. Croskerry und Nimmo

Dargestellt ist ein diagnostisches Schema basierend auf dem Modell des *Dual Processings* nach P. Croskerry und Nimmo<sup>14</sup>. Die Bezeichnungen Type 1 und Type 2 stehen in diesem Schema für die Systeme 1 und 2 in der diagnostischen Entscheidungsfindung. Die blau gestrichelte Querverbindung illustriert den stetigen Wechsel zwischen den beiden Pfaden.

Demnach entsteht eine Diagnose nicht aus einer reinen Abwägung von Wahrscheinlichkeiten aus einem endlosen Katalog von Krankheiten, sondern durch die Interaktion teils bewusster und teils vorbewusster kognitiver Prozesse.

### 2.3. Bedeutung medizinischer Diagnosen

Der Diagnoseprozess charakterisiert die offensichtlichsste Funktion der Diagnosen. Ihre Wirkung geht jedoch über die Festlegung einer Krankheit hinaus und tangiert die Rolle des Arztes, des Patienten und ihr Zusammenspiel.

Die Diagnose besitzt die implizite Funktion, der Medizin Professionalität zu verleihen, ihre gesellschaftliche Anerkennung zu stützen und die verschiedenen Fachbereiche der Medizin zu differenzieren.<sup>1,5,19</sup> Im klinischen Kontext wird durch die Verwendung von Diagnosen beispielsweise entschieden, welchem Fachbereich der entsprechende Patient zugeteilt wird, welche Therapie er benötigt und wie viele Ressourcen für ihn aufgewendet werden.<sup>5</sup>

Die sozialrechtliche Bedeutung medizinischer Diagnosen erwächst aus dem Umstand, dass erst die Feststellung der Diagnose den Krankheitsstatus inklusive etwa der Arbeitsunfähigkeit herbeiführt. Was vorher die persönlichen Beschwerden des Patienten waren, wird so zu einer anerkannten Krankheit.<sup>1</sup> Die bisherige Ungewissheit über seinen Zustand wird, zum Guten oder Schlechten, durch die Diagnose ersetzt.<sup>20</sup> So werden sowohl seine Eigen- als auch die Fremdwahrnehmung verändert. Für den Soziologen Talcott Parsons nimmt der Patient fortan eine besondere soziale Rolle ein. Zu dieser gehören die Freistellung von sozialen Obligationen, die Freistellung von der Eigenverantwortlichkeit für

seinen Krankheitszustand, das Ansehen als vorübergehend krankes Mitglied der Gesellschaft und die Berechtigung zur sozialrechtlichen Unterstützung.<sup>21</sup>

Der umfassende Einfluss, den die Diagnose in der Medizin einnimmt, legt nahe, dass Fehlentscheidungen eine ebenso große Reichweite haben. Dies gilt für die Außenwirkung der Profession, das Erleben des Patienten und besonders für seine medizinische Versorgung.

## **2.4. Bedeutung der Triage**

Unter Triage versteht man die Priorisierung etwa von Verletzungen und Erkrankungen nach ihrer Behandlungsdringlichkeit. Etymologisch lässt sich der Begriff auf das französische Verb „trier“ = „sortieren“ zurückführen und entstammt der Militärmedizin der napoleonischen Ära.<sup>22</sup> Die Herkunft des Begriffs und des Konzepts der Triage lässt sich auch heute noch erkennen. Die Literatur zum Thema ist vermehrt im Kontext von Katastrophen- und Notfallmedizin, häufig in Verbindung mit dem medizinischen Management von Terroranschlägen, zu finden.<sup>22–25</sup>

So beschreibt das Triagieren den Vorgang, zwischen harmlos und gefährlich, zwischen schwer Verwundeten und leicht Verwundeten zu unterscheiden und eine Zuordnung zu den medizinischen Fachdisziplinen vorzunehmen.<sup>5</sup> Diese Differenzierung der Triage erfolgt anhand mehrerer Faktoren, zu denen die Ressourcenverfügbarkeit, das Patientenaufkommen, die Schwere des Patientenleidens sowie die Dringlichkeit seiner Versorgung gezählt werden.<sup>26</sup>

Überträgt man die Triage aus einem Katastrophenkontext in einen allgemeinen medizinischen Anwendungsbereich, sind diese Faktoren unterschiedlich zu gewichten. In der westlichen Welt ist der Faktor der Ressourcenverfügbarkeit zu vernachlässigen, da außerhalb einer Katastrophensituation oder einer schweren Epidemie für jeden individuellen Patienten davon auszugehen ist, dass er nach den Regeln der ärztlichen Kunst behandelt werden kann. Ein zu hohes Patientenaufkommen ist nur in Ausnahmefällen, wie zum Beispiel durch eine übermäßige Inanspruchnahme von ärztlichen Notdiensten und Ambulanzen vorstellbar, wodurch sich eine Versorgung verzögern könnte. Die Faktoren der Schwere des Patientenleidens und der Dringlichkeit sind hingegen auch in der alltäglichen medizinischen Versorgung stets für eine Triage von Bedeutung. So ändert auch eine ideal organisierte Notaufnahme nichts an der Dringlichkeit und Therapie, mit der beispielweise ein Koronarinfarkt versorgt werden muss.

Zum jetzigen Zeitpunkt existiert eine strukturierte Triage in Deutschland vornehmlich in den zentralen Notaufnahmen (ZNA) der Kliniken. Grundsätzlich geht es hier besonders um die Festlegung der Behandlungsdringlichkeit.<sup>24</sup>

Eine wachsende Verbreitung der organisierten Triage in den deutschsprachigen ZNA ist aktuell absehbar. In der Mitte des Jahres 2018 trat die „Regelung [...] zu einem gestuften



System von Notfallstrukturen in Krankenhäusern [...]“ des Gemeinsamen Bundesausschusses (G-BA) in Kraft.<sup>27</sup> In dieser heißt es als Regelung für alle Krankenhäuser der Basisnotfallversorgung im §12: „Es kommt ein strukturiertes und validiertes System zur Behandlungspriorisierung bei der Erstaufnahme von Notfallpatienten zur Anwendung. Alle Notfallpatienten des Krankenhauses erhalten spätestens zehn Minuten nach Eintreffen in der Notaufnahme eine Einschätzung der Behandlungspriorität“.

Zur Anwendung kommen spezialisierte klinische Triage-Systeme, wie die *Australasian* [sic!] *Triage Scale* (ATS), die *Canadian Triage and Acuity Scale* (CTAS) oder die in Deutschland bekannte *Manchester Triage Scale* (MTS). Diese Systeme kategorisieren in meist fünf Dringlichkeitsstufen. Zusätzlich geben sie maximal anzustrebende Wartezeiten an.<sup>24</sup>

Somasundaram et al. (2009) unterscheiden zwischen klinischer und präklinischer Triage: Als Kernziel der klinischen Triage wird die sichere und schnelle Identifizierung der Lage genannt, da das Einschätzen des Krankheitszustands eines Patienten große Bedeutung für seine medizinische Versorgung haben kann.<sup>24</sup> Die präklinische Triage hat demgegenüber zwei Aufgaben. Sie soll den richtigen Versorgungsbedarf im Sinne einer passenden medizinischen Einrichtung einschätzen und die zeitliche Dringlichkeit erkennen helfen. Die präklinische Triage ist bislang kaum in den Fokus der Versorgungsforschung und der wissenschaftlichen Analyse gelangt. Erste Ansätze findet man in Form der Telefontriage und im weiteren Rahmen der Telemedizin.<sup>28,29</sup> Sowohl in der Telefontriage als auch in der Telemedizin ist die zuvor beschriebene Exklusivität der ärztlichen Einschätzung gebrochen. Personal mit medizinischem Vorwissen, wie Pflegekräfte, medizinische Fachangestellte oder Studierende der Humanmedizin können für die Durchführung von Ersteinschätzungen in Betracht gezogen werden.<sup>30-32</sup> Diese Ersteinschätzungen entsprechen lediglich Verdachtsdiagnosen, die als Grundlage für vorläufige Dringlichkeitseinstufungen im Sinne einer Triage dienen. Dabei ist hervorzuheben, dass weder die Verdachtsdiagnosen noch die Dringlichkeitseinstufungen die Verbindlichkeit oder die sozialrechtliche Relevanz besitzen, die ärztliche Diagnosen und Triageempfehlungen innehaben.

## **2.5. Arzt-Patienten-Beziehung**

Für den Patienten sind die Diagnose und die Triage zwei wichtige Bestandteile seiner Versorgung. Sein Zugang zu dieser Versorgung erfolgt dabei durch seinen Kontakt zu einem Arzt. Hierbei wird das Verhältnis zwischen dem behandelnden Arzt und dem Patienten als Arzt-Patienten-Beziehung bezeichnet. Zu diesem Verhältnis gehören emotionale, kommunikative und organisatorische Aspekte, welche dem gesellschaftlichen Wandel unterliegen. Beide Seiten, Arzt und Patient, haben häufig ein unterschiedliches Krankheitsverständnis, verschiedene Kausalitätserklärungen und Wertvorstellungen. Damit sind Diagnosen Gegenstand von Diskurs und Interaktion.<sup>1</sup>

Gesellschaftliche Veränderungen gehen ein in die Kommunikation und Interaktion von Arzt und Patient. Insbesondere die Arzt-Patienten-Beziehung hat sich in den letzten Dekaden deutlich gewandelt.

Besonders markant war der Abschied vom klassischen paternalistischen Modell der Arzt-Patienten-Beziehung in der jüngsten Vergangenheit. Es basierte auf der Annahme einer bevormundenden Arztrolle, in der der Arzt die Entscheidungen im besten Interesse des Patienten zu treffen hatte. Begründet war diese Beziehung in der Vorstellung einer fürsorglichen Aufgabe des Arztes<sup>33</sup>. Der Patient nahm in diesem Modell eine passive, geradezu infantile Rolle ein.

Im starken Kontrast zum Paternalismus steht das Modell des Arztes als Dienstleister. In diesem Modell agiert der Arzt passiv und reagiert auf die Erwartungen des Patienten. Der Patient befindet sich in der Rolle eines Konsumenten, wobei das gegenseitige Vertrauen zwischen Arzt und Patient einer Art Vertragsverhältnis weicht<sup>34,35</sup>.

Das aktuell angestrebte Modell sieht ein partizipatives Verhältnis des Patienten zu seinem Arzt vor. Es basiert auf einer umfassenden Information des Patienten und resultiert in einem sogenannten *shared decision making*, also einer partizipativen Entscheidungsfindung<sup>34,36,37</sup>. Der Arzt hat dabei weder die Funktion eines Dienstleisters inne, noch trifft er Entscheidungen ohne Mitbestimmung des Patienten. Die Abwendung von einer ärztlichen Autokratie reduziert paternalistische Einstellungsmuster, sie lässt Raum für einen Zuwachs an Souveränität auf Seiten des Patienten.<sup>36,38</sup> Unterstützt wird diese Entwicklung durch gesamtgesellschaftliche Trends, die nicht nur den Gesundheitsbereich (Fitness- und Ernährungsbewegungen, Wellness und Body Positivity u.a.) betreffen, sondern auch den politischen und zivilgesellschaftlichen Raum insgesamt (Bürgerbewegungen, partizipatorische Volksentscheide u.a.).

Dem Patienten kommt nun die Aufgabe zu, sich mit einzubringen und zu engagieren. Diese neue Rolle wird allgemein befürwortet und geht Hand in Hand mit der Entwicklung einer größeren Beteiligung des Patienten in allen Fragen des diagnostischen und therapeutischen Handelns<sup>36</sup>. So wird das Einbeziehen des Patienten international als wichtiges Instrument zur verbesserten Bereitstellung von Gesundheitsdiensten angesehen<sup>39</sup>.

Nimmt der Patient seine Rolle in einer Arzt-Patienten-Beziehung auf Augenhöhe an und versucht den damit einhergehenden Erwartungen gerecht zu werden, entsteht für ihn die Notwendigkeit, sich umfassend zu informieren. Nur so kann er das Kompetenzgefälle zwischen sich und dem Arzt vermindern, dessen Expertise hinterfragen und mitentscheiden. Der nötige Zugang zu den medizinischen Informationen hat sich im Laufe der Zeit parallel zur neuen Arzt-Patienten-Beziehung weiterentwickelt.

## 2.6. Der „neue“ Patient

Die Medizin als Fach und Berufsstand hatte in der Vergangenheit die Deutungshoheit über Gesundheit und Krankheit inne. Meinungen außerhalb der Ärzteschaft beeinflussten das gesellschaftliche Verständnis von Krankheit kaum. So lag die organisatorische Entscheidungsmacht zum größten Teil in der Hand der praktizierenden Medizinerschaft selbst.<sup>36</sup>

In den 1970er Jahren begann langsam die Veränderung dieser Verhältnisse. Damals hatte der durchschnittliche Patient allerdings nur wenige Optionen jenseits der Informationsvermittlung durch den Arzt. Medizinisches Wissen war nur in speziellen Beiträgen in Zeitschriften oder Zeitungen, in Büchern für Laien oder durch Kontakte in Selbsthilfegruppen zu finden. Seitdem hat sich das Angebot jedoch stets vergrößert. Spätestens durch das Aufkommen des Internets ist die Informationsbreite auf eine beträchtliche Größe angewachsen<sup>40</sup>. Dadurch haben heutzutage immer mehr Menschen einen Zugriff auf Gesundheitsinformationen. Der Paradigmenwechsel im Zugang zu medizinischen Informationen liegt damit in der Digitalisierung begründet.

Bereits Anfang der 2000er Jahre war es Patienten möglich, mit Hilfe des Internets Informationen zu ihren Beschwerden und deren korrekter Behandlung zu finden, noch bevor sie in direkten Kontakt zu ihrem Arzt traten<sup>38</sup>. Auch Nettleton und Burrows (2003) sahen die Exklusivität des medizinischen Wissens schon zu diesem Zeitpunkt als aufgehoben an. Für sie ergab sich die Vermutung, dass sich Laienwissen mit bestehenden medizinischen Konzepten zu einem neuen medizinischen Wissensstand vermischen könnten<sup>41</sup>.

Seither hat sich das Internet als primäre Informationsquelle des alltäglichen Lebens etabliert. Suchmaschinen wie Google oder Yahoo haben zur Nutzerfreundlichkeit beigetragen und ermöglichen eine schnelle und genaue Suche im Netz. Inzwischen gibt es mehrere Studien, die untersucht haben, wie häufig Patienten das Internet, inklusive Suchmaschinen, verwenden, um Informationen zu Gesundheitsthemen zu finden. Aufgrund der schnellen technischen Weiterentwicklung und Verbreitung des Internets sollten diese Zahlen aber ausschließlich als Orientierung in einem fortschreitenden Prozess verstanden werden.

Im viel zitierten Report „Health Online 2013“ wurde resümiert, dass 72% der erwachsenen Internetnutzenden in den USA nach medizinischen Informationen suchen. Davon gaben 77% an, eine Suchmaschine wie Google, Bing oder andere zu nutzen<sup>42</sup>. Eine Analyse der *Health Information National Trends Survey* (HINTS) ergab, dass das Internet mit mehr als 40% auf Platz 1 der meistgenannten Quellen bei der jeweilig letzten vergangenen Suche der Teilnehmer nach Gesundheitsinformationen stand.<sup>43</sup> Im klinischen Alltag können Suchmaschinen bei komplizierten Krankheitsbildern unter der Voraussetzung passender Suchbegriffe helfen, seltene Differentialdiagnosen zu identifizieren.<sup>44</sup>

Im Verlauf der technischen Entwicklung ist generell davon auszugehen, dass sich die Suchaktivität im Gesundheitsbereich zu großen Teilen vom Computer auf mobile Endgeräte, allen voran das Smartphone, verlagern wird. Dadurch könnte sich folglich auch das Suchverhalten der Patientenschaft verändern.<sup>45</sup>

Besonders die verstärkte Zuwendung der jüngeren Generation hin zu Online-Informationsquellen und mobilen Endgeräten stellt eine Herausforderung für die Medizin dar. Denn diese Generation formt natürlicherweise einen wachsenden Anteil des Patientenkollektivs der Zukunft.

Gerade die Generation Y wird stark mit der Digitalisierung assoziiert. Zu der Generation Y werden allgemein die Geburtenjahrgänge von ungefähr 1980 bis 2000 gezählt.<sup>46,47</sup> Damit folgt sie auf die Generation der Babyboomer und die Generation X und wuchs somit in der Zeit der Globalisierung und der Anfänge des Internets auf.<sup>46</sup> Der Begriff „Generation Y“ beschreibt also eine Generation, die mit einem neuen Verhältnis zur digitalen Welt assoziiert wird. Da die Angehörigen dieser Generation bereits ihr gesamtes Leben mit neuen digitalen Technologien in Kontakt sind, werden sie auch als *Digital Natives* bezeichnet.<sup>48</sup> Ihr Status als digital kompetente Kohorte bestätigt sich in mehreren Beobachtungen. Dem Statistischen Bundesamt nach ist die Häufigkeit der Internetnutzung in den Altersgruppen von 16-24 und 25-44 Jahren am höchsten.<sup>49</sup> Auch in Bezug auf mobile Technologien, wie das Smartphone, zeigt sich eine erhöhte Nutzung gegenüber älteren Personen.<sup>50,51</sup>

Junge Personen nutzen zudem auch explizit häufiger digitale Gesundheitsangebote. Grundsätzlich ist das digitale Suchen nach Gesundheitsinformationen in allen Altersgruppen präsent, unter jüngeren Patienten aber besonders verbreitet.<sup>42,45</sup> In Deutschland wird am häufigsten in der Altersgruppe zwischen 25 und 44 Jahren online nach Informationen zu Gesundheitsthemen gesucht.<sup>49</sup> In einer US-amerikanischen Untersuchung gaben 84% der befragten Jugendlichen zwischen 13 und 18 Jahren an, bereits online Gesundheitsinformationen abgerufen zu haben.<sup>52</sup>

Die heranwachsende Generation von Patienten ist dauerhaft digital verbunden. Ihre Nutzung von Smartphones führt zur Erwartung der Möglichkeit einer schnellen und unmittelbaren Informationssuche.<sup>53,54</sup> Das breite Vorhandensein von Gesundheitsinformationen im Internet könnte so zu einer neuen Anspruchshaltung führen. Diese sofortige Verfügbarkeit von Informationen und Antworten auch zu komplexen Gesundheitsfragen im Netz wird einen „neuen“ Patienten konstituieren mit zahlreichen Konsequenzen für die Balance im Arzt-Patienten-Verhältnis.

## **2.7. Symptomchecker – moderne diagnostische Tools**

Als besonders komplex unter den diversen Gesundheitsfragen ist die Frage nach der eigenen Diagnose hervorzuheben. Ein Patient empfindet durch seine Krankheit einen besonderen Leidensdruck. Unbekannte Symptome oder die Verschlechterung der eigenen

körperlichen Beschwerden führen zu seiner Verunsicherung, und eine Diagnose funktioniert in dieser Situation als Hilfe, um das empfundene Risiko einschätzen zu können.

Um die Konsultation des Arztes aus zeitlichen Gründen zu umgehen oder eine Zweitmeinung einzuholen, können Patienten heute auch durch die zahlreichen Internetquellen nach Diagnosen und Risikoeinschätzungen suchen.<sup>55</sup> Dabei ist die Suche nach punktuellen Informationen im Netz einfach, die Diagnosefindung hingegen komplex. Sie erfordert die Synopsis mehrerer Informationen.

Um diese Aufgabe zu vereinfachen, sind seit einiger Zeit spezialisierte Angebote verfügbar. Diese lassen sich als „Symptomchecker“ bezeichnen. Zu finden sind sie in Form von Applikationen für Smartphones, kurz Apps, wie beispielsweise „Ada – Deine Gesundheitshelferin“ und „Doctor Diagnose“ oder in Form von Webseiten, die über einen Webbrowser zu erreichen sind.<sup>56</sup>

Das Ziel dieser Angebote ist es meist, aus der Angabe von Symptomen eine Reihe von Verdachtsdiagnosen zu generieren. Mit dieser Funktion ersetzen sie eine Kernkompetenz ärztlichen Handelns.<sup>57</sup> Darum stehen sie unter besonderer Beobachtung des professionellen Diskurses.

In den USA sind die Symptomchecker von der *Food and Drug Administration* (FDA) je nach ihren individuellen Funktionen in den meisten Fällen als Medizinprodukte klassifiziert.<sup>57</sup> Damit fallen sie auch unter bestimmte Qualitätsstandards für solche Produkte, sind aber zum gegebenen Zeitpunkt von einer Zulassungskontrolle ausgenommen.<sup>58,59</sup> In Deutschland hingegen existieren bisher noch keine allgemeinen Regulationen der Symptomchecker.

Jedes durch derartige Diagnosehilfen erworbene Vorwissen der Patienten nimmt Einfluss auf das Arztgespräch und kann die Arzt-Patienten-Beziehung nachhaltig beeinflussen. Daher sollten mögliche Marktinteressen der Anbieter beachtet werden, etwa die indirekte Förderung spezifischer Medizinprodukte. Allerdings haben sich medizinische Apps zumindest bis vor einiger Zeit größtenteils vor kritischen Nachfragen bezüglich solcher Interessenkonflikte fernhalten können. Es ist bisher nicht geklärt, ob medizinische Apps Probleme der Medizin zu lösen helfen oder sie selbst verursachen.<sup>60</sup> Es gibt jedoch diverse vorstellbare Vor- und Nachteile, beispielsweise in der Notfallversorgung oder Telemedizin.

Neben der Erstellung einer Verdachtsdiagnose geben Symptomchecker häufig eine Risikoeinschätzung, beziehungsweise Triage ab.<sup>56</sup> Zum einen wird eine geeignete medizinische Anlaufstelle, wie beispielsweise ein medizinischer Erstversorger, ein Facharzt, eine Klinik, die Notaufnahme oder die telefonische Alarmierung des Rettungsdienstes empfohlen. Zum anderen werden zeitliche Orientierungen für die Dringlichkeit einer Behandlung gegeben. Diese reichen von der Empfehlung einer unmittelbaren, sofortigen Konsultation bis hin zu einem Abwarten über Tage oder Wochen hinweg. Eine Ausnahme bei den Empfehlungen bildet die Selbsthilfe. Bei diesem Ergebnis der Triage wird der Patient

dazu aufgefordert, seine Beschwerden mit den ihm zur Verfügung stehenden Mitteln selbst zu lindern. Auch Empfehlungen zum reinen Beobachten des Zustandes fallen in diese Kategorie.

Die Diagnosestellung und die Handlungsempfehlung folgen einem Algorithmus, in dem mehrere Faktoren abgebildet werden, darunter Demografie, Risikofaktoren und Vorerkrankungen, besonders jedoch die Symptomkonstellation.<sup>1</sup>

Ein einzelnes Symptom ließe eine nicht einzugrenzende Zahl an möglichen Differentialdiagnosen zu. Demografische Daten, wie Alter und Geschlecht, ermöglichen vor allem eine Wahrscheinlichkeitseinschätzung altersspezifischer Erkrankungen oder den Ausschluss geschlechtsgebundener Erkrankungen (z.B. Adnexitis). Zum Ausschluss von Differentialdiagnosen ist deshalb die Betrachtung mehrerer Symptome und gegebenenfalls ihrer Ausprägung von Nöten.

Demnach werden Symptomchecker wie folgt definiert:

Symptomchecker sind digitale Angebote, die Patienten bei der Selbstdiagnose helfen. Sie ermöglichen die Zuordnung von Symptomkonstellationen zu Differentialdiagnosen und/oder geben eine Einschätzung zur Notwendigkeit einer ärztlichen Konsultation.

Die aktuelle Forschung mit direktem Bezug auf die Leistung der verfügbaren Symptomchecker ist überschaubar. Einige Untersuchungen haben sich bisher mit dem potenziellen Nutzen und Schaden der Symptomchecker sowie der medizinsoziologischen Bedeutung in der Medizin auseinandergesetzt.<sup>55,60,61</sup> Forschung bezüglich der Qualität der Angebote kommt zu diesem Zeitpunkt vorrangig von der Arbeitsgruppe um Hannah L. Semigran et. al. von der Harvard Medical School in Boston, Vereinigte Staaten.<sup>56</sup>

## **2.8. Fragestellungen und Ziel der Arbeit**

Ziel der Studie ist es, den potenziellen Nutzen der Symptomchecker zu evaluieren. Dazu soll die Qualität der Symptomchecker auf dem deutsch- und englischsprachigen Markt ermittelt werden. Zur Qualitätsbeurteilung der Symptomchecker wird im Einzelnen folgenden Fragestellungen nachgegangen:

1. Wie hoch ist die Treffergenauigkeit im Sinne des Präzisionsgrades korrekter Diagnosen?
2. Mit welcher Trefferwahrscheinlichkeit wird die korrekte Diagnose zumindest in den Differentialdiagnosen überhaupt berücksichtigt?
3. Wie angemessen sind die Handlungsempfehlungen, das heißt, wie sieht die Triagequalität der Symptomchecker aus?

In einem zweiten Schritt sollen die diagnostische und Triage-Qualität von Personen mit medizinischem Vorwissen anhand von studentischen Assessments (Einschätzungen) ermittelt werden.

Dazu stellen sich gleichartige Fragen nach:

1. dem Präzisionsgrad der gestellten Diagnosen (Treffergenauigkeit)
2. der Berücksichtigung von relevanten Diagnosen (Trefferwahrscheinlichkeit)
3. der Triagequalität

Zusätzlich stellt sich bei diesen studentischen Assessments die Frage, ob ein Zuwachs an klinischem Wissen die Qualität der Assessments erhöht.

Darüber hinaus interessieren auch Probleme der Symptompräsentation, das heißt inwiefern die Studierenden mit einer explorativen Symptomsuche zurechtkommen, wenn sie nicht auf eine im Klartext vorgegebene Symptomatik zurückgreifen können.

Die Leistungsfähigkeit der Symptomchecker soll schließlich auch mit der Leistung der studentischen Assessments verglichen werden.

### **3. Material und Methoden**

#### **3.1. Studiendesign**

Es handelt sich um eine monozentrische, prospektive Studie, in der die diagnostischen und Triage-Assessments der auf dem Markt verfügbaren, kostenlosen und allgemeinmedizinischen Symptomchecker sowie von Studierenden der Medizinischen Fakultät der Universität zu Köln evaluiert wurden.

#### **3.2. Stichprobengewinnung**

##### **3.2.1. Aufnahme der Probanden**

Die Probanden wurden im Rahmen des ersten und zweiten Teils des Blockpraktikums des Faches Chirurgie im Sommersemester 2018 (April 2018 - Juli 2018) am Universitätsklinikum der Universität zu Köln rekrutiert. Weitere Medizinstudierende im klinischen Abschnitt der ärztlichen Ausbildung konnten in diesem Zeitraum über persönliche Ansprache und E-Mail-Verteiler erreicht werden.

Voraussetzung für die Teilnahme an der Studie war die Abgabe einer ausgefüllten Einverständniserklärung, mit der einer Erhebung von personenbezogenen Daten zugestimmt wurde. Die Weitergabe, Speicherung und Auswertung dieser Daten erfolgten anonymisiert und unter Einhaltung gesetzlicher Bestimmungen. Die Einverständniserklärung enthielt detaillierte Informationen über Ziel und Verlauf der Studie und klärte über die Möglichkeit auf, die Teilnahme an der Studie vorzeitig zu beenden (Eine Kopie der Einverständniserklärung findet sich im Anhang).

##### **3.2.2. Einschlusskriterien**

Zum Zeitpunkt der Studie mussten die Probanden das achtzehnte Lebensjahr vollendet haben und im klinischen Abschnitt des Medizinstudiums studieren. Einschlusskriterium war somit, dass jeder Proband den ersten Teil der ärztlichen Prüfung bestanden und gleichzeitig den zweiten Teil der ärztlichen Prüfung gemäß der Approbationsordnung für Ärzte (ÄApprO) noch nicht absolviert hatte.

##### **3.2.3. Ausschlusskriterien**

Bei nicht vollständig ausgefüllter Einverständniserklärung oder nicht vollständig ausgefülltem Fragebogen sowie beim frühzeitigen Abbruch der Teilnahme erfolgte der Ausschluss des jeweiligen Probanden.



### **3.3. Material**

#### **3.3.1. Suchstrategie für die Symptomchecker**

Für die Erfassung der englischsprachigen Symptomchecker wurden die Suchergebnisse der Studie von Semigran et al. übernommen.<sup>56</sup> Die Suchparameter für deutschsprachige Angebote wurden zudem in Anlehnung an diese Studie gewählt.

Die deutschsprachige Suche fand im September 2018 statt. Verwendete Quellen waren die Internetsuchmaschine Google sowie zwei Marktplätze für Smartphone-Apps, der Appstore für iOS-Geräte sowie der Google-Playstore für Android-Geräte.

In der Suche wurden die Begriffe „medical diagnosis“ und „symptom checker“ aus der Studie von Semigran et al. verwendet. Um einen zweiten deutschsprachigen Suchbegriff neben „symptom checker“ zu etablieren, wurde zusätzlich nach dem deutschsprachigen Begriff „Krankheitsfinder“ gesucht.

Die Google-Suche wurde im privaten Modus des Browsers Firefox in der zum damaligen Zeitpunkt (September 2018) aktuellen Firefox Version 62 durchgeführt. Der Standort wurde auf Deutschland eingestellt und ausschließlich Seiten auf Deutsch angefordert. Berücksichtigt wurden, wenn vorhanden, die ersten 300 Suchergebnisse, da in vorangegangenen Untersuchungen eine signifikante Abnahme der Suchergebnisrelevanz ab dieser Anzahl an Ergebnissen festgestellt wurde.<sup>62</sup> Gleichartig wurde auf den deutschsprachigen App-Marktplätzen verfahren.

#### **3.3.2. Einschlusskriterien für die Symptomchecker**

Die Symptomchecker mussten auf Deutsch verfügbar sein oder zu den englischsprachigen Symptomcheckern zählen, die in der Studie von Semigran et al. getestet wurden<sup>56</sup>. Es wurden ausschließlich Symptomchecker berücksichtigt, die kostenfrei zur Verfügung standen, öffentlich zugänglich und für die Diagnosefindung bei Menschen vorgesehen waren. Angebote aus dem Bereich der Veterinärmedizin wurden nicht eingeschlossen. Zudem war es Voraussetzung, dass die Symptomchecker ein allgemeinmedizinisches Spektrum an Erkrankungen und ihren Symptomen abdeckten und sich nicht auf einen Fachbereich oder Teilbereich beschränkten.

#### **3.3.3. Ausschlusskriterien für die Symptomchecker**

Ein Ausschluss von Symptomcheckern aus der Studie von Semigran et al. fand bei fehlender Funktionsfähigkeit, fehlender Online-Erreichbarkeit, Vorhandensein einer deutschsprachigen Version oder fehlender Übereinstimmung mit den Einschlusskriterien statt.

Ein Ausschluss von Symptomcheckern aus der deutschsprachigen Suche fand statt, sofern die Funktion der Symptomchecker auf einer Stichwortsuche basierte oder das Suchergebnis nicht zu einem deutschsprachigen Angebot führte.

### 3.3.4. Verwendete Symptomchecker

Die Google-Suche zeigte für den Suchbegriff „Krankheitsfinder“ 100 Suchergebnisse. Für den Suchbegriff „symptom checker“ ergab die Suche 120 Suchergebnisse. Für den Suchbegriff „*medical diagnosis*“ ergab die Suche 171 Suchergebnisse. Die Appstore-Suche ergab für den Suchbegriff „Krankheitsfinder“ 0 Suchergebnisse. Für den Suchbegriff „symptom checker“ ergab die Suche 79 Suchergebnisse. Für den Suchbegriff „*medical diagnosis*“ ergab die Suche eine Anzahl von 240 Suchergebnissen. Die Playstore-Suche ergab für alle drei Suchbegriffe „Krankheitsfinder“, „symptom checker“ und „*medical diagnosis*“ eine Anzahl von 240 Suchergebnissen. Die Suchergebnisse wurden händisch mit den Einschlusskriterien abgeglichen und potenzielle Symptomchecker dokumentiert. Eine Auflistung der verwendeten Symptomchecker findet sich in der Tabelle 1.

Tabelle 1. Liste der untersuchten Symptomchecker mit Kurzbeschreibung

Symptomchecker	Beschreibung	Triageformulierungen/-besonderheiten
deutschsprachig		
Ada	Mobile App für iOS und Android; fragt Alter, Geschlecht und Risikofaktoren ab; Anmeldung notwendig <a href="https://ada.com/de/">https://ada.com/de/</a>	Bei Warnsymptomen direkter Verweis auf den Rettungsdienst; benötigt sofortige medizinische Notfallversorgung; ein Arztbesuch innerhalb der nächsten 2-3 Tage zu empfehlen; kann normalerweise selbst behandelt werden
Diagnosefinder.apotheken.de	Simple Webseite, auf der sich über Körperregionen und Symptome zu Symptomkonstellationen durchgeklickt wird; nicht mehrere Symptome wählbar; Geschlecht wählbar; keine Kinder <a href="https://diagnosefinder.apotheken.de/">https://diagnosefinder.apotheken.de/</a>	Diverse Formulierungen zur Triage; es muss am Ende abhängig von den vorhandenen Symptomen entschieden werden
Netdokter - Symptomchecker	Grobe Alterseinteilung in Kind, Jugendlichen und Erwachsenen; fragt nach Geschlecht; erst Auswahl des Hauptsymptoms, dann zahlreiche Fragen bis zur Diagnose <a href="https://www.netdokter.de/symptom-checker/">https://www.netdokter.de/symptom-checker/</a>	Triage ausschließlich in höchste Dringlichkeitsstufe (EC) bei Warnsymptomen
Symptomate	Frägt nach Alter, Geschlecht und Risikofaktoren. Patientenalter nur ab 18 Jahren auswählbar – keine Kinder <a href="https://symptomate.com/diagnosis/de/">https://symptomate.com/diagnosis/de/</a>	Rufen Sie sofort den Rettungsdienst; Suchen Sie Ihren Arzt in den nächsten 24h auf; Die Symptome sprechen eher nicht für eine schwerwiegende Erkrankung

AskMD Sharecare	Anmeldung erforderlich; nach Angabe des Hauptsymptoms folgt eine Reihe von Fragen mit Antworten zum Ankreuzen; fragt nach Alter, Geschlecht, Maßen und teils Risikofaktoren <a href="https://www.sharecare.com/askmd/get-started">https://www.sharecare.com/askmd/get-started</a>	(Gibt keine Triage)
Doctor Diagnose Android	Mobile App für Android; Alter und Geschlecht je nach Leitsymptom fest zugeordnet aber nicht gefragt; nicht immer Triage gegeben	Consult doctor immediately; seek out medical assistance immediately; call a doctor immediately; consult a doctor; if the pain lasts [...] consult a doctor; keine Triage zur Selbstbehandlung
Drugs.com	Simple Webseite, auf der sich über Körperregionen und Symptome hin zu einer Diagnose geklickt wird; Frage nach Mann, Frau oder Kind vorhanden; nur teilweise Diagnosestellung; Kooperation mit der Harvard Medical School <a href="https://www.drugs.com/symptom-checker/">https://www.drugs.com/symptom-checker/</a>	Diverse Formulierungen zur Triage
Esagil	Keine Fragen nach Alter und Geschlecht; Angabe von bis zu acht Symptomen möglich <a href="http://esagil.com/">http://esagil.com/</a>	(Gibt keine Triage)
Familydoctor.org	Nach Auswahl eines groben Hauptsymptoms werden einige Fragen gestellt; nicht immer wird eine Verdachtsdiagnose gestellt <a href="https://familydoctor.org/your-health-resources/health-tools/symptom-checker/">https://familydoctor.org/your-health-resources/health-tools/symptom-checker/</a>	Diverse Formulierungen zur Triage
Isabelhealthcare	Anmeldung erforderlich; Fragen nach Alter (in Abschnitten), Geschlecht und Standort; Angabe von mehreren Symptomen; Nachfragen nur vor der Triage; seit dem Zeitpunkt der Untersuchung ist eine neue Version erschienen <a href="https://symptomchecker.isabelhealthcare.com/">https://symptomchecker.isabelhealthcare.com/</a>	Triage nach einem Farbtacho - dabei immer Empfehlung zur Konsultation eines Arztes oder Notfallversorgung
Mayoclinic	Unterscheidung nur zwischen Kind und Erwachsenem; keine Geschlechtsfrage;	(Gibt keine Triage)

	nach Auswahl eines Hauptsymptoms können zusätzliche Symptome/Beschreibungen ausgewählt werden <a href="https://www.mayoclinic.org/symptom-checker/select-symptom/itt-20009075">https://www.mayoclinic.org/symptom-checker/select-symptom/itt-20009075</a>	
Medoctor	Fragen nach Geschlecht und Alter ab 12 Jahren; nach Auswahl eines Körperteils werden zahlreiche Fragen gestellt; teilweise doppelte Diagnosen gestellt <a href="https://www.medoctor.com/">https://www.medoctor.com/</a>	(Gibt keine Triage)
Myhealth-alberta – Healthwise	Simple Webseite, auf der sich über Körperregionen und Symptome hin zu einer Triage geklickt wird; keine Verdachtsdiagnosen; Frage nach Mann, Frau oder Kind (nur ausgewählte Symptome) vorhanden <a href="https://myhealth.alberta.ca/Health/pages/symptom-checker.aspx">https://myhealth.alberta.ca/Health/pages/symptom-checker.aspx</a>	You may need care right away – seek care in the next hour; call 911 now emergency care; seek care today; make an appointment; try home treatment
Symcat	Lässt Eingabe mehrerer Symptome zu; fragt anschließend nach Geschlecht, Alter, Ethnie und Risikofaktoren; gibt Verdachtsdiagnosen mit Prozenten für Wahrscheinlichkeit an <a href="http://www.symcat.com/">http://www.symcat.com/</a>	Emergency room; Primary care; Retail clinic; Urgent care
Symptify	Nach Eingabe eines Hauptsymptoms werden mehrere Fragen beantwortet; Frage nach Alter und Geschlecht <a href="https://www.symptify.com/">https://www.symptify.com/</a>	Go to the emergency room; should be seen by a physician; should be seen by a health professional; try home treatment and remedies
Symptoms.webmd	Symptome werden per Suche oder aus Vorschlägen gewählt; Frage nach Alter und Geschlecht <a href="https://symptoms.webmd.com/default.htm#/info">https://symptoms.webmd.com/default.htm#/info</a>	(Gibt keine Triage)

### 3.3.5. Provenienz der klinischen Fallvignetten

In dieser Studie wurden sogenannte Fallvignetten verwendet, die einer Studie zu Symptomcheckern von Semigran et al. entstammen und mit Genehmigung der Autoren für diese Studie weiterverwendet wurden. Dadurch konnte eine bessere Vergleichbarkeit erreicht werden. Es erfolgte eine Übersetzung ins Deutsche durch einen englischsprachigen Muttersprachler. Zusätzlich zu den 45 klinischen Fallvignetten aus der Studie von Semigran

et al. wurden vier weitere Fallvignetten kreiert, die im deutschsprachigen Raum typische Krankheitsbilder abbilden. Diese orientieren sich in Umfang, Form und Stil an den übernommenen 45 Fallvignetten.

### 3.3.6. Verwendete klinische Fallvignetten

Die verwendeten Fallvignetten beinhalten Kurzformen klinischer Fälle inklusive Alter und Geschlecht eines fiktiven Patienten. Zu jeder Fallvignette gibt es auch eine Kurzfassung des entsprechenden Krankheitsbildes. Zudem umfassen die Vignetten teilweise kurze Zusatzinformationen, die über die Symptomkonstellation hinausgehen. In ihrer Länge gestalten sie sich heterogen. Sie unterscheiden sich zudem in ihrer medizinischen Dringlichkeit. Dadurch lassen sie verschiedene Triagierungen zu.

Um die Triagequalität in der Untersuchung bestimmen zu können, ließen sich die Fallvignetten definierten Dringlichkeitsstufen zuordnen. Dies geschah in Anlehnung an die Studie von Semigran et al., in welcher drei Stufen etabliert wurden.<sup>56</sup>

Die in der Studie von Semigran et al. verwendeten Fallvignetten wurden so gewählt, dass aus Sicht der Autoren jeder Dringlichkeitsstufe 15 Fallvignetten zugeordnet werden konnten.<sup>56</sup> Zur vereinfachten Arbeit mit den Fallvignetten wurde jeder ein Kürzel zugewiesen. Dieses bestand aus der Dringlichkeitsstufe, und einer Zahl, die fortlaufend den Fallvignetten zugeordnet wurde. Die vier speziell für diese Untersuchung kreierten Fallvignetten sind mit den Kürzeln EC16, EC 17, NE16 und SC16 kodiert. Zudem enthielt jede Fallvignette eine Kurzform ihres eigentlichen Texts, um den Versuchsleitern einen schnellen Überblick über die Symptomatik zu ermöglichen (siehe Tabelle 2).

Tabelle 2. Standardisierte klinische Fallvignetten in ihrer Kurzfassung

Kürzel	Diagnose	Vereinfachte Fallinformationen
<b>Hohe Dringlichkeit (<i>emergent care</i> = EC)</b>		
EC1	<b>Akutes Leberversagen</b>	48 Jahre, w, Verwirrung, Orientierungslosigkeit, zunehmend schläfrig, Sklerenikterus, leichte Bauchdeckenabwehrspannung im oberen rechten Quadranten, chronische Einnahme von Tylenol/Paracetamol - hat kürzlich mehr eingenommen
EC2	<b>Appendizitis</b>	12 Jahre, w, plötzlich einsetzende starke Bauchschmerzen, Übelkeit, Erbrechen, Durchfall, T = 40 ° C
EC3	<b>Asthma</b>	27 Jahre, w, Asthma in der Vorgeschichte, leichte Atemnot, Keuchen, 3 Tage Husten, Symptome sprechen nicht auf Inhalatoren an, kürzlich zurückliegende Erkältung
EC4	<b>COPD-Anfall</b>	67 Jahre, w, COPD, seit 3 Tagen zunehmende Kurzatmigkeit,

	<b>(schwer)</b>	vermehrtes Husten, grünes Sputum, leichtes Fieber, vermehrter Einsatz von Notfall-Bronchodilatoren
EC5	<b>Tiefe Venenthrombose</b>	65 Jahre, w, seit 5 Tagen Schwellung, Schmerzen im rechten Bein, kürzlicher Krankenhausaufenthalt, Bein schmerzhaft, empfindlich, geschwollen, rot
EC6	<b>Herzinfarkt</b>	64 Jahre, m, seit 1 Tag Schmerzen in der Brust (8/10), nicht-ausstrahlender, substernaler Brustdruck, Schwitzen, Kurzatmigkeit, Engegefühl in der Brust
EC7	<b>Hämolytisch-urämisches Syndrom</b>	4 Jahre, m, seit 7 Tagen, Bauchschmerzen, blutiger Durchfall, aß Hamburger auf Kirmes 3 Tage vor Auftreten der Symptome
EC8	<b>Nierensteine</b>	45 Jahre, m, seit 1 Stunde starker linksseitiger Flankenschmerz, der in die Leistengegend ausstrahlt, Übelkeit, Erbrechen, Schmerzen, die durch Positionswechsel nicht gelindert werden
EC9	<b>Malaria</b>	28 Jahre, m, seit 5 Tagen Schüttelfrost, Rigor, Durchfall, kürzliche Reise ins Ausland (Malariagebiet), von Moskitos gebissen, nahm Malaria-Prophylaxe nicht konsequent
EC10	<b>Meningitis</b>	18 Jahre, m, seit 3 Tagen starke Kopfschmerzen, Fieber, Photophobie, Nackensteifigkeit
EC11	<b>Pneumonie</b>	65 Jahre, m, Bluthochdruck und degenerative Gelenkerkrankungen, seit 3 Tagen produktiver Husten und Fieber (38,3° C)
EC12	<b>Lungenembolie</b>	65 Jahre, m, Kurzatmigkeit über 30 Minuten, Brustschmerz, der sich mit Inspiration verschlechtert, kürzlich zurückliegende Operation mit Bettruhe, Schwellung in der linken Wade, die schmerzempfindlich ist, Fieber
EC13	<b>Rocky-Mountain-Fleckfieber</b>	8 Jahre, m, Fieber, Schüttelfrost, Gelenkschmerzen, Kopfschmerzen, Hautausschläge an Handgelenken und Knöcheln
EC14	<b>Schlaganfall</b>	70 Jahre, m, Übelkeit, Erbrechen, rechtsseitige Schwäche, Armlähmung, Schwierigkeiten beim Sprechen und Verstehen von Sprache
EC15	<b>Tetanus</b>	65 Jahre, m, kann den Mund nicht öffnen, Kontraktion der Muskeln verursacht schmerzhafte Krämpfe für Minuten, Schwitzen, Tachykardie, Schnitt in die Hand während der Gartenarbeit, nicht gegen Tetanus geimpft
EC16	<b>Sigma-divertikulitis</b>	85 Jahre, w, akute Schmerzen im linken Unterbauch, Verstopfung, Übelkeit, Erbrechen, T = 38,5 ° C, CRP erhöht, Leukozyten erhöht
EC17	<b>Ileus</b>	50 Jahre, m, starke Bauchschmerzen, schwallartiges Erbrechen, abdominelle Abwehrspannung, klingende, gesteigerte

Darmgeräusche, CRP erhöht, Leukozyten erhöht

Mittlere Dringlichkeit ( <i>non-emergent care</i> = NE)		
NE1	<b>Akute Otitis media</b>	18 Monate, w, seit 1 Woche Rhinorrhoe, Husten, Verstopfung der Nase, Reizbarkeit, Mangel an Appetit, Fieber, besucht Kindertagesstätte
NE2	<b>Akute Pharyngitis</b>	7 Jahre, w, Fieber (38,5 ° C), Übelkeit, Erbrechen, Halsschmerzen, geschwollene Lymphknoten, Tonsillenexsudat; kein Husten, keine Rhinorrhoe und keine verstopfte Nase
NE3	<b>Akute Pharyngitis</b>	24 Jahre, m, Halsschmerzen, Fieber (39 ° C), Kopfschmerzen, kein Husten, Tonsillenexsudate
NE4	<b>Akute Sinusitis</b>	35 Jahre, w, seit 15 Tagen verstopfte Nase, Gesichtsschmerzen, grüner Nasenausfluss, kein Fieber
NE5	<b>Rücken-schmerzen</b>	35 Jahre, m, Rückenschmerzen nach Schneeschaukeln, Fußheberschwäche links, Symptome seit 3 Wochen, Gefühlsverlust im Fuß
NE6	<b>Zellulitis</b>	45 Jahre, m, Schmerzen und Rötung der Haut, leichtes Fieber, Ödeme und Druckschmerz am Unterschenkel
NE7	<b>COPD-Anfall (mild)</b>	56 Jahre, w, Raucherin, Kurzatmigkeit und Husten seit mehreren Tagen, Rhinorrhoe vor 3 Tagen aufgetreten, weißer Auswurf, kein Schüttelfrost
NE8	<b>Influenza</b>	30 Jahre, w, seit 2 Tagen Fieber, Husten, Kopfschmerzen, Schwäche, bekam keine Gripeschutzimpfung
NE9	<b>Pfeiffer-Drüsenfieber</b>	16 Jahre, w, seit 1 Woche Fieber, Halsschmerzen, Müdigkeit, Schluckbeschwerden, vergrößerte Lymphknoten, Exsudate, makulärer Hautausschlag an Rumpf und Armen
NE10	<b>Magengeschwür</b>	40 Jahre, m, seit 2 Monaten intermittierende Oberbauchschmerzen, stumpfer, nagender Schmerz, wacht nachts auf, Schmerz bessert sich durch Nahrung/Milch/Ranitidin, ähnliche Episode vor 5 Jahren
NE11	<b>Pneumonie</b>	6 Jahre, m, Asthma, seit 5 Tagen Husten, Fieber, Appetit gut, gelbes Sputum, Temperatur 38,7 ° C
NE12	<b>Salmonellen</b>	14 Jahre, m, Übelkeit, Erbrechen, nicht blutiger Durchfall, leichte Bauchkrämpfe (T = 37,8), leichte Bauchschmerzen, Durchfall nach einem Picknick, rohes Hühnchen gegessen
NE13	<b>Gürtelrose</b>	77 Jahre, m, seit 5 Tagen Brennen und Schmerzen auf der rechten Seite der Brust, Erythem, makulopapulöser Ausschlag, Kopfschmerzen, Unwohlsein, Hautausschlag der sich innerhalb 3 bis 5 Tagen zu klaren Vesikeln entwickelt
NE14	<b>Harnwegsinfekt</b>	26 Jahre, w, schmerzhaftes Wasserlassen, dringendes Wasserlassen, häufigeres Wasserlassen seit 2 Tagen, sexuell aktiv;

		kein Fieber, Schüttelfrost, Übelkeit, Erbrechen, Rückenschmerzen, Ausfluss, vaginaler Pruritus
NE15	<b>Schwindel</b>	65 Jahre, w, Schwindel, plötzlicher Beginn, rezidivierend, dauert <30 Sek., anhaltender Trigger, kein Hörverlust, Ohrgeräusche, Muskelschwäche, Gefühlsverlust
NE16	<b>Morbus Basedow</b>	40 Jahre, w, Hitzewallungen, diffuser Haarausfall, innere Unruhe, Zittern, häufiges Schwitzen, vergrößerte Schilddrüse, Ruhepuls 110 Schläge pro Minute, bilateraler Exophthalmus, TSH erniedrigt, fT3 u. fT4 normal

Selbstbehandlung indiziert (*self care* = SC)

SC1	<b>Akute Bronchitis</b>	34 Jahre, w, seit 12 Tagen Husten, anfängliche verstopfte Nase und Halsschmerzen, Husten, kein Fieber
SC2	<b>Akute Bronchitis</b>	61 Jahre, w, seit 4 Tagen Husten, gelbes Sputum, Rhinorrhoe, anfänglich Fieber, jetzt kein Fieber mehr
SC3	<b>Akute Konjunktivitis</b>	14 Jahre, m, seit 3 Tagen rotes, gereiztes Auge (Ausbreitung von rechts nach links), Ausfluss, Symptome eines oberen Atemwegsinfektes, keine Schmerzen oder Lichtempfindlichkeit
SC4	<b>Akute Pharyngitis</b>	26 Jahre, m, seit 2 Tagen Halsschmerzen, Kopfschmerzen, Husten, kein Fieber
SC5	<b>Allergischer Schnupfen</b>	22 Jahre, m, seit 5 Jahren stetige Verschlechterung von verstopfter Nase, Niesen, Nasenjucken schlimmer im Frühjahr, Juckreiz im Auge, Rötung, Tränenfluss, Juckreiz im Gaumen und Hals, Ekzeme in der Kindheit
SC6	<b>Rückenschmerzen</b>	38 Jahre, m, akute Rückenschmerzen nach dem Heben, keine Schmerzen oder Schwäche in den Beinen, kein Fieber, Schüttelfrost, Gewichtsverlust oder kürzlich zurückliegende Infektionen
SC7	<b>Bienenstich ohne Anaphylaxie</b>	9 Jahre, m, Bienenstich, geschwollene und schmerzempfindliche Oberlippe; keine Schwellung der Zunge, kein Speichelfluss, kein Stridor, kein Hautausschlag oder andere Beschwerden
SC8	<b>Aphthe</b>	17 Jahre, m, mit rezidivierenden Ulzerationen, keine respiratorischen, anogenitalen, gastrointestinalen, Augen- oder Hautläsionen, Mutter hatte ähnliche Symptome in der Jugend, keine Medikamente
SC9	<b>Candida-Hefe-Infektion</b>	40 Jahre, w, seit 2 Tagen vaginaler Juckreiz, dicker weißer Ausfluss, keine Bauchschmerzen oder Fieber
SC10	<b>Verstopfung</b>	5 Monate, m, Schwierigkeit / Verspätung beim Passieren von hartem Stuhl, stundenlanger Stress, kann einen Tag verfehlen, schreit beim Stuhlgang und vereinzelte Blutflecken in der Windel, entwöhnt von



		Muttermilch zu Kuhmilch, ernährt sich normal
SC11	<b>Ekzem</b>	12 Jahre, w, trockene, juckende Haut an den Flexoren an der Ellenbeuge, hinter den Knien, am Sprunggelenk, Wangen haben Flecken, trockene, schuppige Haut, Heuschnupfen, Ei und Milchallergie, Bruder hat Asthma und Onkel und Cousins haben Ekzeme
SC12	<b>Hordeolum</b>	30 Jahre, m, seit dem Vortag schmerzhaftes, geschwollenes rechtes Auge, kein Trauma, keine Krustenbildung, keine Sehstörungen, keine Allergien oder Augenerkrankungen, lokalisierte Schmerzempfindlichkeit, Erythem (Rötung)
SC13	<b>Virale Infektion der oberen Atemwege</b>	56 Jahre, m, seit 6 Tagen Husten, verstopfte Nase, grüner Nasenausfluss, Fieber (38,2 ° C), Rhinorrhoe
SC14	<b>Virale Infektion der oberen Atemwege</b>	30 Jahre, m, seit 2 Tagen laufende Nase, Halsschmerzen, heiß, verschwitzt, leichte Kopfschmerzen, Husten mit klarem Auswurf, Muskelschmerzen, kein Fieber oder Nackensteifigkeit
SC15	<b>Erbrechen</b>	2 Jahre, m, leichtes Fieber (T = 38,0 ° C), zweimal erbrochen, erbricht Saft
SC16	<b>Refluxösophagitis</b>	55 Jahre, m, postprandiales, retrosternales Brennen, Druckgefühl in der Magenregion, BMI 30, 25 pack years

### **3.4. Die vollständigen Vignettentexte der Fallvignetten finden sich im Anhang, siehe**

Tabelle 25.

### **3.5. Methoden**

#### **3.5.1. Operationalisierung der Zieldimensionen**

Die Evaluation der diagnostischen Qualität der Assessments von Symptomcheckern und Medizinstudierenden erfolgte über zwei Parameter, zum einen wurde die Treffergenauigkeit, also die Frage nach der Präzision gestellter Diagnosen, dadurch bestimmt, ob die in den Assessments aufgeführten Leitdiagnosen (Top 1) den korrekten Diagnosen entsprachen. Zum anderen wurde die Trefferwahrscheinlichkeit, also die Frage nach der Angemessenheit der diagnostischen Überlegungen, dahingehend definiert, dass das diagnostische Assessment zumindest in den ersten drei Differentialdiagnosen (Top 3) die korrekte Diagnose zeigte.

Die Triagequalität der untersuchten Symptomchecker und der studentischen Assessments wurde daraufhin geprüft, ob sich die Handlungsempfehlungen korrekt zu den drei Dringlichkeitsstufen zuordnen ließen, die als Kategorien in der Harvard-Studie von Semigran et al.<sup>56</sup> Verwendung gefunden hatten.

Die höchste Dringlichkeitsstufe war „*emergent care*“ (EC) und umfasste Triageempfehlungen, die medizinische Notfallversorgung zu kontaktieren oder sofortige medizinische Hilfe aufzusuchen. Darunter fiel auch der Hinweis einen Rettungswagen zu rufen. Ebenfalls zugehörig waren Empfehlungen einen Hausarzt umgehend, beziehungsweise sofort, aufzusuchen. Die explizite Verwendung einer Formulierung, die die Dringlichkeit der empfohlenen Konsultation hervorhebt, war dabei für die Einordnung in diese Dringlichkeitsstufe entscheidend.

Die mittlere Dringlichkeitsstufe war „*non-emergent care*“ (NE) und umfasste jegliche Empfehlung, medizinische Hilfe in Anspruch zu nehmen. Dies beinhaltete die Empfehlung einen Hausarzt oder ähnlichen Erstversorger zu rufen oder einen solchen aufzusuchen. Ebenfalls dazugehörig waren Empfehlungen eine Notfallpraxis oder einen Spezialisten aufzusuchen. Die Konsultation einer Online-Sprechstunde oder einer *retail clinic*, einem medizinischen Service ohne Ärzte in Apotheken etc., wurde auch in diese Kategorie eingeordnet. Beides sind jedoch keine medizinischen Anlaufstellen, die im deutschen Gesundheitssystem üblich sind.

Die niedrigste Dringlichkeitsstufe war „*self care*“ (SC) und umfasste Triageempfehlungen zum Verbleib zu Hause oder zum Aufsuchen einer Apotheke. Bei Triageempfehlungen, medizinische Hilfe in Anspruch zu nehmen, wenn auf Dauer keine Besserung eintritt, die Symptome sich verschlechtern oder Symptome auftreten, die nicht in der Fallvignette genannt wurden, erfolgte ebenfalls eine Einordnung in die niedrigste Dringlichkeitsstufe „*self care*“.

Die Verwendung dieser Dringlichkeitsstufen erlaubte die unmittelbare Vergleichbarkeit mit der Vorgängerstudie von Semigran et al.<sup>56</sup>

### **3.5.2. Versuchsaufbau der Studierendenuntersuchung**

Um die Performance der Studierenden zu ermitteln, wurden ihnen je 6 klinische Fallvignetten zur Bearbeitung vorgelegt. Darunter waren je zwei Fallvignetten hoher Dringlichkeit, zwei Fallvignetten mittlerer Dringlichkeit und zwei Fallvignetten niedriger Dringlichkeit. Je eine der beiden Fallvignetten einer Dringlichkeitsstufe wurde in Form einer Chatanamnese bearbeitet und je eine als sogenannter Textfall. Die Zuteilung der Fallvignetten wurde im Vorfeld randomisiert.

Als Ort für die Untersuchung wurde das Studierendenhaus der Universitätsklinik Köln gewählt. Die Probanden befanden sich dabei allein in einem Raum, in welchem ein Laptop aufgebaut war (siehe Abbildung 2).



Abbildung 2. Versuchsaufbau der Studierendenuntersuchung

Die Bearbeitung der Fallvignetten erfolgte über die Chatfunktion des Instant-Messenger-Dienstes Skype. Weder die Sprach- noch die Videotelefoniefunktion des Dienstes wurden verwendet. Die Kommunikation lief über zwei eigens für diesen Zweck angelegte Nutzerkonten. Während der gesamten Untersuchung bestand kein persönlicher Kontakt zu den Versuchsleitern. Es standen den Probanden keine Hilfsmittel zur Verfügung.

Die Untersuchung fand im Sommersemester 2018 statt. Der Verfahrensablauf ist der Abbildung 3 zu entnehmen.

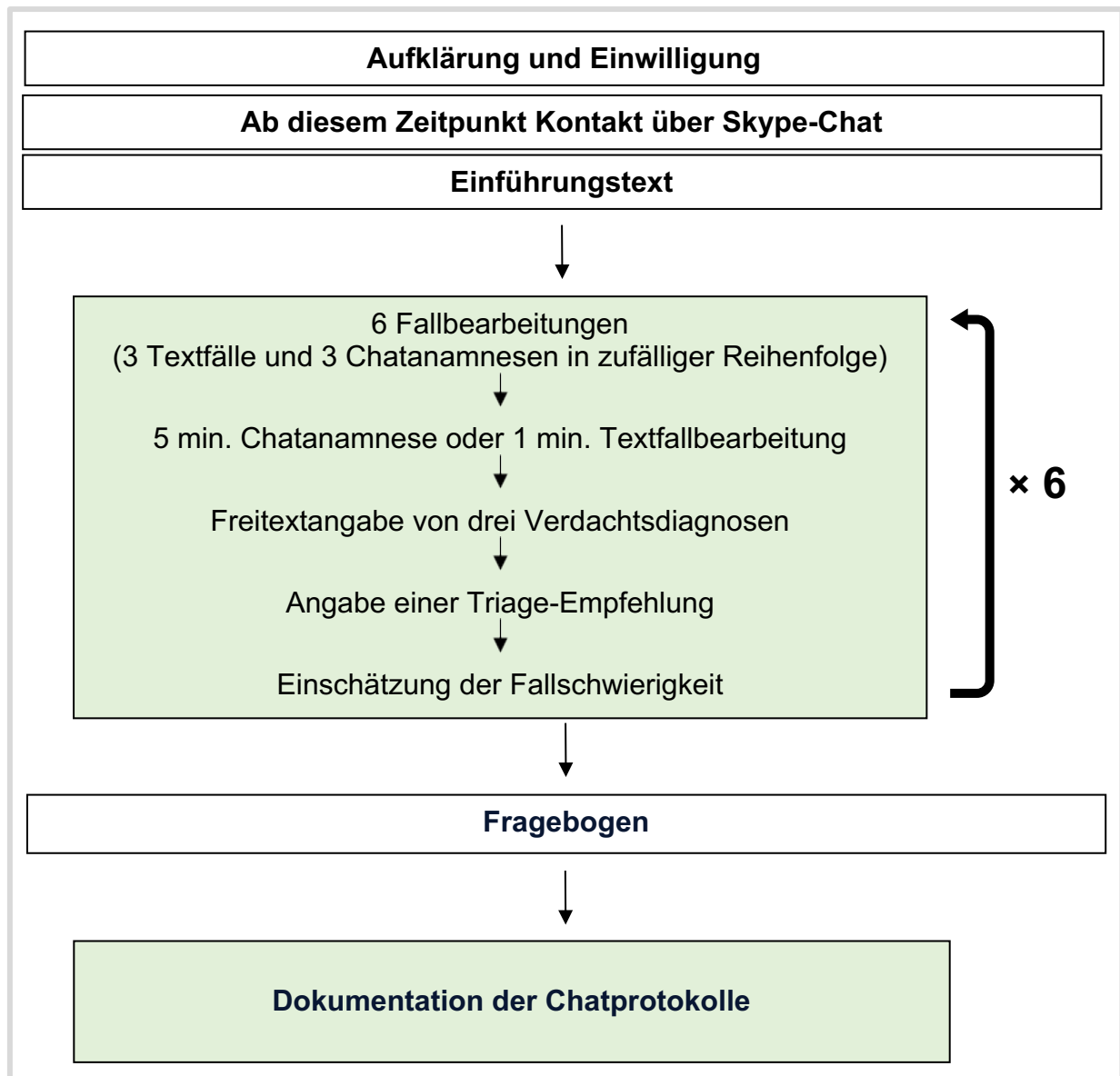


Abbildung 3. Ablauf der Studierendenuntersuchung

Nach einem Einführungstext, der den Ablauf der Fallbearbeitungen beschrieb, wurden von den Probanden je 6 Fälle bearbeitet. Davon wurden drei per Chatanamnese bearbeitet und drei als Textfälle.

Die Chatanamnese umfasste den Vorgang, dass den Probanden fünf Minuten Zeit zur Verfügung stand, um durch Fragen Informationen von einem fiktiven Patienten sammeln zu können. Die Antworten wurden von den Versuchsleitern gegeben. Um die Antwortzeit zu verkürzen und eine möglichst standardisierte Formulierung der Informationen zu erreichen, wurde auf vorformulierte Textbausteine zurückgegriffen. Bedeutungsgleiche Fragestellungen wurden identisch beantwortet. Fragen nach Informationen oder Symptomen, die nicht in der jeweiligen Vignette angegeben waren, wurden beantwortet, als könne der Patient keine

Auskunft geben (z.B. „Das weiß ich nicht“ oder „Das kann ich Ihnen nicht sagen“). Nach Ablauf der Bearbeitungszeit wurden keine weiteren Fragen beantwortet.

Eine Bearbeitung als Textfall bedeutete in diesem Zusammenhang, dass den Probanden der gesamte Text der jeweiligen Vignette vorlag. Die Bearbeitungszeit betrug dabei eine Minute. Rückfragen wurden nicht beantwortet.

Nach Ablauf der Bearbeitungszeit des jeweiligen Falles wurde um die Angabe von drei Verdachtsdiagnosen, geordnet von der wahrscheinlichsten hin zur unwahrscheinlichsten Diagnose, gebeten. Anschließend wurde die Einschätzung der Dringlichkeit des Falles erbeten. Zu wählen war unter folgenden Optionen:

- A. Hohe Dringlichkeit (Krankenhouseinweisung)
- B. Mittlere Dringlichkeit (Aufsuchen eines niedergelassenen Arztes)
- C. Geringe Dringlichkeit (Zuhause selbst auskurieren)

Zuletzt sollte die individuell empfundene Schwierigkeit der Fallbearbeitung auf einer numerischen Skala von 1 bis 10 angegeben werden.

Diese Reihenfolge der Fallbearbeitung wurde sechs Mal wiederholt. Daraufhin erreichten die Probanden über einen Hyperlink einen Fragebogen in Form eines Google-Forms-Formular. Dieser Fragebogen beinhaltete Fragen zur Vorerfahrung mit Computern und digitalen Lernmedien sowie zur Bewertung von E-Learning und dessen Implementierung in die Lehre.

### **3.5.3. Versuchsaufbau der Symptomcheckeruntersuchung**

Um die Performance der Symptomchecker zu ermitteln, wurde jeder der eingeschlossenen Symptomchecker anhand jeder der 49 Fallvignetten getestet.

Die Testung fand nach Einstellung des privaten Modus im Internetbrowser Firefox statt.

Aufgrund der unterschiedlichen Funktionsweisen der Symptomchecker wurde im Verlauf der Untersuchung die jeweils erforderliche Information in Übereinstimmung mit der jeweiligen Fallvignette eingegeben. Bei Nachfragen von Seite der Symptomchecker wurde ebenfalls nur in Übereinstimmung mit der Fallvignette geantwortet. Dies galt ebenfalls für Fragen zu Intensität oder Dauer von Symptomen.

Die Konsultation eines Symptomcheckers wurde als „nicht zu bearbeiten“ eingestuft, sofern sich keine Verdachtsdiagnosen beziehungsweise Triagierung ergab. In einigen Fällen ließen sich zudem nur Fallvignetten ab einem definierten Patientenalter eintragen. Bei fehlender Übereinstimmung des Mindestalters mit der Fallvignette wurde die Konsultation eines Symptomcheckers ebenfalls als „nicht zu bearbeiten“ vermerkt.

### **3.5.4. Dokumentations- und Auswertungsmethoden**

Eine Übersicht der in der Untersuchung verwendeten Programme und Geräte findet sich in Tabelle 3.

Tabelle 3. Verwendete Programme und Geräte der Studie

Programm	Verwendungszweck	Hersteller
Microsoft Windows Vista	Betriebssystem	Microsoft Corporation, Redmond, WA, USA
Skype Version 7.41	Fallbearbeitung Studierende	Microsoft Corporation, Redmond, WA, USA
Microsoft Windows 10	Betriebssystem	Microsoft Corporation, Redmond, WA, USA
Microsoft Excel für Office 365 MSO (16.0.12325.20280) 32-bit	Tabellenkalkulation	Microsoft Corporation, Redmond, WA, USA
Microsoft Word für Office 365 MSO (16.0.12325.20280) 32-bit	Textverarbeitung	Microsoft Corporation, Redmond, WA, USA
IBM SPSS Statistics Version 25	Tabellenkalkulation	IBM Corporation, Armonk, NY, USA
Acer Notebook	Texteingabe	Acer Group, Taipeh, Taiwan

Die Erhebung des Fragebogens wurde mithilfe eines Google-Forms-Formulars durchgeführt und gespeichert. Die Chatprotokolle wurden nach Testabschluss der Probanden als Textdatei gespeichert. Erst nach Abschluss der Studie erfolgte die Auswertung.

Die Ergebnisse der Symptomcheckeruntersuchung wurden in Protokollen dokumentiert. Orthografiefehler und Formulierungsvarianten wurden dabei nicht abgeändert. Die Formulierungen der Triageempfehlungen wurden ebenfalls dokumentiert. Erst nach Abschluss der Untersuchung erfolgte die Auswertung.

### 3.5.5. Statistische Methoden

Aus den Protokollen der Chatverläufe der Probanden wurden die Verdachtsdiagnosen, Dringlichkeitseinstufungen, die Antworten zu den Fragen zur subjektiven Schwierigkeit und die Anzahl gestellter Fragen in SPSS Statistics (Version 25) übertragen. Die Auswertung des *Data sheets* erfolgte im gleichen Programm.

Aus den Protokollen der Testung der Symptomchecker wurden die Verdachtsdiagnosen und Dringlichkeitseinstufungen Fragen in SPSS Statistics (Version 25) übertragen. Die Auswertung des *Data sheets* erfolgte im gleichen Programm.

Die Angaben aus dem Posttest-Fragebogen wurden in SPSS Statistics (Version 25) übertragen. Die Auswertung des *Data sheets* erfolgte im gleichen Programm.

## 4. Ergebnisse

### 4.1. Symptomchecker-Assessments

#### 4.1.1. Diagnosequalität der Symptomchecker-Assessments

Die Treffergenauigkeit, das heißt die Präzision der Assessments, gibt die Wahrscheinlichkeit wieder, dass die korrekte Diagnose an erster Stelle (Top 1) gelistet ist. Dies war in 38.02% der Fallbearbeitungen zutreffend (siehe Tabelle 4). Die Präzision der Assessments variierte nicht wesentlich zwischen den Dringlichkeitsstufen (EC, NE, SC). Die Treffergenauigkeit war bei der höchsten Stufe der Interventionsdringlichkeit am geringsten.

Tabelle 4. Diagnostische Treffergenauigkeit der Symptomchecker

Symptomchecker gesamt					
	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)	P-Wert	Chi-Quadrat
Top 1	49 (100.00)	254/668	38.02 (34.73–47.54)		
Art der Vignetten					
EC	17 (34.69)	86/236	36.44 (30.49–42.49)	.627	0.93
NE	16 (32.65)	83/223	37.22 (30.90–43.43)		
SC	16 (32.65)	85/209	40.67 (33.81–47.39)		

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose in 56.44% der Fälle im Sinne der Trefferwahrscheinlichkeit gelistet (siehe Tabelle 5). Auch bezüglich der Trefferwahrscheinlichkeit waren die Diagnosen in der höchsten Dringlichkeitsstufe am wenigsten häufig korrekt.

Tabelle 5. Diagnostische Trefferwahrscheinlichkeit der Symptomchecker

Symptomchecker gesamt					
	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)	P-Wert	Chi-Quadrat
Top 3	49 (100.00)	377/668	56.44 (52.45–60.19)		
Art der Vignetten					
EC	17 (34.69)	123/236	52.11 (45.89–58.52)	.072	5.25
NE	16 (32.65)	123/223	55.16 (48.28–61.70)		
SC	16 (32.65)	131/209	62.68 (56.31–69.16)		

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

#### 4.1.2. Triagequalität der Symptomchecker-Assessments

Die korrekte Triagierung wurde in 59.19% der Fallbearbeitungen genannt. Dabei unterschied sich die Triagequalität der Symptomchecker innerhalb der verschiedenen Dringlichkeitsstufen signifikant,  $p < .001$  (siehe Tabelle 6).

Tabelle 6. Triagequalität der Symptomchecker

Symptomchecker gesamt					
	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)	P-Wert	Chi-Quadrat
Triage	49 (100.00)	264/446	59.19 (54.67–63.68)		
Art der Vignetten					
EC	17 (34.69)	116/162	71.60 (63.74–78.57)	.001	48.37
NE	16 (32.65)	98/143	68.53 (60.88–76.16)		
SC	16 (32.65)	50/141	35.46 (27.50–43.62)		

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

#### 4.1.3. Differenzielle Ergebnisse der Symptomchecker-Assessments

Insgesamt wurde von den deutschsprachigen Symptomcheckern die korrekte Diagnose in 48.90% an erster Stelle genannt (Top 1). Die englischsprachigen Symptomchecker nannten die korrekte Diagnose nur in 33.95% an erster Stelle, siehe Tabelle 7. Dieser Unterschied war signifikant,  $p = .001$ . Zusätzlich unterschieden sich deutschsprachige und englischsprachige Symptomchecker signifikant in der Genauigkeit in den Dringlichkeitsstufen Emergent Care,  $p = .027$  und Self Care,  $p = .008$ .

Tabelle 7. Vergleich der diagnostischen Treffergenauigkeit deutsch- und englischsprachiger Symptomchecker

Vergleich deutsch- und englischsprachiger Symptomchecker							
Vignetten	Anzahl an Vignetten (%)	deutschsprachig		englischsprachig		P-Wert	Chi-Quadrat
		Verhältnis	% (95% KI)	Verhältnis	% (95% KI)		
Top 1							
Alle	49 (100.00)	89/182	48.90 (42.11–56.34)	165/486	33.95 (29.73–38.21)	.001	12.56
EC	17 (34.69)	31/65	47.69 (35.48–58.93)	55/171	32.16 (25.53–39.18)	.027	4.90
NE	16 (32.65)	26/59	44.07 (31.38–57.62)	57/164	34.76 (27.41–42.26)	.204	1.61
SC	16 (32.65)	32/58	55.17 (42.00–68.75)	53/151	35.10 (27.97–43.12)	.008	7.00

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose von den deutschsprachigen Symptomcheckern in 69.23% der Fälle gelistet. Von den



englischsprachigen Symptomcheckern wurde innerhalb der ersten drei gegebenen Diagnosen die korrekte Diagnose in 51.65% der Fälle gelistet. Dieser Unterschied war ebenfalls signifikant,  $p = .001$ . Zusätzlich unterschieden sich deutschsprachige und englischsprachige Symptomchecker bezüglich der diagnostischen Trefferwahrscheinlichkeit signifikant in den Dringlichkeitsstufen Emergent Care,  $p = .003$ , und Self Care,  $p = .006$  (siehe Tabelle 8).

Tabelle 8. Vergleich der diagnostischen Trefferwahrscheinlichkeit deutsch- und englischsprachiger Symptomchecker

Vergleich deutsch- und englischsprachiger Symptomchecker							
Vignetten	Anzahl an Vignetten (%)	deutschsprachig		englischsprachig		P-Wert	Chi-Quadrat
		Verhältnis	% (95% KI)	Verhältnis	% (95% KI)		
Top 3							
Alle	49 (100.00)	126/182	69.23 (62.22–75.29)	251/486	51.65 (47.24–56.03)	.001	16.65
EC	17 (34.69)	44/65	67.69 (55.56–78.69)	79/171	46.20 (38.61–53.45)	.003	8.72
NE	16 (32.65)	37/59	62.71 (50.00–74.64)	86/164	52.44 (44.51–60.37)	.174	1.85
SC	16 (32.65)	45/58	77.59 (66.68–88.46)	86/151	56.95 (48.63–64.33)	.006	7.63

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

Die richtige Triagierung wurde von den deutschsprachigen Symptomcheckern in 65.22% der Fallbearbeitungen und von den englischsprachigen Symptomcheckern in 56.49% korrekt genannt. Dabei unterschied sich die Triagequalität der Symptomchecker nur innerhalb der Dringlichkeitsstufe Self Care signifikant,  $p = .049$  (siehe Tabelle 9).

Tabelle 9. Vergleich der Triagequalität deutsch- und englischsprachiger Symptomchecker

Vergleich deutsch- und englischsprachige Symptomchecker							
Vignetten	Anzahl an Vignetten (%)	deutschsprachig		englischsprachig		P-Wert	Chi-Quadrat
		Verhältnis	% (95% KI)	Verhältnis	% (95% KI)		
Triage							
Alle	49 (100.00)	90/138	65.22 (57.50–72.81)	174/308	56.49 (50.66–62.26)	.083	3.00
EC	17 (34.69)	37/52	71.15 (57.50–83.63)	79/110	71.82 (63.30–79.44)	.930	0.01
NE	16 (32.65)	33/44	75.00 (61.54–87.17)	65/99	65.66 (55.86–75.27)	.267	1.23
SC	16 (32.65)	20/42	47.61 (32.56–62.50)	30/99	30.30 (21.78–39.80)	.049	3.86

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

Treffergenauigkeit (Top 1) und -wahrscheinlichkeit (Top 3) variierten bei den einzelnen untersuchten Symptomchecker erheblich. Die präzise Diagnose (Top 1) etwa wurde

zwischen 9% und 59% der Fallbearbeitungen gefunden, auch zeigte sich eine große Variabilität in Bezug auf die Trefferwahrscheinlichkeit (Top 3). Ein Symptomchecker stellte keine Diagnosen (siehe Tabelle 10).

Tabelle 10. Differenzielle Diagnosequalität der Symptomchecker

Symptomchecker	Diagnosequalität			
	Top 1		Top 3	
	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)
<b>deutschsprachig</b>				
Ada	29/49	59 (44–73)	40/49	82 (70–92)
	25/49	51 (36–65)	34/49	69 (56–83)
Diagnosefinder.apotheken				
.de				
Netdoktor	- 16/49	33 (20–47)	28/49	57 (42–71)
symptomchecker				
Symptomate	19/35	54 (39–71)	24/35	69 (52–84)
<b>englischsprachig</b>				
AskMD Sharecare	14/48	29 (16–43)	26/48	54 (40–68)
Doctor Diagnose Android	14/36	39 (22–54)	15/36	42 (26–57)
Drugs.com	17/35	49 (32–65)	20/35	57 (41–73)
Esagil	7/48	15 (5–25)	11/48	23 (11–36)
Familydoctor.org	12/34	35 (19–52)	14/34	41 (24–58)
Isabelhealthcare	27/49	55 (42–70)	33/49	67 (54–80)
Mayoclinic	13/49	27 (15–39)	28/49	57 (42–70)
Medoctor	4/45	9 (2–18)	17/45	38 (24–52)
Myhealth.alberta	- —	—	—	—
Healthwise				
Symcat	16/49	33 (20–46)	31/49	63 (50–76)
Symptify	20/44	46 (30–60)	23/44	52 (38–68)
Symptoms.webmd	21/49	43 (29–58)	33/49	67 (54–80)

Hinsichtlich der Triagierung gaben fünf Symptomchecker keine Handlungsempfehlungen. Korrekte Dringlichkeitseinstufungen wurden zwischen 49% und 80% der Fallvignetten gegeben. Dabei zeigten sich die besten Ergebnisse im Erkennen von Notfallsituationen (siehe Tabelle 11).

Tabelle 11. Differenzielle Triagequalität der Symptomchecker nach Dringlichkeitsstufen

Symptomchecker	Triagequalität nach Dringlichkeitsstufen							
	Alle		Emergent		Non-emergent		Self care	
	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)
<b>deutschsprachig</b>								
Ada	32/49	65 (52–78)	13/17	77 (55–95)	9/16	56 (30–80)	10/16	63 (36–86)
Diagnosefinder.	30/49	61 (48–74)	9/17	53 (27–77)	13/16	81 (60–100)	8/16	50 (25–77)
<b>apotheken.de</b>								
Netdoktor	- 4/5	80 (0–100)	4/4	100 (0–100)	0/1	0 (0–0)	—	—
<b>symptomchecker</b>								
Symptomate	24/35	69 (53–82)	11/14	79 (53–100)	11/11	100 (0–100)	2/10	20 (0–50)
<b>englischsprachig</b>								
AskMD Sharecare	—	—	—	—	—	—	—	—
Doctor Diagnose	9/21	43 (22–65)	7/10	70 (38–100)	2/4	50 (0–100)	0/7	0 (0–0)
<b>Android</b>								
Drugs.com	28/49	57 (43–71)	9/17	53 (28–77)	10/16	63 (36–86)	9/16	56 (31–81)
Esagil	—	—	—	—	—	—	—	—
Familydoctor.org	29/47	62 (47–75)	9/17	53 (29–79)	11/15	73 (50–94)	9/15	60 (31–86)
Isabelhealthcare	29/49	59 (46–74)	16/17	94 (81–100)	13/16	81 (60–100)	0/16	0 (0–0)
Mayoclinic	—	—	—	—	—	—	—	—
Medoctor	—	—	—	—	—	—	—	—
Myhealth.alberta	- 29/49	59 (45–73)	15/17	88 (70–100)	8/16	50 (25–77)	6/16	38 (11–63)
<b>Healthwise</b>								
Symcat	24/49	49 (34–63)	10/17	59 (33–83)	14/16	88 (69–100)	0/16	0 (0–0)
Symptify	26/44	59 (45–73)	13/15	87 (65–100)	7/16	44 (19–69)	6/13	46 (18–75)
Symptoms.webmd	—	—	—	—	—	—	—	—

## 4.2. Studierenden-Assessments

### 4.2.1. Demographische und Fachsemesterdaten der Studierenden

Insgesamt nahmen 143 Studierende (53 männlich, 89 weiblich, 1 andere; mittleres Alter = 23.77, range: 20-34 Jahre) teil. Für detaillierte demographische Angaben siehe Tabelle 12.

Tabelle 12. Demographische und Fachsemesterdaten Studierenden

N = 143	Männer (n = 53)	Frauen (n = 89)	Andere (n = 1)
Alter (Mittelwert, SD)	23.64 (2.19)	23.85 (2.40)	23.00
Fachsemester			
1. Klinisches Semester	2	1	0
2. Klinisches Semester	16	25	1
3. Klinisches Semester	2	6	0
4. Klinisches Semester	4	6	0
5. Klinisches Semester	20	43	0
6. Klinisches Semester	9	8	0

#### 4.2.2. Diagnosequalität der Studierenden-Assessments

Von den Studierenden wurde die korrekte Diagnose in 64.10% der Fallbearbeitungen an erster Stelle (Top 1) gelistet (siehe Tabelle 13). Präzise Diagnosen der Studierenden zeigten keinen Zusammenhang mit den Dringlichkeitsstufen.

Tabelle 13. Diagnostische Treffergenauigkeit der Studierenden-Assessments

		Studierende (Textbearbeitung)	
Vignetten	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)
Top 1	49 (100.00)	275/429	64.10 (59.56–68.63)
Art der Vignetten			
EC	17 (34.69)	87/143	60.84 (52.70–68.89)
NE	16 (32.65)	98/143	68.53 (60.67–76.03)
SC	16 (32.65)	90/143	62.94 (55.56–71.31)

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose von den Studierenden in 56.44% der Fälle gelistet, auch hier ergab sich kein systematischer Zusammenhang mit den Dringlichkeitsstufen (siehe Tabelle 14).

Tabelle 14. Diagnostische Trefferwahrscheinlichkeit der Studierenden-Assessments

		Studierende (Textbearbeitung)	
Vignetten	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)
Top 3	49 (100.00)	344/429	80.19 (76.46–83.59)
Art der Vignetten			
EC	17 (34.69)	98/143	68.53 (60.67–76.03)
NE	16 (32.65)	124/143	86.71 (80.77–92.31)
SC	16 (32.65)	122/143	85.31 (79.22–91.25)

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

### 4.2.3. Triagequalität der Studierenden-Assessments

Die korrekte Triagierung wurde von den Studierenden in 59.19% der Fallbearbeitungen genannt. Es ergaben sich Zusammenhänge zwischen der Triagequalität und der Dringlichkeitsstufe der jeweiligen Fallvignetten (siehe Tabelle 15).

Tabelle 15. Triagequalität der Studierenden-Assessments

Vignetten	Anzahl an Vignetten (%)	Studierende (Textbearbeitung)	
		Verhältnis	% (95% KI)
Triage	49 (100.00)	257/429	59.91 (54.99–64.46)
Art der Vignetten			
EC	17 (34.69)	124/143	86.71 (80.77–92.31)
NE	16 (32.65)	87/143	60.84 (52.70–68.89)
SC	16 (32.65)	46/143	32.17 (24.63–40.00)

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

### 4.2.4. Studierenden-Assessments und Studienfortschritt

Zur Untersuchung von Zusammenhängen der Diagnoseleistung und dem Studienfortschritt wurden folgende Berechnungen durchgeführt: Die Modalitäten (Text und Chat), die Dringlichkeitsstufen der Vignetten (Emergent Care, Non-emergent Care, Self Care) und der Studienfortschritt der Studierenden (erste Hälfte oder zweite Hälfte des klinischen Studienabschnitts) wurden mithilfe von *Generalized estimating equations* (GEE) analysiert, in einem  $2 \times 3 \times 2$  - Design (Modalität  $\times$  Dringlichkeitsstufen  $\times$  Studienfortschritt). Neben signifikanten Effekten für die Modalitäten ( $b = -0.253$ , Wald  $\chi^2(1) = 13.37$ ,  $p < .001$ , siehe Abbildung 4) und der Dringlichkeitsstufen ( $b = -0.981$ , Wald  $\chi^2(2) = 10.84$ ,  $p = .004$ ), ergab sich ein signifikanter Haupteffekt für den Studienfortschritt ( $b = -0.347$ , Wald  $\chi^2(1) = 6.37$ ,  $p = .012$ ). Dies bedeutet, dass Studierende in der zweiten Hälfte des klinischen Studienabschnitts für mehr Fallbearbeitungen korrekte Diagnosen fanden, als Studierende in der ersten Hälfte des klinischen Studienabschnitts. Darüber hinaus gab es noch einen signifikanten Interaktionseffekt zwischen Modalität und Dringlichkeitsstufen ( $b = -0.893$ , Wald  $\chi^2(2) = 12.47$ ,  $p = .002$ ). Alle anderen Interaktionseffekte waren nicht signifikant, alle Wald  $\chi^2 < 1.43$ .

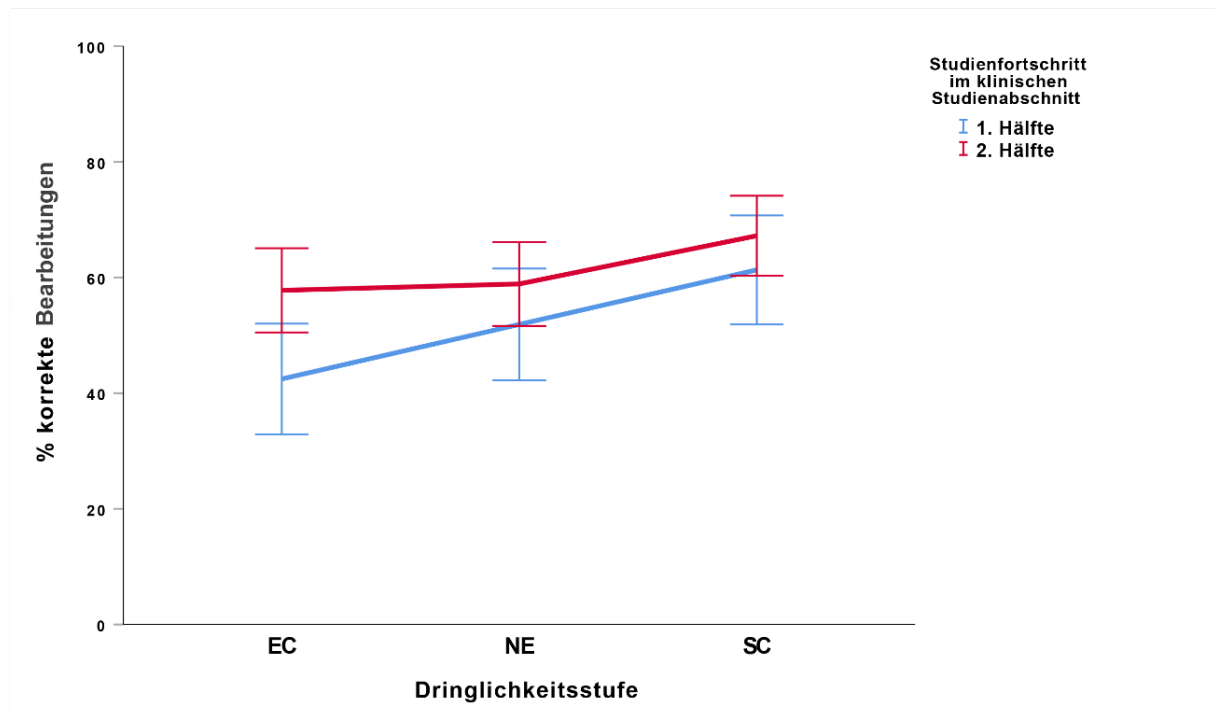


Abbildung 4. Vergleich der Diagnosekorrektheit in Textbearbeitung und Chatbearbeitung für Studierende in der ersten und zweiten Hälfte des klinischen Studienabschnitts

Zur Untersuchung von Zusammenhängen der Triagequalität und dem Studienfortschritt wurden folgende Berechnungen durchgeführt. Die Modalitäten (Text und Chat), die Dringlichkeitsstufen der Vignetten (Emergent Care, Non-emergent Care, Self Care) und der Studienfortschritt der Studierenden (erste Hälfte oder zweite Hälfte des klinischen Studienabschnitts) wurden mithilfe der GEE analysiert, in einem  $2 \times 3 \times 2$  - Design (Modalität  $\times$  Dringlichkeitsstufen  $\times$  Studienfortschritt). Neben einem signifikanten Haupteffekt für Dringlichkeitsstufen ( $b = 1.35$ , Wald  $\chi^2(2) = 100.86$ ,  $p < .001$ , siehe Abbildung 5), ergab sich ein signifikanter Interaktionseffekt zwischen Modalität und Dringlichkeitsstufen ( $b = 1.68$ , Wald  $\chi^2(2) = 22.13$ ,  $p < .001$ ). Alle anderen Haupt- und Interaktionseffekte waren nicht signifikant, alle Wald  $\chi^2 < 5.90$ . Dies bedeutet, dass Studierende in der zweiten Hälfte des klinischen Studienabschnitts in mehr Fallbearbeitungen richtig triagierten, als Studierende in der ersten Hälfte des klinischen Studienabschnitts.

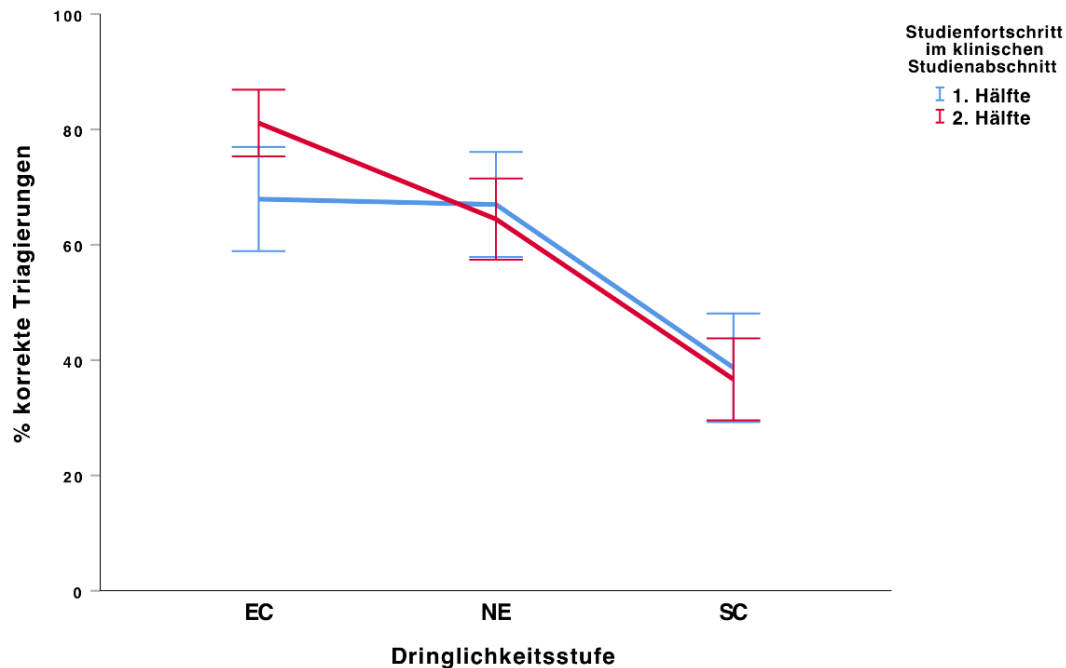


Abbildung 5. Vergleich der Triagequalität in Textbearbeitung und Chatbearbeitung für Studierende in der ersten und zweiten Hälfte des klinischen Studienabschnitts

#### 4.2.5. Studierenden-Assessments und Symptompräsentation

Um zu untersuchen, ob sich die Modalität (Text oder Chat) und die Dringlichkeitsstufen (Emergent, Non-emergent, Self Care) der Vignetten auf die Wahrscheinlichkeit der Studierenden die Vignetten präzise zu lösen auswirken, wurde eine Analyse mittels der GEE in einem  $2 \times 3$  Design (Modalität  $\times$  Dringlichkeitsstufen) durchgeführt. Es gab einen signifikanten Haupteffekt für Modalität,  $b = -0.185$ , Wald  $\chi^2(1) = 13.30$ ,  $p < .001$ . Es gab außerdem einen signifikanten Haupteffekt für Dringlichkeitsstufen,  $b = -0.982$ , Wald  $\chi^2(2) = 10.13$ ,  $p = .006$ . Darüber hinaus gab es noch einen signifikanten Interaktionseffekt zwischen Modalität und Dringlichkeitsstufen,  $b = -0.893$ , Wald  $\chi^2(2) = 14.23$ ,  $p = .001$ . Um diesen Interaktionseffekt genauer zu untersuchen, wurden  $\chi^2$ -Tests durchgeführt.

Eine signifikant größere Anzahl an Studierenden löste die Vignetten mit der Dringlichkeitsstufe Emergent korrekt, wenn die Vignetten als Text bearbeitet wurden, 87 von 143 (60.84%), im Vergleich zur Bearbeitung in Chatform, 62 von 143 (43.36%),  $\chi^2(1, N = 286) = 8.76$ ,  $p = .003$ , für Ergebnisse siehe Abbildung 6.

Eine signifikant größere Anzahl an Studierenden löste die Vignetten mit der Dringlichkeitsstufe Non-emergent korrekt, wenn die Vignetten als Text bearbeitet wurden, 98 von 143 (68.53%), im Vergleich zur Bearbeitung in Chatform, 63 von 143 (44.06%),  $\chi^2(1, N = 286) = 17.41$ ,  $p < .001$ .

Der Unterschied an korrekten Lösungen der Vignetten in der Dringlichkeitsstufe Self Care zwischen der Bearbeitung in Textform, 90 von 143 (62.94%), und der Bearbeitung in Chatform, 96 von 143 (67.13%), war nicht signifikant,  $\chi^2(1, N = 286) = 0.55, p = .457$ .

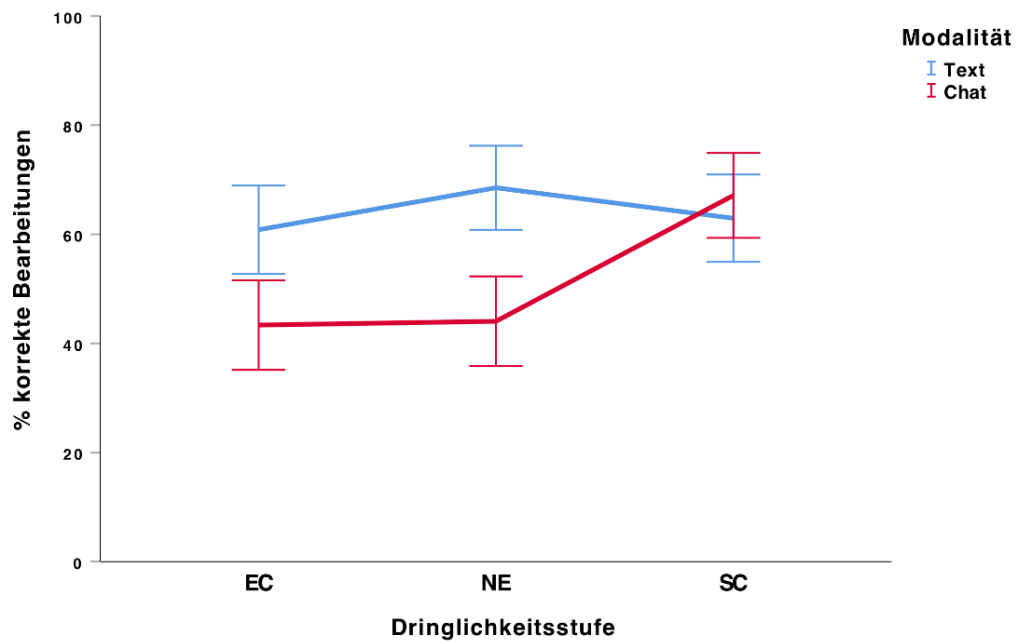


Abbildung 6. Korrekte Diagnosestellungen nach Dringlichkeitsstufen in den Modalitäten Text und Chat

Dargestellt ist der prozentuale Anteil korrekter Diagnosestellungen für die Vignetten in den verschiedenen Dringlichkeitsstufen für die Modalitäten Text und Chat. (EC = Emergent Care; NE = Non-emergent Care; SC = Self Care)

Um zu untersuchen, ob sich die Modalität (Text oder Chat) und die Dringlichkeitsstufen (Emergent, Non-emergent, Self Care) der Vignetten auf die Wahrscheinlichkeit der Studierenden die Vignetten richtig zu triagieren auswirken, wurde eine Analyse mit *Generalized estimating equations* (GEE) in einem  $2 \times 3$  Design (Modalität  $\times$  Dringlichkeitsstufen) durchgeführt. Es gab keinen signifikanten Haupteffekt für Modalität,  $b = -0.450$ , Wald  $\chi^2(1) = 0.60, p = .438$ . Es gab einen signifikanten Haupteffekt für Dringlichkeitsstufen,  $b = 0.947$ , Wald  $\chi^2(2) = 113.61, p < .001$ . Darüber hinaus gab es noch einen signifikanten Interaktionseffekt zwischen Modalität und Dringlichkeitsstufen,  $b = 1.68$ , Wald  $\chi^2(2) = 23.00, p < .001$ . Um diesen Interaktionseffekt genauer zu untersuchen, wurden  $\chi^2$ -Tests durchgeführt.

Eine signifikant größere Anzahl an Studierenden triagierte die Vignetten mit der Dringlichkeitsstufe Emergent korrekt, wenn die Vignetten als Text bearbeitet wurden, 124



von 143 (86.71%), im Vergleich zur Bearbeitung in Chatform, 94 von 143 (65.73%),  $\chi^2(1, N = 286) = 17.36, p < .001$ , für Ergebnisse siehe Abbildung 7.

Der Unterschied an korrekten Triagen der Vignetten in der Dringlichkeitsstufe Non-emergent zwischen der Bearbeitung in Textform, 87 von 143 (60.83%), und der Bearbeitung in Chatform, 100 von 143 (69.93%), war nicht signifikant,  $\chi^2(1, N = 286) = 2.61, p = .106$ .

Der Unterschied an korrekten Triagen der Vignetten in der Dringlichkeitsstufe Self Care zwischen der Bearbeitung in Textform, 46 von 143 (32.17%), und der Bearbeitung in Chatform, 61 von 143 (42.66%), war nicht signifikant,  $\chi^2(1, N = 286) = 3.36, p = .067$ .

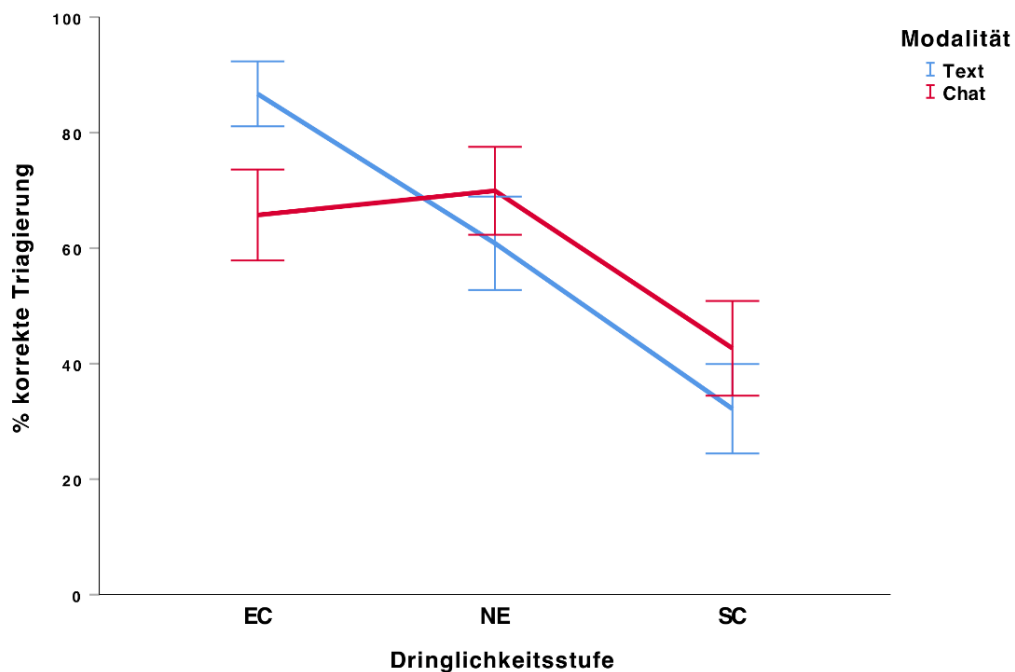


Abbildung 7. Triagequalität nach Dringlichkeitsstufe in den Modalitäten Text und Chat

Dargestellt ist der prozentuale Anteil an korrekten Triagierungen für die Vignetten in den verschiedenen Dringlichkeitsstufen für die Modalitäten Text und Chat. (EC = Emergent Care; NE = Non-emergent Care; SC = Self Care)

#### 4.2.6. Studierenden-Assessments und subjektive Bewertungsdimensionen

Die subjektiv wahrgenommene Schwierigkeit der Text- und der Chatbedingung wurde mit einem  $t$ -Test für abhängige Stichproben verglichen. Hierbei bewerteten die Studierenden die Schwierigkeit der Textbedingung ( $M = 5.57, SD = 1.75$ ) als signifikant geringer als die Schwierigkeit der Chatbedingung ( $M = 6.27, SD = 1.77$ ),  $t(142) = 4.35, p < .001, d = 0.397$ .

Im Fragebogen wurde auch der Informationsbedarf im Chat abgefragt. Auf einer Skala von 1 (sehr viele zusätzliche Informationen) bis 10 (überhaupt keine zusätzlichen Informationen) gaben die Studierenden einen zusätzlichen Bedarf von durchschnittlich 3.91 ( $SD = 1.98$ ) an.

Um den Zusammenhang zwischen der subjektiven Schwierigkeit der Textfälle und der diagnostischen Treffergenauigkeit und -wahrscheinlichkeit sowie der Triagequalität zu untersuchen, wurde die subjektiv empfundene Schwierigkeit der gestellten Fragen mit den korrekten Antworten auf den Top 1 und Top 3 Diagnosen sowie der Triagierung korreliert (siehe Tabelle 16). Hierbei zeigte sich ein signifikanter Zusammenhang zwischen der subjektiven Schwierigkeit der Textfälle und der korrekten Diagnose hinsichtlich Treffergenauigkeit (Top 1) und Trefferwahrscheinlichkeit (Top 3). Dies bedeutet, dass Studierende, die die Textfälle als leichter einstufen, mehr Fälle korrekt diagnostizierten als Studierende, die die Textfälle als schwieriger einstufen. Selbstverständlich zeigte sich hier auch ein signifikanter Zusammenhang zwischen Treffergenauigkeit (Top 1) und Trefferwahrscheinlichkeit (Top 3). Es zeigte sich kein signifikanter Zusammenhang zwischen den korrekten Diagnosen und der korrekten Triagierung.

Tabelle 16. Zusammenhang korrekter Textfallbearbeitung und subjektiv empfundener Schwierigkeit

	Korrekte Top 1	Korrekte Top 3	Triagierung
Subjektive Schwierigkeit Textfall	<b>-.351**</b>	<b>-.370**</b>	-.163
Korrekte Top 1		<b>.708**</b>	.109
Korrekte Top 3			.039

\*  $p < .05$ , \*\*  $p < .01$

Um den Zusammenhang zwischen der subjektiven Schwierigkeit der Chatfälle und der diagnostischen Treffergenauigkeit und -wahrscheinlichkeit sowie der Triagequalität zu untersuchen, wurde die subjektiv empfundene Schwierigkeit der gestellten Fragen mit den korrekten Antworten auf den Top 1 und Top 3 Diagnosen sowie der Triagierung korreliert (siehe Tabelle 17). Hierbei zeigte sich ein signifikanter Zusammenhang zwischen der subjektiven Schwierigkeit der Chatfälle und der korrekten Diagnose hinsichtlich Treffergenauigkeit (Top 1) und Trefferwahrscheinlichkeit (Top 3). Dies bedeutet, dass Studierende, die die Chatfälle als leichter einstufen, mehr Fälle korrekt diagnostizierten als Studierende, die die Chatfälle als schwieriger einstufen. Selbstverständlich zeigte sich auch hier ein signifikanter Zusammenhang zwischen Treffergenauigkeit (Top 1) und Trefferwahrscheinlichkeit (Top 3). Es zeigte sich kein signifikanter Zusammenhang zwischen der Treffergenauigkeit (Top 1) und der korrekten Triagierung.

Tabelle 17. Zusammenhang korrekter Chatfallbearbeitung und subjektiv empfundener Schwierigkeit

	Korrekte Top 1	Korrekte Top 3	Triagierung
Subjektive Schwierigkeit Textfall	<b>-.270**</b>	<b>-.171*</b>	-.028
Korrekte Top 1		<b>.628**</b>	.161
Korrekte Top 3			<b>.172*</b>

\*  $p < .05$ , \*\*  $p < .01$

Um den Zusammenhang zwischen der Anzahl gestellter Fragen im Chat und der diagnostischen Treffergenauigkeit und -wahrscheinlichkeit sowie der Triagequalität zu untersuchen, wurde die Anzahl gestellter Fragen mit den korrekten Antworten in den Top 1 und Top 3 Diagnosen sowie der Triagierung korreliert (siehe Tabelle 18). Hierbei zeigte sich kein signifikanter Zusammenhang zwischen der Anzahl der gestellten Fragen und der Treffergenauigkeit (Top 1), Trefferwahrscheinlichkeit (Top 3) oder Triagequalität. Es zeigte sich aber ein signifikanter Zusammenhang zwischen der korrekten Diagnose im Rahmen der Top 1 und Top 3. Darüber hinaus zeigte sich ein signifikanter Zusammenhang zwischen der Trefferwahrscheinlichkeit (Top 3) und der korrekten Triagierung.

Tabelle 18. Zusammenhang der Fragenanzahl und korrekten Chatfallbearbeitungen

	Korrekte Top 1	Korrekte Top 3	Triagierung
Anzahl Fragen	.086	.029	-.161
Korrekte Top 1		<b>.628**</b>	.161
Korrekte Top 3			<b>.172*</b>

\*  $p < .05$ , \*\*  $p < .01$

Im Fragebogen gaben die Studierenden im Durchschnitt auf einer Skala von 1 bis 10 (1 = sehr schnell, 10 = sehr langsam) eine Tippgeschwindigkeit von 4.63 ( $SD = 2.01$ ) an. Die subjektive Einschränkung aufgrund der individuellen Tippfähigkeiten wurde von den Studierenden im Durchschnitt auf einer Skala von 1 bis 10 (1 = überhaupt nicht eingeschränkt, 10 = sehr stark eingeschränkt) mit 3.44 ( $SD = 2.19$ ) angegeben.

### 4.3. Vergleich Symptomchecker- und Studierenden-Assessments

Insgesamt wurde von den Symptomcheckern die korrekte Diagnose in 38.02% an erster Stelle genannt (Top 1). Die Studierenden nannten die korrekte Diagnose in 64.10% an erster Stelle, siehe Tabelle 19. Dieser Unterschied war signifikant,  $p = .001$ . Zusätzlich unterschieden sich die Symptomchecker und die Studierenden signifikant in der diagnostischen Treffergenauigkeit (Top 1) in den Dringlichkeitsstufen Emergent Care,  $p = .001$ , Non-emergent Care,  $p = .001$  und Self Care,  $p = .001$ .

Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose von den Symptomcheckern in 56.44% der Fälle gelistet. Von den Studierenden wurde innerhalb der ersten drei gegebenen Diagnosen die korrekte Diagnose in 80.19% der Fälle gelistet. Dieser Unterschied war ebenfalls signifikant,  $p = .001$ . Zusätzlich unterschieden sich die Symptomchecker und die Studierenden signifikant in der diagnostischen Trefferwahrscheinlichkeit (Top 3) in den Dringlichkeitsstufen Emergent Care,  $p = .002$ , Non-emergent Care,  $p = .001$  und Self Care,  $p = .001$ .

Die richtige Triagierung wurde von den Symptomcheckern in 59.19% der Fallbearbeitungen und von den Studierenden in 59.91% korrekt genannt. Dabei unterschied sich die Triagequalität der Symptomchecker nur innerhalb der Dringlichkeitsstufe Emergent Care signifikant ( $p = .001$ ) zugunsten der Studierenden-Assessments (siehe (EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)).

Tabelle 19. Vergleich Symptomchecker- und Studierenden-Assessments

Vergleich Assessments Symptomchecker gesamt und Studierende							
Vignetten	Symptomchecker gesamt			Studierende (Textbearbeitungen)		P-Wert	Chi-Quadrat
	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)		
<b>Top 1</b>							
Alle	49 (100.00)	254/668	38.02 (34.73–47.54)	275/429	64.10 (59.56–68.63)	.001	71.16
EC	17 (34.69)	86/236	36.44 (30.49–42.49)	87/143	60.84 (52.70–68.89)	.001	21.37
NE	16 (32.65)	83/223	37.22 (30.90–43.43)	98/143	68.53 (60.67–76.03)	.001	34.17
SC	16 (32.65)	85/209	40.67 (33.81–47.39)	90/143	62.94 (55.56–71.31)	.001	16.84
<b>Top 3</b>							
Alle	49 (100.00)	377/668	56.44 (52.45–60.19)	344/429	80.19 (76.46–83.59)	.001	65.41
EC	17 (34.69)	123/236	52.11 (45.89–58.52)	98/143	68.53 (60.67–76.03)	.002	9.87
NE	16 (32.65)	123/223	55.16 (48.28–61.70)	124/143	86.71 (80.77–92.31)	.001	39.54
SC	16 (32.65)	131/209	62.68 (56.31–69.16)	122/143	85.31 (79.22–91.25)	.001	21.52
<b>Triage</b>							
Alle	49 (100.00)	264/446	59.19 (54.67–63.68)	257/429	59.91 (54.99–64.46)	.830	0.05
EC	17 (34.69)	116/162	71.60 (63.74–78.57)	124/143	86.71 (80.77–92.31)	.001	10.34
NE	16 (32.65)	98/143	68.53 (60.88–76.16)	87/143	60.84 (52.70–68.89)	.174	1.85
SC	16 (32.65)	50/141	35.46 (27.50–43.62)	46/143	32.17 (24.63–40.00)	.557	0.34

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

#### 4.3.1. Deutschsprachige Symptomchecker- und Studierenden-Assessments

Insgesamt wurde von den deutschsprachigen Symptomcheckern die korrekte Diagnose in 48.09% an erster Stelle genannt (Top 1). Die Studierenden nannten die korrekte Diagnose in 64.10% an erster Stelle, siehe Tabelle 20. Dieser Unterschied war signifikant,  $p = .001$ .

Dabei unterschied sich die Treffergenauigkeit (Top 1) nur innerhalb der Dringlichkeitsstufe Non-emergent Care signifikant,  $p = .001$ .

Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose von den deutschsprachigen Symptomcheckern in 69.23% der Fälle gelistet. Von den Studierenden wurde innerhalb der ersten drei gegebenen Diagnosen die korrekte Diagnose in 80.19% der Fälle genannt. Dieser Unterschied war ebenfalls signifikant,  $p = .003$ . Die Trefferwahrscheinlichkeit unterschied sich auch hier nur innerhalb der Dringlichkeitsstufe Non-emergent Care signifikant,  $p = .001$ .

Die richtige Triageurteilung wurde von den deutschsprachigen Symptomcheckern in 65.22% der Fallbearbeitungen und von den Studierenden in 59.91% korrekt genannt. Dabei unterschied sich die Triagequalität nur innerhalb der Dringlichkeitsstufe Emergent Care signifikant,  $p = .011$  (siehe Tabelle 20).

Tabelle 20. Vergleich deutschsprachiger Symptomchecker- und Studierenden-Assessments

Vergleich deutschsprachiger Symptomchecker- und Studierenden-Assessments							
Vignetten	deutschsprachige Symptomchecker			Studierende (Textbearbeitungen)		P-Wert	Chi-Quadrat
	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)		
<b>Top 1</b>							
Alle	49 (100.00)	89/182	48.90 (42.11–56.34)	275/429	64.10 (59.56–68.63)	.001	12.26
EC	17 (34.69)	31/65	47.69 (35.48–58.93)	87/143	60.84 (52.70–68.89)	.076	3.15
NE	16 (32.65)	26/59	44.07 (31.38–57.62)	98/143	68.53 (60.67–76.03)	.001	10.55
SC	16 (32.65)	32/58	55.17 (42.00–68.75)	90/143	62.94 (55.56–71.31)	.307	1.04
<b>Top 3</b>							
Alle	49 (100.00)	126/182	69.23 (62.22–75.29)	344/429	80.19 (76.46–83.59)	.003	8.64
EC	17 (34.69)	44/65	67.69 (55.56–78.69)	98/143	68.53 (60.67–76.03)	.904	0.02
NE	16 (32.65)	37/59	62.71 (50.00–74.64)	124/143	86.71 (80.77–92.31)	.001	14.87
SC	16 (32.65)	45/58	77.59 (66.68–88.46)	122/143	85.31 (79.22–91.25)	.185	1.75
<b>Triage</b>							
Alle	49 (100.00)	90/138	65.22 (57.50–72.81)	257/429	59.91 (54.99–64.46)	.265	1.24
EC	17 (34.69)	37/52	71.15 (57.50–83.63)	124/143	86.71 (80.77–92.31)	.011	6.41
NE	16 (32.65)	33/44	75.00 (61.54–87.17)	87/143	60.84 (52.70–68.89)	.087	2.94
SC	16 (32.65)	20/42	47.61 (32.56–62.50)	46/143	32.17 (24.63–40.00)	.066	3.38

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

#### 4.3.2. Englischsprachige Symptomchecker- und Studierenden-Assessments

Insgesamt wurde von den englischsprachigen Symptomcheckern die korrekte Diagnose in 33.95% an erster Stelle genannt (Top 1). Die Studierenden nannten die korrekte Diagnose in

64.10% an erster Stelle, siehe Tabelle 21. Dieser Unterschied war signifikant,  $p = .001$ . Zusätzlich unterschieden sich die Assessments der Symptomchecker und der Studierenden signifikant auf allen Dringlichkeitsstufen (Emergent Care,  $p = .001$ , Non-emergent Care,  $p = .001$  und Self Care,  $p = .001$ ).

Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose von den englischsprachigen Symptomcheckern in 51.65% der Fälle gelistet. Von den Studierenden wurde innerhalb der ersten drei gegebenen Diagnosen die korrekte Diagnose in 80.19% der Fälle gelistet. Dieser Unterschied war ebenfalls signifikant,  $p = .001$ . Zusätzlich zeigten sich die signifikanten Differenzen auf allen drei Dringlichkeitsstufen (Emergent Care,  $p = .001$ , Non-emergent Care,  $p = .001$  und Self Care,  $p = .001$ ).

Die richtige Triagierung wurde von den Symptomcheckern in 56.49% der Fallbearbeitungen und von den Studierenden in 59.91% korrekt genannt. Dabei war die korrekte Empfehlung der Symptomchecker nur auf der Dringlichkeitsstufe Emergent Care signifikant ( $p = .003$ ) überlegen (siehe Tabelle 21).

Tabelle 21. Vergleich englischsprachiger Symptomchecker- und Studierenden-Assessments

Vergleich englischsprachiger Symptomchecker- und Studierenden-Assessments							
Vignetten	englischsprachige Symptomchecker			Studierende (Textbearbeitungen)		P-Wert	Chi-Quadrat
	Anzahl an Vignetten (%)	Verhältnis	% (95% KI)	Verhältnis	% (95% KI)		
<b>Top 1</b>							
Alle	49 (100.00)	165/486	33.95 (29.73–38.21)	275/429	64.10 (59.56–68.63)	.001	82.99
EC	17 (34.69)	55/171	32.16 (25.53–39.18)	87/143	60.84 (52.70–68.89)	.001	25.85
NE	16 (32.65)	57/164	34.76 (27.41–42.26)	98/143	68.53 (60.67–76.03)	.001	34.86
SC	16 (32.65)	53/151	35.10 (27.97–43.12)	90/143	62.94 (55.56–71.31)	.001	22.78
<b>Top 3</b>							
Alle	49 (100.00)	251/486	51.65 (47.24–56.03)	344/429	80.19 (76.46–83.59)	.001	81.62
EC	17 (34.69)	79/171	46.20 (38.61–53.45)	98/143	68.53 (60.67–76.03)	.001	15.79
NE	16 (32.65)	86/164	52.44 (44.51–60.37)	124/143	86.71 (80.77–92.31)	.001	41.52
SC	16 (32.65)	86/151	56.95 (48.63–64.33)	122/143	85.31 (79.22–91.25)	.001	28.55
<b>Triage</b>							
Alle	49 (100.00)	174/308	56.49 (50.66–62.26)	257/429	59.91 (54.99–64.46)	.354	0.86
EC	17 (34.69)	79/110	71.82 (63.30–79.44)	124/143	86.71 (80.77–92.31)	.003	8.70
NE	16 (32.65)	65/99	65.66 (55.86–75.27)	87/143	60.84 (52.70–68.89)	.446	0.58
SC	16 (32.65)	30/99	30.30 (21.78–39.80)	46/143	32.17 (24.63–40.00)	.759	0.09

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

#### 4.4. Zusammenfassung der Ergebnisse

Insgesamt zeigt sich die Leistungsfähigkeit der Symptomchecker anhand der drei Parameter Treffergenauigkeit, Trefferwahrscheinlichkeit und Korrektheit der Triage. Die korrekte Diagnose wurde in 38.02% der Fälle an erster Stelle (Top 1) im Sinne der Treffergenauigkeit genannt. Innerhalb der ersten drei gegebenen Diagnosen (Top 3) wurde die korrekte Diagnose in 56.44% der Fälle im Sinne der Trefferwahrscheinlichkeit gelistet. Die Diagnoseleistung unterschied sich für die verschiedenen Dringlichkeitsstufen nicht signifikant, siehe Tabelle 22. Die korrekte Triagierung wurde in 59.19% der Fallbearbeitungen genannt. Dabei unterschied sich die Triagequalität der Symptomchecker innerhalb der verschiedenen Dringlichkeitsstufen signifikant,  $p < .001$ .

Tabelle 22. Zusammenfassung der Leistung der Symptomchecker

Symptomchecker gesamt		
	% (95% KI)	Signifikanz
<b>Top 1</b>	<b>38.02</b> (34.73–47.54)	
Art der Vignetten		
EC	36.44 (30.49–42.49)	
NE	37.22 (30.90–43.43)	nicht signifikant
SC	40.67 (33.81–47.39)	
<b>Top 3</b>	<b>56.44</b> (52.45–60.19)	
Art der Vignetten		
EC	52.11 (45.89–58.52)	
NE	55.16 (48.28–61.70)	nicht signifikant
SC	62.68 (56.31–69.16)	
<b>Triage</b>	<b>59.19</b> (54.67–63.68)	
Art der Vignetten		
EC	71.60 (63.74–78.57)	
NE	68.53 (60.88–76.16)	signifikant
SC	35.46 (27.50–43.62)	

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)

Die Treffergenauigkeit (Top 1) der deutschsprachigen Symptomchecker überstieg die der englischsprachigen Symptomchecker signifikant, siehe Tabelle 23. Dies galt auch für die Trefferwahrscheinlichkeit (Top 3).

Tabelle 23. Zusammenfassung des Leistungsvergleichs deutsch- und englischsprachiger Symptomchecker

Vergleich deutsch- und englischsprachiger Symptomchecker			
	deutschsprachig	englischsprachig	
	% (95% KI)	% (95% KI)	Signifikanz
Top 1	48.90 (42.11–56.34)	33.95 (29.73–38.21)	signifikant
Top 3	69.23 (62.22–75.29)	51.65 (47.24–56.03)	signifikant
Triage	65.22 (57.50–72.81)	56.49 (50.66–62.26)	nicht signifikant

Die Leistung der Symptomchecker lässt sich auch im Vergleich zur Leistung der Studierenden beurteilen. Die diagnostische Treffergenauigkeit (Top 1) der Symptomchecker war signifikant niedriger als in den Studierendenassessments, siehe Tabelle 24. Dies galt ebenfalls für die Trefferwahrscheinlichkeit (Top 3). Der Anteil korrekter Triagierungen unterschied sich zwischen den Symptomcheckern und Studierenden nicht signifikant. Dies war nur in der Dringlichkeitsstufe Emergent Care der Fall.

Tabelle 24. Zusammenfassung Vergleich Symptomchecker- und Studierenden-Assessments

Vergleich Assessments Symptomchecker gesamt und Studierende			
	Symptomchecker gesamt	Studierende (Textbearbeitungen)	
Vignetten	% (95% KI)	% (95% KI)	Signifikanz
Top 1			
Alle	38.02 (34.73–47.54)	64.10 (59.56–68.63)	signifikant
EC	36.44 (30.49–42.49)	60.84 (52.70–68.89)	signifikant
NE	37.22 (30.90–43.43)	68.53 (60.67–76.03)	signifikant
SC	40.67 (33.81–47.39)	62.94 (55.56–71.31)	signifikant
Top 3			
Alle	56.44 (52.45–60.19)	80.19 (76.46–83.59)	signifikant
EC	52.11 (45.89–58.52)	68.53 (60.67–76.03)	signifikant
NE	55.16 (48.28–61.70)	86.71 (80.77–92.31)	signifikant
SC	62.68 (56.31–69.16)	85.31 (79.22–91.25)	signifikant
Triage			
Alle	59.19 (54.67–63.68)	59.91 (54.99–64.46)	nicht signifikant
EC	71.60 (63.74–78.57)	86.71 (80.77–92.31)	signifikant
NE	68.53 (60.88–76.16)	60.84 (52.70–68.89)	nicht signifikant
SC	35.46 (27.50–43.62)	32.17 (24.63–40.00)	nicht signifikant

(EC=Emergent Care, NE=Non-emergent Care, SC=Self Care)



## 5. Diskussion

### 5.1. Wissenschaftskontext der Studie

Die Bedeutung der Diagnose hat sich in der Entwicklung der Medizin gewandelt. Über ihre ursprünglichen klassifikatorischen Funktionen hinaus tragen Diagnosen heute maßgeblich zum ärztlichen Handeln und der klinischen Patientenversorgung bei. Daneben lassen sich sowohl sozialrechtliche als auch soziologische Funktionen identifizieren, sodass Diagnosen heutzutage sowohl für Ärzte als auch Patienten elementar sind.<sup>1,21</sup>

Klassischerweise erfolgt die Diagnosestellung durch den Arzt, welcher insbesondere durch diese Kompetenz seinen professionellen Status erhält. Umgekehrt verleiht die Berechtigung zur Diagnosestellung dem Arzt Autorität und unterstreicht seine Stellung in der Gesundheitsversorgung<sup>1,5</sup>. Diese Stellung, besonders im Verhältnis zum Patienten, hat sich in den letzten Jahrzehnten jedoch stetig verändert. Die Arzt-Patienten-Beziehung hat sich vom paternalistischen, bevormundenden Modell hin zu einer partizipativen Arzt-Patienten-Beziehung auf Augenhöhe entwickelt, in der das *shared decision making* dem Patienten mehr Mitbestimmung gewährt<sup>33,34,37</sup>.

Diese Entwicklung beruht auch darauf, dass der Patient durch einen einfacheren und umfassenderen Zugang zu medizinischen Informationen besser informiert ist<sup>63</sup>. Mit der Digitalisierung hat sich die Quelle solcher medizinischer Informationen für Patienten zunehmend ins Internet verlagert. Besonders die jüngeren Generationen, wie die Generation Y, nutzen vornehmlich das Internet zur Informationsbeschaffung.<sup>40,42,64</sup>

Neue digitale Gesundheitsprogramme tragen diesem Trend Rechnung. Die Symptomchecker vereinfachen die komplexe Recherche im Internet und zeigen den Anspruch, Patienten bei der Selbstdiagnose zu helfen. Dazu streben sie die Kondensation von Symptomkonstellationen hin zu Differentialdiagnosen an und/oder geben eine Triageempfehlung ab.

Solche Symptomchecker haben sich zum jetzigen Zeitpunkt bereits etabliert, ihre Qualität bleibt allerdings weitestgehend unklar<sup>56,60</sup>. Um sowohl Risiken als auch Chancen der Symptomchecker zum gegebenen Zeitpunkt abschätzen zu können, bedarf es einer differenziellen Untersuchung ihrer Leistungsfähigkeit. Erst so lassen sich mögliche Anwendungsszenarien evaluieren. An die Symptomchecker wird die Erwartung gerichtet, dass sie Patienten im digitalen Zeitalter den Zugang zu medizinischen Informationen erleichtern, sich dabei in das bestehende Gefüge der Medizin einfügen und keine zusätzlichen Probleme kreieren.

Mögliche Vorteile könnten beispielsweise eine Entlastung des Gesundheitssystems in Form von Kosten- und Ressourceneinsparungen bei nicht behandlungsbedürftigen Beschwerden oder eine Unterstützung der Versorgung in ländlichen Gebieten sein<sup>56,61</sup>. Zudem können

unnötige Warte- und Fahrtzeiten der Patienten vermieden werden. Fehleinschätzungen und Fehldiagnosen durch die Symptomchecker können jedoch auch zu einer Überlastung von Notaufnahmen führen und bergen Risiken, wie das Nichterkennen ernster Krankheitszustände und Notfälle<sup>65</sup>.

Diesen Chancen und Risiken medizinischer Diagnoseprogramme versucht die vorliegende Studie eine empirische Beurteilungsgrundlage zu geben.

## **5.2. Diskussion der Ergebnisse**

### **5.2.1. Die diagnostische Leistungsfähigkeit der Symptomchecker**

Anhand von 49 standardisierten klinischen Fallvignetten wurde eine möglichst realistische Nutzung der frei zugänglichen und marktgängigen Symptomchecker realisiert. Dabei wurden die Assessments von 16 auf dem Markt befindlichen Symptomcheckern untersucht (4 deutschsprachige, 12 englischsprachige in Übersetzung)

Die präzise Diagnose (Treffergenauigkeit) wurde von den Symptomcheckern in 38.02% genannt. Innerhalb der ersten drei Verdachtsdiagnosen, also den Top 3, wurden die Fallvignetten in 56.44% noch korrekt diagnostiziert (Trefferwahrscheinlichkeit). Dabei ist zu diskutieren, welches der beiden Ergebnisse die Leistung der Symptomchecker besser beschreibt. Zum einen ist die erstgenannte Diagnose die wahrscheinlichste und dürfte auch vom Patienten als solche aufgenommen werden. Sie wird für ihn in den meisten Fällen die vorübergehende Erklärung für sein Leiden sein und die nächsten Handlungen, wie eine weitere Recherche oder einen Arztbesuch, bestimmen. Zum anderen werden auch im Diagnoseprozess eines Arztes verschiedene Hypothesen aufgestellt, die eine weitere Klärung durch Untersuchungen erforderlich machen.<sup>12</sup> Im Sinne von Differentialdiagnosen, die weiterhin in Betracht gezogen werden müssen, haben somit auch die weniger wahrscheinlichen Diagnosen ihre Bedeutung. Trotzdem entspricht auch ein Ergebnis von 56.44% richtiger Diagnosen in den Top 3 einer Fehldiagnoserate von über 43%, die Rate bei der Treffergenauigkeit, also bei den präzisen (Top 1-) Diagnosen liegt noch darunter, sodass die Fehlerquote hier auf ca. 62% steigt.

Betrachtet man also das gesamte Kollektiv an getesteten Symptomcheckern, spricht dieses Ergebnis für eine schlechte Leistung, da ihre Fehldiagnoserate ungleich höher ist als in der klinischen Realität. Denn die Häufigkeit von Diagnosefehlern ist nach Elstein in der klinischen Medizin auf nur ungefähr 15% zu beziffern.<sup>66</sup> Diese Schätzung halten auch Berner und Graber für zutreffend.<sup>67</sup> Nach ihnen lassen sich darüber hinaus sehr unterschiedliche Angaben zu Fehldiagnosen machen. Je nach Fachbereich und Studienfokus variiert deren Häufigkeit zwischen niedrigen Prozentzahlen von 0,6% bis 12% in Notaufnahmesituationen und bis zu 45% für einzelne Krankheitsbilder wie der Lungenembolie oder 18,6% für die Appendizitis.<sup>67-69</sup> Eine Untersuchung zu Fehldiagnosen im ambulanten Behandlungssetting

in den USA ermittelte auf der Basis dreier großer Beobachtungsstudien an Erwachsenen eine Häufigkeit von nur 5.08%.<sup>70</sup>

Als Grund für diese große Varianz nicht korrekter Leitdiagnosen im klinischen Feld sind vor allem die Unterschiede in den Definitionen der Fehldiagnose und ihrer Untersuchungsmethoden zu nennen.<sup>71,72</sup> Die Begriffe der Fehldiagnose und des diagnostischen Fehlers werden teilweise synonym verwendet und variieren je nach Studienausrichtung.<sup>71</sup> Nach Kirch und Engwicht geht der Begriff Fehldiagnose vor allem darin auf, dass die Grundlagen therapeutischen Handelns und prognostischer Überlegungen falsch gewählt werden.<sup>65</sup>

Andere Autoren sprechen im Kontext des diagnostischen Fehlers lediglich von einer Diagnose, die verzögert, falsch oder nie gestellt wurde.<sup>73,74</sup> Damit liegen sie in Übereinstimmung mit der National Academy of Medicine, die einen diagnostischen Fehler als das Versäumnis definiert, eine rechtzeitige und genaue Erklärung für die Gesundheitsprobleme des Patienten zu finden.<sup>12</sup>

Die Herangehensweisen zur Erhebung diagnostischer Fehler in der Praxis umfassen Autopsiestudien, Patienten- und Versorgerumfragen, Zweitbeurteilungen in der visuellen Diagnostik (z.B. Dermatologie, Pathologie, Radiologie), Überprüfungen diagnostischer Testverfahren, Analysen von Arzthaftungsfällen, Fallreviews, freiwillige Fehlermeldungen und Studien mit standardisierten Patienten.<sup>72</sup> Davon kommt die Verwendung von standardisierten Patienten in realen Behandlungssituationen unserem Studiendesign mit standardisierten Fallvignetten am nächsten und kann zudem belastbare Zahlen generieren.<sup>72</sup> In einer Studie, die diese Methode anwandte, konnte festgestellt werden, dass es bei der Präsentation typischer internistischer Leiden, wie COPD, Diabetes Mellitus und weiteren Krankheitsbildern in 13% der Fälle zu einer Fehldiagnose aufgrund eines diagnostischen Fehlers kam.<sup>75</sup>

Es wird in unserer Studie deutlich, dass die Fehldiagnoserate der Symptomchecker die genannten Häufigkeiten trotz ihrer großen Diskrepanzen weit übersteigt. Dies gilt nicht nur für die Leistung aller Symptomchecker mit ungefähr 62% Fehldiagnosen, sondern auch für die Symptomchecker, die die besten Ergebnisse zeigten. Dazu gehört die Ada-App mit ca. 40% Fehldiagnosen sowie der Symptomchecker „Symptomate“ mit ca. 46% Fehldiagnosen. Die Leistung der Symptomchecker erscheint insgesamt zu schlecht, als dass sie mit der Diagnosefindung in der Praxis oder Klinik mithalten könnte. Zu einem vergleichbaren Ergebnis kam die Arbeitsgruppe um Semigran et al., auf deren Arbeit unsere Studie aufbaut.<sup>56</sup> Die Autoren verwendeten 45 standardisierte klinische Fallvignetten, die in größtenteils identischer Form auch in der vorliegenden Untersuchung verwendet wurden. Auch hier waren die Fallvignetten für eine einfachere Eingabe auf die wichtigsten Symptome kondensiert und zudem mit einer feststehenden korrekten Diagnose verknüpft. Anschließend

wurden die Fallvignetten in 23 Symptomchecker vom englischsprachigen Markt eingegeben und die Verdachtsdiagnosen und Dringlichkeitseinstufungen dokumentiert. Die Autoren fanden dabei heraus, dass in 34% die korrekte Diagnose an erster Stelle genannt wurde, die Fehlerquote also mit 66% ähnlich hoch war wie in der vorliegenden Studie mit 62% Fehldiagnosen.

In 51% befand sich die korrekte Diagnose unter den ersten drei gegebenen Verdachtsdiagnosen, vergleichbar mit einer Trefferwahrscheinlichkeit von 56% in unserer Studie. Die leicht erhöhte Trefferwahrscheinlichkeit in unserer Untersuchung könnte dadurch zustande gekommen sein, dass, so der an uns kommunizierte Hinweis der Harvard-Autoren, möglicherweise nach der Veröffentlichung der Vorstudie eine Implementierung der Fallvignetten in einzelne Symptomchecker erfolgte.

Es ist auch zu berücksichtigen, dass in unserer Studie erstmals deutschsprachige Symptomchecker systematisch getestet wurden. Betrachtet man ausschließlich das Kollektiv dieser deutschsprachigen Angebote, ist festzustellen, dass sie eine signifikant höhere Rate an korrekten Diagnosen nannten als die englischsprachigen Symptomchecker. Dies gilt sowohl für die Betrachtung der Trefferwahrscheinlichkeit (69,23% gegenüber 51,65%), als auch der Treffergenauigkeit (48,90% gegenüber 33,95%). Dies spricht zum einen für den modernen Entwicklungsstand der deutschsprachigen Symptomchecker, die möglicherweise fortschrittlichere Algorithmen verwenden. Zum anderen dürfte die geringe Anzahl von lediglich 4 auf dem deutschen Markt befindlichen Symptomcheckern dazu führen, dass durch das starke Abschneiden der Symptomchecker Ada (59% korrekte Top 1 und 82% korrekte Top 3) und Symptomate (54% korrekte Top 1 und 69% korrekte Top 3) die Leistungsfähigkeit des Gesamtangebots angehoben wird.

Generell ist zu beachten, dass die Leistungen der Symptomchecker einer starken Varianz unterliegen (korrekte Top 1 zwischen 9 und 59%), was Semigran et al. ebenfalls beobachteten.<sup>56</sup> Dies ist durch die Diversität der Symptomchecker zu erklären. Sie unterscheiden sich stark in ihrem Aufbau, in ihrem Vorgehen zur Sammlung von Symptomen und Informationen sowie ihrer technischen Ausführung. Es ist davon auszugehen, dass Symptomchecker von Herausgebern mit größeren finanziellen Mitteln in Zukunft bessere Ergebnisse erzielen werden als Angebote, die mit geringen finanziellen Mitteln die Marktreife erlangen sollen.

Die hohe Rate an Fehldiagnosen, die bei den gegenwärtig auf dem Markt befindlichen Symptomcheckern festzustellen ist, lässt annehmen, dass sie bestehende Probleme der Diagnosestellung aggravieren könnten.

Zum einen haben Fehldiagnosen allgemein das Potential, den betroffenen Patienten erheblichen körperlichen Schaden zuzufügen.<sup>76-79</sup> Dieser Aspekt ist mit dem Aufkommen der Patientensicherheitsbewegung vor ungefähr zwei Jahrzehnten in den Brennpunkt der

Aufmerksamkeit gelangt.<sup>14</sup> Das Bestreben, die Sicherheit der Patienten zu erhöhen, beschränkte sich anfangs auf Medikamentensicherheit und weitere einfach zu behebbende Probleme, rückte jedoch zunehmend auch Fehler der Diagnosestellung in den Fokus.<sup>80</sup> In diesem Zuge wurden im Jahr 2008 Fehldiagnosen auch vom Programm für Patientensicherheit der WHO zum Forschungsschwerpunkt erklärt.<sup>81</sup> Ihre Reduktion lässt sich auch in Deutschland als wichtiger Baustein zum Erreichen einer höheren Patientensicherheit identifizieren.<sup>82</sup>

Zum anderen verursachen Fehldiagnosen neben körperlichen Schäden auch beträchtliche Kosten. In den USA wurde in einer retrospektiven Untersuchung berechnet, dass im Zeitraum von 1986 bis 2010 Diagnose-assoziierte Fehlbehandlungen zu einer Auszahlung von Entschädigungen in einer Höhe von 38.8 Mrd. US-Dollar führten.<sup>83</sup> In einer Datenbank (1985-2007) der Physicians Insurers Association of America machten Diagnosefehler mit 37% den größten Anteil aller Schadensersatzansprüche in Notaufnahmen aus.<sup>78</sup> In Deutschland waren im Jahr 2018 ebenfalls über ein Fünftel, knapp 22,72%, der 10.647 vermuteten Arzthaftungsfälle (die durch die Bundesärztekammer von ihren Gutachterkommissionen und Schlichtungsstellen registriert wurden) Diagnose-assoziiert.<sup>84</sup>

Die Verminderung von Fehldiagnosen könnte folglich zu Einsparungen von Kosten führen, was in den Gesundheitssystemen vieler Länder der Welt ein wesentliches Ziel darstellt.<sup>36</sup> Nach Berner & Graber (2008) ließe sich eine solche Reduktion von Fehldiagnosen auch mittels höherer diagnostischer Kompetenz erreichen.<sup>67</sup> Diese Fähigkeit lässt sich aus der Diagnoseleistung der Symptomchecker in unserer Studie nicht ableiten.

### **5.2.2. Die präklinische Triagefähigkeit der Symptomchecker**

Um den Nutzen der Symptomchecker in der Versorgung bewerten zu können, wurde ebenfalls ihre Triagefähigkeit anhand der standardisierten Fallvignetten geprüft. Jeder der Fallvignetten war eine der drei Dringlichkeitsstufen *Emergent Care* (EC), *Non-emergent Care* (NE) oder *Self Care* (SC) zugeteilt. Die Empfehlungen der Symptomchecker wurden der höchsten Dringlichkeit *Emergent Care* bei Aufforderung zur unverzüglichen Vorstellung beim Arzt, im Krankenhaus oder zur Verständigung des Rettungsdienstes zugeordnet. Bei einer Aufforderung zur Konsultation eines Arztes, die nicht unverzüglich stattfinden muss, erfolgte eine Zuteilung zu *Non-emergent Care* und im Falle einer Empfehlung zur Selbsthilfe oder zum Abwarten zur Kategorie *Self Care*. Diese Aufteilung findet auch in der Praxis Anwendung und erfolgte, um die diversen Triagestufen der Symptomchecker untereinander vergleichbar zu machen.<sup>28</sup>

Insgesamt triagierten die Symptomchecker in 59.19% richtig. Dies bedeutet, dass 6 von 10 Patienten direkt auf die passende Anlaufstelle verwiesen, beziehungsweise zur Selbstbehandlung aufgefordert wurden. Eine äußerst ähnliche Triagequalität durch Symptomchecker zeigte sich in der Studie von Semigran et al. mit 57% korrekt triagierten

Fallvignetten<sup>56</sup>. Vergleichswerte zur Triage im Gesundheitssektor lassen sich in der Literatur nicht finden, da die Triage ein Bestandteil der Versorgung ist, der im Regelfall direkt in der Arztpraxis oder Klinik stattfindet.

Die Handlungsempfehlungen der Symptomchecker imponieren als eine Form der präklinischen Triage, da sie vor dem persönlichen Kontakt in der Klinik oder Arztpraxis stattfinden. Diese Art der Triage ermöglicht es, den richtigen Versorgungsbedarf des Patienten im Sinne einer passenden medizinischen Einrichtung einzuschätzen und darüber hinaus die zeitliche Dringlichkeit einer Intervention zu erkennen. Grundsätzlich können Symptomchecker den Patienten, die keine ärztliche Behandlung benötigen, empfehlen zu Hause zu bleiben. Dadurch könnten Ressourcen gespart und mögliche Überbehandlungen vermieden werden.<sup>56</sup>

Demgegenüber stehen deutliche Risiken, wenn eine präklinische Triage qualitativ unzureichend ist: Fehleinschätzungen können Notdienste zusätzlich belasten, indem risikoscheue Symptomchecker Patienten ohne tatsächliche Indikation dem Rettungsdienst oder den Notaufnahmen zuführen. Dies bedeutete einen erhöhten Zeitaufwand für sowohl Arzt als auch Patient und konsekutiv hohe Kosten für Klinik und Gesellschaft.<sup>56</sup>

Die möglichen Vorteile präklinischer Triagierung haben zu diversen Ansätzen geführt, den Versorgungsbedarf der Patienten schon vor dem Eintreffen in der Klinik zu steuern. Hier ist die Telefontrriage das augenscheinlich prominenteste Anwendungsbeispiel für eine organisierte präklinische Triage.<sup>28,29</sup> Sie wird neben der Einrichtung hausärztlicher Notfallpraxen in der Nähe von ZNAs als probates Mittel zur Steuerung von dringlichen und nichtdringlichen Fällen gesehen.<sup>31,85</sup>

Die Telefontrriage fällt unter den Oberbegriff der Telekonsultation, die auch die telemedizinische Beratung in Akutsituationen umfassen kann.<sup>32</sup> Bei einer solchen Telefontrriage wird eine Ersteinschätzung des Gesundheitszustands des Patienten anhand einer telefonischen Anamnese erhoben. Diese entscheidet dann darüber, wie mit dem Patienten weiter verfahren wird, ob und wann er persönlich vorstellig werden sollte oder ob er eine medizinische Notfallversorgung benötigt. Gegebenenfalls kann auch hier eine Beratung zur Selbstbehandlung erfolgen.<sup>32</sup> Damit ähnelt die Telefontrriage in ihrer Funktion den Symptomcheckern. Das Potenzial der Symptomchecker lässt sich demnach im Vergleich zur Telefontrriage messen, deren Nutzung bereits verbreitet ist.

Telefontrriagesysteme kommen seit längerem vor allem in den USA, Großbritannien, den Niederlanden, oder der Schweiz zum Einsatz und ihre Einführung wird vor dem Hintergrund steigender Belastung der Notaufnahmen auch in Deutschland diskutiert.<sup>31,32</sup> Denn ein Einsatz der Telefontrriage könnte sowohl die Allgemeinarztpraxen als auch Notaufnahmen entlasten.<sup>28</sup> Untersuchungen weisen darauf hin, dass sich ein Großteil der Patienten in den

ZNAs ambulant behandeln ließen und in einer Triage nicht als Notfall zu definieren wären.<sup>31,86</sup>

Obwohl die Telefontragesysteme mancherorts bereits etabliert sind, lassen sich allgemeingültige Aussagen über ihre Güte nur schwer treffen.<sup>28,87</sup> Die Teilaspekte Sicherheit, Angemessenheit und klinisches Outcome stehen bei einem Vergleich zu den Symptomcheckern im Vordergrund. Diese drei finden sich auch unter den neun Teilaspekten, die von Lake et al. in ihrem Überblick mehrerer Reviews zur Qualität, Sicherheit und Verwaltung von Triagediensten untersucht wurden.<sup>28</sup>

Bezüglich der Sicherheit resümieren die Autoren, dass die Telefontriage in den meisten Fällen sicher ist<sup>28</sup>. Angaben zu falsch niedrigen Triagierungen mit einem gesundheitlichen Schadenspotenzial wurden von Blank et al. in zwei von ihnen untersuchten Studien gefunden und auf 1.3%, beziehungsweise 3.2% beziffert.<sup>88</sup> Carrasqueiro et al. nennen eine Häufigkeit von 4% bis 10% für Fälle, in denen eine notwendige Versorgung nicht empfohlen wurde.<sup>89</sup> Im Umkehrschluss wird im Review von Huibers et al. ein Anteil von 97% sicheren Konsultationen außerhalb regulärer Sprechzeiten angeführt. Als Beispiel einer flächendeckenden Implementierung lässt sich die in der Schweiz seit langem etablierte überregionale Telefontriage nennen.<sup>90</sup> Dort ist die Telemedizin seit vielen Jahren selbstverständlicher Bestandteil des Gesundheitssystems und nimmt für viele Patienten heute die Funktion eines Hausarztes ein.<sup>91</sup> Auch hier konnten kritische Situationen mit möglicher Gesundheitsgefährdung nur in unter 5% der Fälle festgestellt werden.<sup>92</sup> Von McKinstry et al. wird jedoch angeführt, dass es in Telefonkonsultationen weniger wahrscheinlich als bei normalen Arztkontakten ist, dass ausreichend Informationen zusammengetragen wurden, um wichtige gefährliche Differentialdiagnosen ausschließen zu können. Dies führen die Autoren auf kürzere Konsultationszeiten sowie weniger Nachfragen, Beratung und Beziehungsaufbau als in persönlichen Gesprächen zurück.<sup>93</sup> Mehr als bei der Sicherheit zeigen Angaben zur Angemessenheit der gegebenen Triage eine beträchtliche Varianz. Im diesbezüglichen Review von Blank et al. variiert diese zwischen 44 und 98% bei einem Median von 75% verhältnismäßiger Triageentscheidungen.<sup>88</sup> Ebenso ist das klinische Outcome ein Teilaspekt, der sich bisher nicht ausreichend evaluieren lässt, was unter anderem durch mangelnde Forschung, kleine Stichproben oder schwache Ergebnisse bedingt ist.<sup>28,93,94</sup>

Die unterschiedlichen Aussagen zum Nutzen der Telefontriage führen zur Überlegung, dass für den sinnvollen Einsatz einer präklinischen Triage nicht nur ihre flächendeckende Einführung nötig ist, sondern auch ihre letztendliche Ausgestaltung, sprich ihre Funktionsweise und Qualität, den Nutzen mitbestimmen wird. Dem entsprechen Forderungen nach kontinuierlichen Qualitätssicherungsmaßnahmen der Telefontriage.<sup>31,92</sup> In der präklinischen Triage würde dies einen Vorteil der Symptomchecker bedeuten, deren

Kommunikationsqualität nicht auf der Grundlage der Ausbildung von Fachpersonal beruht, sondern auf Algorithmen, die sich zentral anpassen ließen.

Doch auch für sie gilt, dass ihre Qualität ausreichend sein muss, da die korrekte Einschätzung des Krankheitszustands eines Patienten große Bedeutung für seine medizinische Versorgung haben kann.

An dieser Stelle zeigen sich die besonderen Herausforderungen einer Triage über Distanz. Im Falle von Fehleinschätzungen könnten Patienten sich dazu entscheiden, zuhause zu bleiben und keine nötige Versorgung aufzusuchen. Werden wegweisende Untersuchungen in der Folge zu spät durchgeführt oder Therapien zu spät angesetzt, kann dies zu irreversiblen Schäden führen und Morbidität sowie Mortalität ansteigen lassen. Entscheidend ist neben einer schnellen Versorgung außerdem, in welcher Einrichtung ein Patient behandelt wird. So können hochakute Krankheitsbilder, wie ein Schlaganfall oder ein Herzinfarkt, ausschließlich in einer Klinik mit entsprechenden Fachabteilungen adäquat behandelt werden. Ein besonderes Risiko besteht damit bei Fällen mit hoher zeitlicher Dringlichkeit, wie sie in dieser Studie als *Emergent Care* eingestuft wurden. Betrachtet man diese Fälle gesondert, konnten wir zeigen, dass die Symptomchecker 71.60% von ihnen richtig triagierten. Dies bedeutet gleichzeitig, dass über ein Viertel der Notfälle nicht einer entsprechenden Notfallversorgung zugeführt worden wären. Auch die sofortige Vorstellung bei einem niedergelassenen Arzt würde in solchen Fällen zu einem Nachteil führen.

Auf der anderen Seite liegt in der Funktion der Symptomchecker, wie zuvor erwähnt, das Potenzial, die Gesundheitsversorgung zu entlasten. Dazu wäre es nötig, Patienten mit leichten Beschwerden eine Selbstbehandlung zu Hause zu empfehlen. Beschwerden dieser Art entsprechen der Kategorie *Self Care* unter den Fallvignetten in dieser Studie. Die Auswertung ergab jedoch, dass nur 35.46% dieser Fälle richtig triagiert wurden. Die deutschsprachigen Symptomchecker schnitten in dieser Kategorie mit 47.61% erneut besser ab als die englischsprachigen mit durchschnittlich nur 30.30%.

Dies bedeutet, dass knapp zwei Dritteln der Patienten mit harmlosen Beschwerden die Konsultation eines Arztes empfohlen worden wäre. Somit bestätigt sich die Einschätzung einer risikoscheuen Dringlichkeitseinstufung seitens der Symptomchecker durch Semigran et al..<sup>56</sup> Denn durch eine höhere Bewertung der Dringlichkeit wird das Risiko, einen behandlungsbedürftigen Krankheitszustand zu verkennen, verringert. Es sind wohl ethische und haftungsrechtliche Motive, die dieses „Upgrading“ der Triagierung von Symptomcheckern begründen.

Die daraus resultierenden finanziellen und sozialen Konsequenzen, dass sich gegebenenfalls mehr Patienten beim Arzt vorstellen, werden nicht von den Symptomcheckern getragen, sodass für sie kein wirtschaftlicher Anreiz besteht, möglichst genau zwischen behandlungsbedürftig und harmlos zu entscheiden. Das würde dazu führen,



dass im bestehenden Gesundheitssystem Patienten, welche ohne akuten Krankheitszustand Beratung suchen, die ohnehin schon vielerorts überlasteten Notaufnahmen weiter strapazieren.<sup>95</sup> Derzeit scheint der potenzielle Nutzen der Symptomchecker aufgrund ihrer risikoscheuen Triage nicht zum Tragen zu kommen. Für eine direkte Vergleichbarkeit von Symptomcheckern und Telefontriage bezüglich präklinischer Triage bedarf es weiterer Forschung. Zu diesem Zeitpunkt sind die Symptomchecker jedoch keine Empfehlungsoption für Patienten.

### **5.2.3. Leistungsfähigkeit der Studierenden**

Bei den Symptomcheckern entstehen die Diagnose und Triagierung in der direkten Interaktion zwischen Programm und Patient. Überlegungen zu möglichen Vor- und Nachteilen der Symptomchecker basieren immer auch auf einem indirekten Vergleich zur Leistungsfähigkeit des bisherigen Systems, in dem Diagnosestellung und Triage klassischerweise auf zwischenmenschlicher Kommunikation beruhen. So sind es Ärzte und Ärztinnen, die Diagnosen stellen und die Dringlichkeit eines Falls einstufen. Für beide Aufgaben kann, abgesehen von approbierten Ärzten, auch medizinisch geschultes Personal wie Pflegekräfte als Vergleichsgruppe in Betracht gezogen werden.<sup>30</sup> Auch Medizinstudierende stellen ein Beispiel für nicht-ärztliches Fachpersonal dar, dessen Expertise auch in der telemedizinischen Konsultation gefragt ist.<sup>31,32</sup> Deshalb wurden in dieser Studie die 49 standardisierten Fallvignetten von Studierenden der Humanmedizin bearbeitet.

Die Bearbeitung der Fallvignetten erfolgte in dieser Studie in zwei Modalitäten, zum einen durch eine reine Vorlage des Vignettentexts und zum anderen mittels Chatanamnese. Dabei waren die Studierenden aufgefordert per Chat innerhalb von 5 Minuten eine Anamnese durchzuführen. Durch diese Bearbeitungsform sollte eine Interaktivität ermöglicht werden, die der Anamneseführung in der Praxis ähnelt. Die Bearbeitungszeit von 5 Minuten wurde orientierend an der durchschnittlichen Konsultationszeit in der Erstaufnahmesituation weltweit und in Deutschland gewählt.<sup>96-98</sup> Zudem erfolgte eine Verkürzung auf 5 Minuten unter Berücksichtigung der geringeren Fülle an verfügbaren Zusatzinformationen der „Patienten“ in den Fallvignetten sowie der entfallenden Begrüßungs- und Verabschiedungsgepflogenheiten. Wie bei der Testung der Symptomchecker wurden die von den Studierenden gestellten Fragen anhand der Fallvignetten beantwortet. Somit waren in den beiden Bearbeitungsmodalitäten „Text“ und „Chat“ letztendlich die gleichen Informationen verfügbar. In den Ergebnissen zeigten sich zwischen ihnen jedoch signifikante Unterschiede.

Die Studierenden stellten die richtige Diagnose signifikant häufiger, wenn sie den Fall in Textform vorliegen hatten. Dies gilt für die Kategorien EC (Text 60.84% gegenüber Chat

43.36%) und NE (Text 68.35% gegenüber Chat 44.06%). Für die Fallvignetten der Kategorie *Self Care* zeigte sich kein signifikanter Unterschied. Eine Hypothese ist es, dass die einfacheren und harmloseren Symptome dieser Fallvignetten bereits in den ersten Fragen eines Chatanamnesegesprächs erfragt werden. Die bessere Bearbeitung der Textfälle in den beiden Kategorien EC und NE könnte darüber hinaus durch die generelle Herausforderung einer guten Anamnese zu erklären sein. Die Textform offenbart alle Informationen und Symptome eines Falls, wohingegen in der Chatanamnese vergessen werden kann, vorhandene Symptome abzufragen oder genannte Symptome vertiefend zu explorieren. Diese Vermutung unterstützen auch die Triageergebnisse des Vergleichs zwischen Text- und Chatmodus. Diese zeigten eine geringere Differenz zwischen den beiden Modalitäten, die Textbearbeitung führte in den EC-Fällen aber mit 86.71% gegenüber der Chatbearbeitung mit 65.73% zu einem signifikant besseren Ergebnis. Da in der Textbearbeitung alle Informationen zu einem Zeitpunkt zur Verfügung stehen, ist zu vermuten, dass hierdurch die Identifikation von Warnsymptomen einfacher fiel und so häufiger eine Notfallversorgung empfohlen wurde.

Um die Selbsteinschätzung der Studierenden bewerten zu können, wurde ihre subjektive Einschätzung der Schwierigkeit mit ihren Ergebnissen korreliert. Es konnte gezeigt werden, dass die Studierenden, die die Fallvignetten als einfacher einstufen, diese auch signifikant häufiger richtig diagnostizierten. Dies spricht für eine gute Selbsteinschätzung der Studierenden und gegen eine Verunsicherung durch eine der beiden Modalitäten. Gleichzeitig stellte der Arbeitsschritt der Anamnese eine zusätzliche Anstrengung für die Studierenden dar. Dies zeigte sich daran, dass die Bearbeitungsmodalität „Chat“ als signifikant schwieriger eingeschätzt wurde als die Textfallbearbeitung. Als möglicher Bias im Vergleich der Modalitäten wurde von uns die Tippgeschwindigkeit der Teilnehmer identifiziert. Bei der Bearbeitung via Chatanamnese hätte dies zu einer Beeinträchtigung der Bearbeitungsfähigkeit führen können. Die Studierenden gaben die Einschränkungen durch ihre Tippgeschwindigkeit jedoch mit 3.44 von 10 (10 = sehr stark eingeschränkt) und einer Standardabweichung von SD = 2.19 durchweg als niedrig an. Somit ist dieser Faktor aus unserer Sicht zu vernachlässigen.

Um die Auswirkung der Anzahl gestellter Fragen auf die Wahrscheinlichkeit einer korrekten Diagnosestellung zu ermitteln, wurden diese beiden Parameter miteinander korreliert. Dabei zeigte sich kein signifikanter Zusammenhang. Die Hypothese, dass mehr Fragen zu einer gründlicheren und damit erfolgreicherer Anamnese führen, bestätigt sich somit nicht. Demnach ist die zielführende Fragestellung bedeutsam, der Erfolg stellt sich schneller ein, wenn wenige gezielte Fragen anstatt vieler unpräziser Fragen gestellt werden.<sup>99</sup> Auch eine offene Frage, wie „Haben Sie sonstige Beschwerden?“, kann essenzielle Symptome erfassen und ersetzt damit gegebenenfalls das einzelne Abfragen von Körpersystemen.<sup>100</sup>

Gleichwohl könnte in den Fällen, wo viele Fragen gestellt wurden, letztendlich doch noch eine richtige Diagnose generiert worden sein, sodass sich die beiden Mechanismen gegenseitig aufheben und keine Korrelation entsteht.

Die teilnehmenden Studierenden in dieser Studie befanden sich alle im klinischen Abschnitt ihres Studiums. Dieser hat eine Dauer von sechs Semestern und wurde in der vorliegenden Studie in eine erste und zweite Hälfte unterteilt. Dabei zeigte sich, dass die fortgeschrittenen Studierenden in der zweiten Hälfte signifikant häufiger die richtige Diagnose nannten als die Studierenden in der ersten. Damit bestätigt sich die Annahme, dass mehr Basiswissen zu einer besseren Diagnoseleistung führt. Die Daten zur Leistungsfähigkeit der Studierenden in unserer Studie können mit der Arbeit der Arbeitsgruppe von Semigran et al. aus dem Jahr 2016 verglichen werden, in der die Diagnose- und Triagefähigkeiten von Ärzten mit Hilfe der Online-Plattform *HumanDx* getestet wurden.<sup>101</sup> Auf dieser stellen Ärzte als gemeinschaftliches Projekt Verdachtsdiagnosen für diverse klinische Fälle. Über die Plattform wurden die Fallvignetten Ärzten vorgelegt und jede einzelne mindestens 20 Mal bearbeitet. Da dies jedoch online geschieht, stand der Versuchsaufbau und die Qualifikation der teilnehmenden Personen nur unter eingeschränkter Kontrolle.

Vergleicht man die Ergebnisse der Ärzte mit den Studierenden zeigt sich bei den Diagnosen auf der Top 1 (Ärzte 72.1% gegenüber Studierende 64.10%) und in den Top 3 (Ärzte 84.3% gegenüber Studierende 80.19%) nur eine geringfügig bessere Leistung dieser Vergleichsgruppe. Als Erklärung für diese knappe Überlegenheit könnte ihre klinische Erfahrung dienen, da sie auch eine Rolle im System 1 Denken spielt.<sup>14</sup> Es ist anzunehmen, dass die Studierenden der höheren klinischen Semester hingegen nur geringfügig mehr Erfahrung durch Famulaturen und klinische Praktika sammeln können. Trotzdem ist nicht auszuschließen, dass dies einen Einfluss auch auf ihre Diagnosefähigkeit hat.

#### **5.2.4. Leistungsvergleich der Symptomchecker und der Studierenden**

Um Symptomchecker und Studierende unter möglichst gleichen Bedingungen zu vergleichen, wurden den Studierenden die Fallvignetten in Textform vorgelegt. Die Studierenden stellten mit 64.10% die präzise Verdachtsdiagnose (Symptomchecker 38.02%) und mit 80.19% eine zumindest noch korrekte Differentialdiagnose (Symptomchecker 56.44%), womit sie eine signifikant höhere Treffergenauigkeit und -wahrscheinlichkeit als die Symptomchecker erreichten. Dies gilt auf allen Triagestufen (EC, NE, SC). Vergleicht man die studentischen Assessments ausschließlich mit den deutschsprachigen Symptomcheckern, fällt auf, dass ein signifikanter Unterschied nur in der Kategorie *Non-emergent Care* vorhanden ist. Hier wäre eine mögliche Erklärung, dass die Nuancen zwischen harmlos und gefährlich schwerer für die Programme zu erkennen sind. Es fehlen zum einen Warnsignale oder *red flags*, zum anderen sind aber auch nicht nur unbedenkliche Symptome, wie Schnupfen und Müdigkeit, vorhanden, die nur einen geringen Krankheitswert

haben, sodass eine Verharmlosung oder Übergewichtung der Symptome erfolgt und eine falsche Verdachtsdiagnose resultiert.

In der Triage schnitten die Studierenden ebenfalls besser ab. Eine signifikante höhere Rate an richtigen Dringlichkeitseinstufungen war aber im Vergleich zu den Symptomcheckern nur in der Kategorie *Emergent Care* zu finden (86.71% gegenüber 71.60%). Dies widerspricht der zuvor erwähnten Einschätzung eines risikoscheuen Verhaltens der Symptomchecker, da sie für die EC-Fälle weniger häufig die höchste Dringlichkeitsstufe wählten. Ein Erklärungsansatz für die Überlegenheit der Studierenden könnte sein, dass sie eher in der Lage waren auch harmlose Symptome in einer speziellen Konstellation einer Notsituation zuordnen zu können. Ebenso konnten sie eventuell starke Ausprägungen eines sonst harmlosen Symptoms besser erkennen und es dadurch richtig bewerten. Beispielsweise sind Schmerzen je nach Stärke unterschiedlich zu bewerten und ein Symptom wie Kopfschmerzen erst in Kombination mit Lichtscheu und Nackensteifigkeit verdächtig für eine Meningitis.

Wurde in unserer Studie ein Vergleich von Studierenden und Symptomcheckern angestrengt, erfolgte dieser in der Studie von Semigran et al. zuvor zwischen Ärzten und Symptomcheckern. Jene Untersuchung ergab, dass die Ärzte mit 72,1% deutlich häufiger als die Symptomchecker die korrekte Diagnose an erster Stelle nannten.<sup>101</sup> Innerhalb der ersten drei Verdachtsdiagnosen, wurde die korrekte Diagnose von ihnen in 84,3% der Fälle genannt. Als Kernaussage des Papers ergibt sich somit, dass die Diagnosekorrektheit von Ärzten deutlich höher ist als die der Symptomchecker. Die Ergebnisse unserer Studie weisen darauf hin, dass ihre Fähigkeiten zurzeit die der Symptomchecker noch deutlicher übersteigen, jedoch auch die Studierenden mit medizinischem Basiswissen und kaum klinischer Erfahrung ähnlich gute Ergebnisse erzielen.

### **5.2.5. Methodische Limitationen und Aussagekraft der Studie**

Beim Betrachten der Ergebnisse unserer Studie sind Grenzen der Untersuchung zu beachten. Sie können die Vergleichbarkeit der Resultate mit anderen Forschungsergebnissen beeinflussen. Das schmale Zeitfenster der Untersuchung sowie die monozentrische Erhebung der Daten schränkt die Repräsentativität für größere Studierendengruppen und größere Zeitabschnitte ein. Die Aussagekraft dieser Studie beruht ebenfalls stark auf dem Zeitpunkt, zu dem der Markt für Symptomchecker analysiert wurde. Durch den Auswahlprozess wurden anhand festgelegter Kriterien Symptomchecker in die Studie eingeschlossen. Dabei unterschieden sich die Kriterien für die englischsprachigen und deutschsprachigen Symptomchecker nicht, um den Vergleich mit bisherigen Studien zu erleichtern.

Allerdings ist zu beachten, dass die Entwicklung dieser Angebote ein junges Phänomen ist und deshalb bloß eine geringe Anzahl an Symptomcheckern untersucht wurde. Aus diesem

Grund werden auch in Zukunft Studien nötig sein, um die Qualität in der Entwicklung der Symptomchecker zu evaluieren. Dabei ist stets zu differenzieren, ob sich die Schlussfolgerungen einer Studie auf ein Gesamtangebot von Symptomcheckern beziehen oder bloß eine Auswahl.<sup>102</sup>

Die standardisierten klinischen Fallvignetten, die in dieser Studie Verwendung fanden, waren aus den Untersuchungen von Semigran et al. übersetzt und adaptiert worden. Ergänzt wurden sie um vier weitere allgemeinmedizinische Fälle. Die Fallvignetten dienten als Annäherung an einen Patientenfall, können dabei aber nicht die Fülle an Informationen bereitstellen, die ein realer Patient geben könnte. Die Studierenden gaben entsprechend im Fragebogen den Bedarf an zusätzlichen Informationen während der Fallbearbeitung mit einem Durchschnittswert von 3.91 an (Skala von 1 = „sehr viele zusätzliche Informationen“ bis 10 = „überhaupt keine zusätzlichen Informationen“). Dieser hohe Bedarf an Zusatzinformationen kann damit erklärt werden, dass die Fallvignetten im Gegensatz zu einem realen Patienten weniger Fragen eindeutig beantworten können.

Auch aus unserer Sicht weisen die Fallvignetten Schwächen auf, da sie teils Antworten zu wichtigen Differentialdiagnosen nicht verneinen und somit Interpretationsspielraum lassen. Es wurde außerdem Kritik an der Auswahl der Krankheitsbilder geäußert, da diese einen hohen Anteil seltener Diagnosen beinhalten, und sich Medizinstudierende auf diesen Umstand besser einstellen können als die Programme, die stets die Wahrscheinlichkeit einer Diagnose mit einbeziehen.<sup>102</sup>

Die Drop-Out-Quote lag bei 0%, sodass es zu keinen Verzerrungseffekten durch Abbruch oder Ausschluss kam. Es ist möglich, dass es zu einer Stichprobenverzerrung durch die Selbstselektion der teilnehmenden Studierenden kam (Self-Selection-Bias). Hierbei ist davon auszugehen, dass sich dadurch vor allem am Thema interessierte Personen zur Verfügung stellten. Damit ist eine Übertragung auf die gesamte Studierendenschaft nicht mehr möglich. Eine Schweigeverzerrung (non-response-Bias) kann angenommen werden und ist zumindest nicht auszuschließen. Damit ist gemeint, dass eine Untersuchung unter Studierenden, die nicht teilgenommen haben, zu abweichenden Ergebnissen geführt hätte. Aus ethischer Perspektive wäre jedoch eine Teilnahmeverpflichtung von zufällig ausgewählten Studierenden nicht vertretbar. In zukünftigen Studien sollte darauf geachtet werden, entsprechende Untersuchungen an einem möglichst repräsentativen Feld von Teilnehmenden durchzuführen.

#### **5.2.6. Resümee und Ausblick**

Diese Untersuchung fokussierte sich auf die Leistungsfähigkeit der Symptomchecker, mit der Absicht eine Grundlage zu schaffen für Überlegungen zum potenziellen Nutzen und Schaden durch diese Programme. Die möglichen Auswirkungen reichen vom Erleben des einzelnen Patienten bis hin zur Interaktion mit dem existenten Gesundheitssystem.

Die Symptomchecker richten sich mit ihrer Funktion in erster Linie an Patienten. Durch die Verbreitung von Symptomcheckern könnte einer großen Zahl von Menschen eine schnelle und kostengünstige Möglichkeit zur Selbstdiagnose ermöglicht und der Zugang zu medizinischen Informationen weiter erleichtert werden. Damit bilden die Symptomchecker die Fortsetzung einer Entwicklung, in der das Internet zunehmend größere Bedeutung für Patienten gewinnt. Eine Untersuchung in den USA ergab, dass von allen Erwachsenen 35% das Internet speziell zur Diagnosefindung für sich oder andere Personen nutzten. In knapp mehr als der Hälfte der Fälle wurde von den Suchenden anschließend ein Arzt konsultiert.<sup>42</sup>

Eine deutsche Studie, basierend auf einer Umfrage in Nordrhein-Westfalen, fand heraus, dass im Jahr 2011 bereits 67,9% der Internetnutzer Gesundheitsinformationen im Netz suchten. Über die Hälfte der Nutzer suchte nach solchen Informationen, da sie selbst erkrankt waren. Weitere 21,4% suchten explizit nach Symptomen.<sup>64</sup>

Bestimmte Bevölkerungsgruppen nehmen dabei unterschiedlich häufig Onlinequellen in Anspruch. So ist zu erkennen, dass in sozial schwachen Schichten deutlich weniger auf medizinische Informationsquellen online zurückgegriffen wird, als in Mittel- oder Oberschicht.<sup>42,43,64</sup> Beobachten lässt sich dieser Unterschied auch zwischen Patienten mit geringerem gegenüber Patienten mit höherem Bildungsniveau.<sup>103</sup>

Ebenso gibt es eine Korrelation zwischen der starken Inanspruchnahme von Leistungen der Gesundheitsversorgung und der Wahrscheinlichkeit, dass sich ein Patient auch online informiert hat.<sup>64</sup> Ob das Informieren im Internet Patienten jedoch zusätzlich motiviert, die Gesundheitsversorgung in Anspruch zu nehmen, bleibt unklar.

Mit Blick in die Zukunft ist davon auszugehen, dass besonders auch mobile Endgeräte Einfluss auf das Suchverhalten der Patienten haben werden.<sup>45</sup> Schon im Jahr 2012 hatten bereits 19% der Smartphone-Nutzer eine Gesundheitsapp auf ihrem Gerät.<sup>104</sup> In einer weiteren Befragung im Jahr 2015 waren es 34% der Smartphone-Nutzer.<sup>105</sup> Unter diesen mobilen Applikationen sind unter anderem Programme zur Fitnessüberwachung, Gewichtsreduzierung, Testung von Sehfähigkeit, Rauchentwöhnung oder Symptomkontrolle zu finden.<sup>106</sup> Es ist somit davon auszugehen, dass mobile Endgeräte die Bedeutung des Internets in Gesundheitsfragen stark verändern werden. Die ständige Verfügbarkeit eines Smartphones könnte beispielsweise zu einer noch stärkeren Nutzung führen und dadurch auch die Bedeutung der Symptomchecker in den kommenden Jahren weiter steigern.

Wie zuvor erwähnt, ist die aktuelle Forschung mit Bezug auf die Leistung der verfügbaren Symptomchecker überschaubar. Die beiden bisherigen Untersuchungen kommen von Semigran et al. und betrachten den englischsprachigen Markt für Symptomchecker.<sup>56,101</sup> Als Limitationen dieser Studien sind die Auswahl der Probanden sowie die begrenzte Differenzierung in der Betrachtung der Ergebnisse zu nennen. Der Anspruch unserer Studie

war es deshalb, auch einen Überblick über die Leistung deutschsprachiger Symptomchecker zu geben und sie in einem Vergleich zu menschlicher Leistungsfähigkeit zu unterziehen. Über eine Leistungsüberprüfung der Symptomchecker hinausgehend existiert Forschung zu ihrer Bedeutung und ihren Charakteristika. So enthalten die Symptomchecker in vielen Fällen einen Hinweis, wie die Verdachtsdiagnosen gewertet werden sollen. Dabei wird häufig betont, dass diese nicht den Besuch eines Arztes ersetzen und im Zweifelsfall immer ein Mediziner kontaktiert werden sollte. Ebenfalls wird auf die Notwendigkeit einer Notfallbehandlung bei schwerwiegenden Symptomen hingewiesen. Diese Diskrepanz zwischen beworbener Funktion und rechtlicher Absicherung wird von Jutel und Lupton (2015) kritisiert<sup>55,57</sup>. Denn sie kommen auch zu der Erkenntnis, dass in einer Großzahl der Angebote gezielt medizinische oder informationstechnische Fachtermini verwendet werden, die darauf abzielen, Glaubwürdigkeit und Autorität auszustrahlen.<sup>55</sup> Es ist denkbar, dass sich Patienten in der Folge verunsichert fühlen. Denn durch eine Selbstdiagnose kann der Patient in seinem persönlichen Erleben beeinflusst werden. Stigmata könnten verstärkt und Zukunftsängste ausgelöst werden.<sup>57</sup> Es liegt dann am Patienten selbst, die Glaubwürdigkeit des Symptomcheckers und die Dringlichkeit der Handlungsempfehlung gegeneinander abzuwägen. Das Ermitteln von Informationen zu Herausgebern, Nachweisen zur fachlichen Expertise der Programmierer oder den kommerziellen Interessen hinter dem Angebot gestaltet sich allerdings schwierig und ist zeitaufwendig.<sup>55</sup> In der zukünftigen Forschung ist es somit nötig, neben der Leistung auch die Ausgestaltung der Symptomchecker zu untersuchen. Dabei ergibt sich ein interdisziplinäres Forschungsfeld, das Fragen zu Datensicherheit, Interessenskonflikten, Haftungsfragen, Nutzerfreundlichkeit und medizinischer Expertise umfasst.

## 6. Literaturverzeichnis

- 1 Jutel A. Sociology of diagnosis: a preliminary review. *Sociology of Health & Illness* 2009; **31**: 278–99.
- 2 Pape W. Handwörterbuch der griechischen Sprache, 3rd edn. Braunschweig: Vieweg & Sohn, 1914.
- 3 Oxford University Press. Diagnosis, N. OED online. 2015. <https://www.oed.com/view/Entry/51836?redirectedFrom=diagnosis&> (accessed April 24, 2019).
- 4 Veith I. Historical reflections on the changing concepts of disease. *Calif Med* 1969; **110**: 501–6.
- 5 Jutel A. Classification, Disease, and Diagnosis. *Perspectives in Biology and Medicine* 2011; **54**: 189–205.
- 6 Bloch H. Thomas Sydenham, MD (1624-1689): the father of clinical observation. *Journal of Family Practice* 1994; **38**: 80–2.
- 7 Sydenham T, SWAN J. The Entire Works of Dr. Thomas Sydenham, Newly Made English ... To which are Added, Explanatory and Practical Notes, from the Best Medical Writers. By John Swan. London: Edward Cave, 1742.
- 8 O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring Diagnoses: ICD Code Accuracy. *Health Services Research* 2005; **40**: 1620–39.
- 9 Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Medical Education* 2015; **49**: 961–73.
- 10 Bowen JL. Educational Strategies to Promote Clinical Diagnostic Reasoning. *New England Journal of Medicine* 2006; **355**: 2217–25.
- 11 MFT Medizinischer Fakultätentag der Bundesrepublik Deutschland. Nationaler Kompetenzbasierter Lernzielkatalog Medizin. Berlin, 2015.
- 12 National Academies of Sciences, Engineering, and Medicine. Improving Diagnosis in Health Care. Washington, D.C.: National Academies Press, 2015 DOI:10.17226/21794.
- 13 Young M, Thomas A, Lubarsky S, *et al.* Drawing Boundaries: The Difficulty in Defining Clinical Reasoning. *Academic Medicine* 2018; **93**: 990–5.
- 14 Croskerry P, Nimmo GR. Better clinical decision making and reducing diagnostic error. *Journal of the Royal College of Physicians of Edinburgh*. 2011; **41**: 155–62.
- 15 Djulbegovic B, Hozo I, Beckstead J, Tsalatsanis A, Pauker SG. Dual processing model of medical decision-making. *BMC Medical Informatics and Decision Making* 2012; **12**: 94.



- 16 Croskerry P. A Universal Model of Diagnostic Reasoning. *Academic Medicine* 2009; **84**: 1022–8.
- 17 Thammasitboon S, Cutrer WB. Diagnostic decision-making and strategies to improve diagnosis. *Current Problems in Pediatric and Adolescent Health Care*. 2013; **43**: 232–41.
- 18 Nendaz M, Perrier A. Diagnostic errors and flaws in clinical reasoning: mechanisms and prevention in practice. *Swiss Medical Weekly* 2012; **142**. DOI:10.4414/smw.2012.13706.
- 19 De Swaan A. The reluctant imperialism of the medical profession. *Social Science & Medicine* 1989; **28**: 1165–70.
- 20 Rosenberg CE. The Tyranny of Diagnosis: Specific Entities and Individual Experience. *The Milbank Quarterly* 2002; **80**: 237–60.
- 21 Parsons T. Illness and the Role of the Physician: a Sociological Perspective. *American Journal of Orthopsychiatry* 1951; **21**: 452–60.
- 22 Benjamin JR, George EE. Disaster Medicine: Impact On Critical Care Operations. In: Parsons PE, Wiener-Kronish JP, eds. *Critical Care Secrets: Fifth Edition*, 5th edn. Philadelphia, PA: Elsevier/Mosby, 2013: 521–6.
- 23 Frykberg ER. Medical management of disasters and mass casualties from terrorist bombings: How can we cope? *Journal of Trauma* 2002; **53**: 201–12.
- 24 Somasundaram R, Ale Abaei A, Koch M. Triage in zentralen Notaufnahmen. *Notfall + Rettungsmedizin* 2009; **12**: 250–5.
- 25 Ellebrecht N. Triage. Charakteristika und Gegenwart eines ordnungsstiftenden Verfahrens. *Sociologia Internationalis* 2009; **47**: 229–57.
- 26 Haller HL, Wurzer P, Peterlik C, Gabriel C, Cancio LC. Burn Management in Disasters and Humanitarian Crises. *Total Burn Care* 2018; : 36-49.e2.
- 27 Gemeinsamer Bundesausschuss. Regelungen des Gemeinsamen Bundesausschusses zu einem gestuften System von Notfallstrukturen in Krankenhäusern gemäß § 136c Absatz 4 SGB V Stand: 19. April 2018. 2018; : 8.
- 28 Lake R, Georgiou A, Li J, *et al.* The quality, safety and governance of telephone triage and advice services - An overview of evidence from systematic reviews. *BMC Health Services Research* 2017; **17**: 1–10.
- 29 Boggan JC, Shoup JP, Whited JD, *et al.* Effectiveness of Acute Care Remote Triage Systems: a Systematic Review. *Journal of General Internal Medicine*. 2020; : 1–10.
- 30 Jiwa M, Mathers N, Campbell M. The effect of GP telephone triage on numbers seeking same-day appointments. *Br J Gen Pract* 2002; **52**: 390–1.
- 31 Graf von Stillfried D, Czihal T, Meer A. State of the art report—structured medical first assessment in Germany (SmED). *Notfall und Rettungsmedizin*. 2019; **22**: 578–88.

- 32 Braga AV. Die telemedizinische Konsultation. In: Pfannstiel M, Da-Cruz P, Mehlich H, eds. *Digitale Transformation von Dienstleistungen im Gesundheitswesen I*. Wiesbaden: Springer Gabler, 2017: 93–108.
- 33 Dörner K. *Der gute Arzt: Lehrbuch der ärztlichen Grundhaltung*, 2nd edn. Stuttgart: Schattauer, 2003.
- 34 Myfanwy M. The doctor-patient relationship. In: Scambler G, ed. *Sociology as Applied to Medicine*, 6th edn. Edinburgh, London, New York, Oxford, Philadelphia, St.Louis, Sydney, Toronto: Saunders, 2008: 55–69.
- 35 Geisler L. Arzt-Patient-Beziehung im Wandel – Stärkung des dialogischen Prinzip. In: Deutscher Bundestag, ed. *Abschlussbericht der Enquête-Kommission Recht und Ethik der modernen Medizin*. Opladen, 2002: 216–20.
- 36 Bury M, Taylor D. Towards a Theory of Care Transition: From Medical Dominance to Managed Consumerism. *Social Theory and Health* 2008; **6**: 201–19.
- 37 Faller H. Patientenorientierte Kommunikation in der Arzt-Patient-Beziehung. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2012; **55**: 1106–12.
- 38 McKinlay JB, Marceau LD. The end of the golden age of doctoring. *International Journal of Health Services* 2002; **32**: 379–416.
- 39 Barello S, Graffigna G, Vegni E. Patient Engagement as an Emerging Challenge for Healthcare Services: Mapping the Literature. *Nursing Research and Practice* 2012; **2012**. DOI:10.1155/2012/905934.
- 40 Lupton D. The digitally engaged patient: Self-monitoring and self-care in the digital health era. *Social Theory & Health* 2013; **11**: 256–70.
- 41 Nettleton S, Burrows R. E-Scaped Medicine? Information, Reflexivity and Health. *Critical Social Policy* 2003; **23**: 165–85.
- 42 Fox S, Duggan M. *Health Online 2013*. Washington D.C., 2013.
- 43 Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. *Cogent Social Sciences* 2017; **3**: 1302785.
- 44 Tang H, Ng JHK. Googling for a diagnosis - Use of Google as a diagnostic aid: Internet based study. *British Medical Journal* 2006; **333**: 1143–5.
- 45 Branscum P, Hayes L, Wallace L. DIRECT OBSERVATION OF SEARCHING FOR ONLINE HEALTH INFORMATION: A SYSTEMATIC REVIEW OF CURRENT EVIDENCE. *American Journal of Health Studies* 2016; **31**: 222–32.
- 46 DeVaney SA. Understanding the millennial generation. *Journal of Financial Service Professionals* 2015; **69**: 11–4.

- 47 Bolton RN, Parasuraman A, Hoefnagels A, *et al.* Understanding Generation Y and their use of social media: A review and research agenda. *Journal of Service Management* 2013; **24**: 245–67.
- 48 Prensky M. Digital Natives, Digital Immigrants. *On the horizon* 2001; **9**: 1–6.
- 49 Statistisches Bundesamt (Destatis). Erhebung über die private Nutzung von Informations- und Kommunikationstechnologien. Wiesbaden, 2018.
- 50 Tsetsi EL. Digital Divide 3.0: The Mobile Revolution, Smartphone Use, and the Emerging Device Gap. 2016.
- 51 van Deursen AJ, van Dijk JA. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New Media & Society* 2019; **21**: 354–75.
- 52 Wartella E, Rideout V, Montague H, Beaudoin-Ryan L, Lauricella A. Teens, Health and Technology: A National Survey. *Media and Communication* 2016; **4**: 13–23.
- 53 Walsh A. Mobile information literacy: a preliminary outline of information behaviour in a mobile environment. *Journal of Information Literacy* 2012; **6**. DOI:10.11645/6.2.1696.
- 54 Liu Z, Huang X, Fu Y. Information behavior in the mobile environment: A study of undergraduate smartphone users in China. *Chinese Journal of Library and Information Science* 2014; **7**: 1–15.
- 55 Lupton D, Jutel A. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. *Social Science & Medicine* 2015; **133**: 128–35.
- 56 Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: Audit study. *BMJ (Online)* 2015; **351**: h3480.
- 57 Jutel A, Lupton D. Digitizing diagnosis: a review of mobile applications in the diagnostic process. *Diagnosis* 2015; **2**: 89–96.
- 58 Food and Drug Administration. Mobile Medical Applications - Guidance for Industry and Food and Drug Administration Staff. 2015; : 13–5.
- 59 Food and Drug Administration. Examples of Software Functions for Which the FDA Will Exercise Enforcement Discretion. 2019. <https://www.fda.gov/medical-devices/device-software-functions-including-mobile-medical-applications/examples-software-functions-which-fda-will-exercise-enforcement-discretion> (accessed Nov 27, 2019).
- 60 Krieger WH. Medical Apps: Public and Academic Perspectives. *Perspectives in Biology and Medicine* 2013; **56**: 259–73.
- 61 Morita T, Rahman A, Hasegawa T, Ozaki A, Tanimoto T. The Potential Possibility of Symptom Checker. *Int J Health Policy Manag* 2017; **6**: 615–6.

- 62 Orizio G, Merla A, Schulz PJ, Gelatti U. Quality of online pharmacies and websites selling prescription drugs: A systematic review. *Journal of Medical Internet Research*. 2011; **13**: e74.
- 63 Aujoulat I, d'Hoore W, Deccache A. Patient empowerment in theory and practice: Polysemy or cacophony? *Patient Education and Counseling* 2007; **66**: 13–20.
- 64 Nölke L, Mensing M, Krämer A, Hornberg C. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. *BMC Public Health* 2015; **15**: 31.
- 65 Kirch W, Engwicht A. Definition und Häufigkeit der Fehldiagnose. In: Kirch W, ed. *Fehldiagnosen und Patientensicherheit*. Berlin, Heidelberg: Springer, 2005: 1–52.
- 66 Schwartz A, Elstein AS. Clinical reasoning in medicine. In: Higgs J, Jones M, Loftus S, Christensen N, eds. *Clinical Reasoning in the Health Professions*, 3rd edn. Edinburgh: Elsevier, 2008: 223–34.
- 67 Berner ES, Graber ML. Overconfidence as a Cause of Diagnostic Error in Medicine. *American Journal of Medicine* 2008; **121**: S2–23.
- 68 Pineda LA, Hathwar VS, Grant BJB. Clinical Suspicion of Fatal Pulmonary Embolism. *Chest* 2001; **120**: 791–5.
- 69 Graff L, Russell J, Seashore J, *et al.* False-negative and False-positive Errors in Abdominal Pain Evaluation Failure to Diagnose Acute Appendicitis and Unnecessary Surgery. *Academic Emergency Medicine* 2000; **7**: 1244–55.
- 70 Singh H, Meyer AND, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014; **23**: 727–31.
- 71 Zwaan L, Singh H. The challenges in defining and measuring diagnostic error. *Diagnosis* 2015; **2**: 97–103.
- 72 Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf* 2013; **22 Suppl 2**: ii21–7.
- 73 Graber ML, Franklin N, Gordon R. Diagnostic Error in Internal Medicine. *Archives of Internal Medicine* 2005; **165**: 1493.
- 74 Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. *BMJ Qual Saf* 2017; **26**: 484–94.
- 75 Glassman; JPLJB. Assessing the Accuracy of Administrative Data in Health Information Systems. *Medical Care* 2004; **42**: 1066–72.
- 76 Gandhi TK, Kachalia A, Thomas EJ, *et al.* Missed and Delayed Diagnoses in the Ambulatory Setting: A Study of Closed Malpractice Claims. *Annals of Internal Medicine* 2006; **145**: 488.

- 77 Zwaan L, Thijs A, Wagner C, Van Der Wal G, Timmermans DRM. Relating faults in diagnostic reasoning with diagnostic errors and patient harm. *Academic Medicine* 2012; **87**: 149–56.
- 78 Brown TW, McCarthy ML, Kelen GD, Levy F. An Epidemiologic Study of Closed Emergency Department Malpractice Claims in a National Database of Physician Malpractice Insurers. *Academic Emergency Medicine* 2010; **17**: 553–60.
- 79 Panesar SS, Desilva D, Carson-Stevens A, *et al.* How safe is primary care? A systematic review. *BMJ Quality and Safety*. 2016; **25**: 544–53.
- 80 Croskerry P. Perspectives on diagnostic failure and patient safety. *Healthc Q* 2012; **15 Spec No**: 50–6.
- 81 World Health Organization. Global Priorities for Research in Patient Safety. Genf, 2008.
- 82 Hölscher UM, Gausmann P, Haindl H, *et al.* Übersichtsartikel: Patientensicherheit als nationales Gesundheitsziel: Status und notwendige Handlungsfelder für die Gesundheitsversorgung in Deutschland. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 2014; **108**: 6–14.
- 83 Saber Tehrani AS, Lee H, Mathews SC, *et al.* 25-Year summary of US malpractice claims for diagnostic errors 1986-2010: an analysis from the National Practitioner Data Bank. *BMJ Qual Saf* 2013; **22**: 672–80.
- 84 Bundesärztekammer. Behandlungsfehler-Statistik der Gutachterkommissionen und Schlichtungsstellen 2018. Berlin, 2019.
- 85 Van den Heede K, Van de Voorde C. Interventions to reduce emergency department utilisation: A review of reviews. *Health Policy* 2016; **120**: 1337–49.
- 86 Scherer M, Lühmann D, Kazek A, Hansen H, Schäfer I. Patients attending emergency departments—a cross-sectional study of subjectively perceived treatment urgency and motivation for attending. *Dtsch Arztebl Int* 2017; **114**: 645–52.
- 87 Downes MJ, Mervin MC, Byrnes JM, Scuffham PA. Telephone consultations for general practice: A systematic review. *Systematic Reviews* 2017; **6**: 128.
- 88 Blank L, Coster J, O’Cathain A, *et al.* The appropriateness of, and compliance with, telephone triage decisions: a systematic review and narrative synthesis. *Journal of Advanced Nursing* 2012; **68**: 2610–21.
- 89 Carrasqueiro S, Oliveira M, Encarnaço P. Evaluation of telephone triage and advice services: A systematic review on methods, metrics and results. In: *Studies in Health Technology and Informatics*. IOS Press, 2011: 407–11.
- 90 Christ M, Bingisser R. Telefontriage in der Schweiz: Patienten sind zufrieden. *Deutsches Ärzteblatt International* 2017; **114**: A1472–3.

- 91 Von Gossler C, Klauser C. Telemedizin in der Schweiz – das Beispiel Medgate. *DMW - Deutsche Medizinische Wochenschrift* 2017; **142**: 334–9.
- 92 Meer A, Gwerder T, Duembgen L, Zumbrunnen N, Zimmermann H. Is computer-assisted telephone triage safe? A prospective surveillance study in walk-in patients with non-life-threatening medical conditions. *Emergency Medicine Journal* 2012; **29**: 124–8.
- 93 McKinstry B, Hammersley V, Burton C, *et al.* The quality, safety and content of telephone and face-to-face consultations: A comparative study. *Quality and Safety in Health Care* 2010; **19**: 298–303.
- 94 Huibers L, Smits M, Renaud V, Giesen P, Wensing M. Safety of telephone triage in out-of-hours care: A systematic review. *Scandinavian Journal of Primary Health Care*. 2011; **29**: 198–209.
- 95 Pines JM, Hilton JA, Weber EJ, *et al.* International Perspectives on Emergency Department Crowding. *Academic Emergency Medicine* 2011; **18**: 1358–70.
- 96 Valverde Bolívar FJ, Pedregal González M, Moreno Martos H, Cózar García I, Torío Durántez J. Communication with patients and the duration of family medicine consultations. *Atencion Primaria* 2018; **50**: 621–8.
- 97 Deveugele M, Derese A, van den Brink-Muinen A, Bensing J, de Maeseneer J. Consultation length in general practice: Cross sectional study in six European countries. *BMJ* 2002; **325**: 472.
- 98 Irving G, Neves AL, Dambha-Miller H, *et al.* International variations in primary care physician consultation time: A systematic review of 67 countries. *BMJ Open*. 2017; **7**: e017902.
- 99 Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club*. 1995; **123**: A12-3.
- 100 Silverman J, Kurtz S, Draper J. *Skills for communicating with patients*, 3rd edn. London: Radcliffe, 2013.
- 101 Semigran HL, Levine DM, Nundy S, Mehrotra A. Comparison of Physician and Computer Diagnostic Accuracy. *JAMA Internal Medicine* 2016; **176**: 1860–1.
- 102 Fraser HSF, Clamp S, Wilson CJ. Limitations of study on symptom checkers. *JAMA Internal Medicine*. 2017; **177**: 740–1.
- 103 Tennant B, Stellefson M, Dodd V, *et al.* eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. *J Med Internet Res* 2015; **17**: e70.
- 104 Fox S, Duggan M. *Pew Internet & American Life Project: Mobile Health 2012*. Washington D.C., 2012.

- 105 Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach. *J Med Internet Res* 2017; **19**: e125.
- 106 Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J Public Health Inform* 2014; **5**: 229.

## **7. Anhang**

### **7.1. Abbildungsverzeichnis**

<b>Abbildung 1. Schema des Dual Processing nach P. Croskerry und Nimmo</b>	<b>14</b>
<b>Abbildung 2. Versuchsaufbau der Studierendenuntersuchung</b>	<b>34</b>
<b>Abbildung 3. Ablauf der Studierendenuntersuchung</b>	<b>35</b>
<b>Abbildung 4. Vergleich der Diagnosekorrektheit in Textbearbeitung und Chatbearbeitung für Studierende in der ersten und zweiten Hälfte des klinischen Studienabschnitts</b>	<b>45</b>
<b>Abbildung 5. Vergleich der Triagequalität in Textbearbeitung und Chatbearbeitung für Studierende in der ersten und zweiten Hälfte des klinischen Studienabschnitts</b>	<b>46</b>
<b>Abbildung 6. Korrekte Diagnosestellungen nach Dringlichkeitsstufen in den Modalitäten Text und Chat</b>	<b>47</b>
<b>Abbildung 7. Triagequalität nach Dringlichkeitsstufe in den Modalitäten Text und Chat</b>	<b>48</b>

### **7.2. Tabellenverzeichnis**

<b>Tabelle 1. Liste der untersuchten Symptomchecker mit Kurzbeschreibung</b>	<b>25</b>
<b>Tabelle 2. Standardisierte klinische Fallvignetten in ihrer Kurzfassung</b>	<b>28</b>
<b>Tabelle 3. Verwendete Programme und Geräte der Studie</b>	<b>37</b>
<b>Tabelle 4. Diagnostische Treffergenauigkeit der Symptomchecker</b>	<b>38</b>
<b>Tabelle 5. Diagnostische Trefferwahrscheinlichkeit der Symptomchecker</b>	<b>38</b>
<b>Tabelle 6. Triagequalität der Symptomchecker</b>	<b>39</b>
<b>Tabelle 7. Vergleich der diagnostischen Treffergenauigkeit deutsch- und englischsprachiger Symptomchecker</b>	<b>39</b>
<b>Tabelle 8. Vergleich der diagnostischen Trefferwahrscheinlichkeit deutsch- und englischsprachiger Symptomchecker</b>	<b>40</b>
<b>Tabelle 9. Vergleich der Triagequalität deutsch- und englischsprachiger Symptomchecker</b>	<b>40</b>
<b>Tabelle 10. Differenzielle Diagnosequalität der Symptomchecker</b>	<b>41</b>
<b>Tabelle 11. Differenzielle Triagequalität der Symptomchecker nach Dringlichkeitsstufen</b>	<b>42</b>
<b>Tabelle 12. Demographische und Fachsemesterdaten Studierenden</b>	<b>43</b>
<b>Tabelle 13. Diagnostische Treffergenauigkeit der Studierenden-Assessments</b>	<b>43</b>
<b>Tabelle 14. Diagnostische Trefferwahrscheinlichkeit der Studierenden-Assessments</b>	<b>43</b>
<b>Tabelle 15. Triagequalität der Studierenden-Assessments</b>	<b>44</b>
<b>Tabelle 16. Zusammenhang korrekter Textfallbearbeitung und subjektiv empfundener Schwierigkeit</b>	<b>49</b>
<b>Tabelle 17. Zusammenhang korrekter Chatfallbearbeitung und subjektiv empfundener Schwierigkeit</b>	<b>50</b>
<b>Tabelle 18. Zusammenhang der Fragenanzahl und korrekten Chatfallbearbeitungen</b>	<b>50</b>
<b>Tabelle 19. Vergleich Symptomchecker- und Studierenden-Assessments</b>	<b>51</b>
	<b>79</b>



<b>Tabelle 20. Vergleich deutschsprachiger Symptomchecker- und Studierenden-Assessments</b>	<b>52</b>
<b>Tabelle 21. Vergleich englischsprachiger Symptomchecker- und Studierenden-Assessments</b>	<b>53</b>
<b>Tabelle 22. Zusammenfassung der Leistung der Symptomchecker</b>	<b>54</b>
<b>Tabelle 23. Zusammenfassung des Leistungsvergleichs deutsch- und englischsprachiger Symptomchecker</b>	<b>55</b>
<b>Tabelle 24. Zusammenfassung Vergleich Symptomchecker- und Studierenden-Assessments</b>	<b>55</b>
<b>Tabelle 25. Auflistung der standardisierten klinischen Fallvignetten auf Deutsch in ihrer vollständigen Form</b>	<b>81</b>

### 7.3. Standardisierte klinische Fallvignetten auf Deutsch in ihrer vollständigen Form

Tabelle 25. Auflistung der standardisierten klinischen Fallvignetten auf Deutsch in ihrer vollständigen Form

Kürzel Diagnose		Vereinfachte Fallinformationen
Hohe Dringlichkeit ( <i>emergent care</i> = EC)		
<b>EC1</b>	<b>Akutes Leberversagen</b>	<p>Eine 48-jährige Frau mit Migräne-Kopfschmerzen in der Vergangenheit präsentiert sich in der Notaufnahme mit in den letzten Stunden verändertem mentalen Status. Sie wurde von ihrem Ehemann früher am Tag als sehr desorientiert und zunehmend schläfrig empfunden. Bei der körperlichen Untersuchung, hat sie ikterische Skleren, Bauchdeckenabwehrspannung im oberen rechten Quadranten sowie Asterixis. Im vorläufigen Labor sind ein Serum-ALT von 6498 Einheiten/L, Gesamt-Bilirubin von 5,6 mg/dL und INR von 6,8 auffällig. Ihr Ehemann berichtet, dass sie seit Jahren Schmerztabletten konsumiert und vor einigen Tagen zusätzliche 500 mg Paracetamol-Tabletten eingenommen hat, um Rückenschmerzen zu lindern. Die weitere Anamnese zeigt eine Medikamentenliste mit mehreren Acetaminophen-haltigen Präparaten.</p>
<b>EC2</b>	<b>Appendizitis</b>	<p>Ein 12-jähriges Mädchen stellt sich mit plötzlich einsetzenden schweren generalisierten Bauchschmerzen vor, die mit Übelkeit, Erbrechen und Durchfall einhergehen. Bei der Untersuchung erscheint sie krank und hat eine Temperatur von 40 ° C. Ihr Bauch ist angespannt. Zusätzlich zeigt sich eine generalisierte Bauchdeckenabwehrspannung. Es sind keine Darmgeräusche zu hören.</p>
<b>EC3</b>	<b>Asthma</b>	<p>Eine 27-jährige Frau mit mittelschwerem, persistierendem Asthma präsentiert sich in der Notaufnahme mit fortschreitender Verschlechterung der Kurzatmigkeit, Keuchen und Husten über die letzten 3 Tage. Sie berichtet von einem Kontakt mit einer Person, die eine laufende Nase und Husten hatte. Sie erhielt keine nennenswerte Erleichterung von ihrem Notfallinhalator mit sich verschlechternden Symptomen, trotz erhöhter Dosierung. Sie ist mit ihrem Asthma-Erhaltungsprogramm, das aus einem inhalativen Corticosteroid und einem Leukotrien-Rezeptor-Antagonisten für die Erhaltungstherapie und Albuterol als Notfalltherapie besteht, compliant. Ihr Husten unterbricht ihren Schlafrythmus und als Folge davon leidet sie unter Schläfrigkeit tagsüber, was ihre Arbeitsleistung beeinträchtigt.</p>
<b>EC4</b>	<b>COPD-Anfall (schwer)</b>	<p>Eine 67-jährige Frau mit COPD zeigt nach 3 Tagen eine Verschlechterung ihrer Dyspnoe und eine erhöhte Hustenfrequenz. Ihr Husten ist jetzt produktiv mit grünem, eitrigem Sputum. Die Patientin hat 100 pack years (py). Sie hatte in den letzten 3 Tagen intermittierend leichtes Fieber von 37,7 ° C und ihr Appetit ist schlecht. Um die Symptome zu kontrollieren, benutzt sie vermehrt die Notfall-Bronchodilatator-Medikation zusätzlich zu ihrer</p>

		Erhaltungsmedikation.	
<b>EC5</b>	<b>Tiefe Venenthrombose</b>	Eine 65-jährige Frau stellt sich mit seit 5 Tagen bestehenden Schmerzen im rechten Bein vor. Das Bein ist angeschwollen. Die Patientin leidet an Bluthochdruck und einer milden kongestiven Herzinsuffizienz. Die Patientin hatte vor kurzem einen Krankenhausaufenthalt wegen einer Lungenentzündung. Sie hatte sich zu Hause erholt, aber als sie mit der Mobilisierung begann, wurde das rechte Bein schmerzhaft und schwoll an. Bei der Untersuchung ist der Umfang der rechten Wade 4 cm größer als der Umfang der linken Wade (gemessen 10 cm unterhalb der Tuberositas tibiae). Oberflächliche Venen im Bein sind am rechten Fuß dilatiert. Das rechte Bein ist etwas röter als das linke Bein. In der rechten Kniekehle liegt eine leichte Palpationsempfindlichkeit vor.	65 Jahre, w, seit 5 Tagen Schwellung, Schmerzen im rechten Bein, kürzlicher Krankenhausaufenthalt, Bein schmerzhaft, empfindlich, geschwollen, rot
<b>EC6</b>	<b>Herzinfarkt</b>	Herr Y. ist ein 64 Jahre alter Mann, der sich mit seit 24 Stunden bestehenden Schmerzen in der Brust vorstellt. Die Symptome begannen einen Tag vor der Vorstellung beim Arzt mit einem nicht ausstrahlendem, substernalen Brustdruck (8 von 10), Schweißausbrüchen und Kurzatmigkeit. Der Schmerz besserte sich anfänglich mit Paracetamol, jedoch verschlechterten sich seine Symptome in den folgenden 24 Stunden. Der Patient ging zu seinem Hausarzt, wo ein EKG durchgeführt wurde, das ST-Hebungen in den Ableitungen V2-V6 zeigte.	64 Jahre, m, seit 1 Tag Schmerzen in der Brust (8/10), nicht-ausstrahlender, substernaler Brustdruck, Schwitzen, Kurzatmigkeit, Engegefühl in der Brust
<b>EC7</b>	<b>Hämolytisches-urämisches Syndrom</b>	Ein 4-jähriger Junge stellt sich mit seit 7 Tagen bestehenden Bauchschmerzen und wässrigen Durchfällen vor, die nach dem ersten Tag blutig wurden. Drei Tage vor Beginn der Symptome hatte er mit seiner Familie die Kirmes besucht und einen Hamburger gegessen. Die körperliche Untersuchung zeigt eine leichte Anämie.	4 Jahre, m, seit 7 Tagen, Bauchschmerzen, blutiger Durchfall, aß Hamburger auf Kirmes 3 Tage vor Auftreten der Symptome
<b>EC8</b>	<b>Nierensteine</b>	Ein 45-jähriger Mann stellt sich in der Notaufnahme mit vor einer Stunde plötzlich aufgetretenen, linksseitigen Flankenschmerzen vor, die sich in Richtung seiner Leistenbeuge ausbreiten. Der Patient krümmt sich unter Schmerzen, die durch Positionswechsel nicht gelindert werden. Er klagt auch über Übelkeit und Erbrechen.	45 Jahre, m, seit 1 Stunde starker linksseitiger Flankenschmerz, der in die Leistengegend ausstrahlt, Übelkeit, Erbrechen, Schmerzen, die durch Positionswechsel nicht gelindert werden
<b>EC9</b>	<b>Malaria</b>	Ein 28-jähriger Mann stellt sich bei seinem Arzt mit seit 5 Tagen bestehendem Fieber, Schüttelfrost, Rigor und Durchfall vor. Die Einnahme von Paracetamol führte zu keiner Besserung. Er war 3 Monate durch Mittelamerika gereist und kehrte vor 8 Wochen zurück. Er war mehrmals von Moskitos gebissen worden, und, obwohl er anfangs eine Malariaphylaxe einnahm, unterbrach er diese wegen leichter Übelkeit. Die Details seiner prophylaktischen Therapie kann er nicht berichten. Bei der Untersuchung hat er eine Temperatur von 38 ° C und ist leicht tachykard bei einem Blutdruck von 126/82 mmHg. Der restliche Befund der körperlichen Untersuchung ist normal.	28 Jahre, m, seit 5 Tagen Schüttelfrost, Rigor, Durchfall, kürzliche Reise ins Ausland mit Malaria, von Moskitos gebissen, nahm Malaria-Prophylaxe nicht konsequent
<b>EC10</b>	<b>Meningitis</b>	Ein 18-jähriger männlicher Student leidet seit 3 Tagen unter starken Kopfschmerzen und Fieber. Die Untersuchung zeigt Fieber, Photophobie und Nackensteifigkeit.	18 Jahre, m, seit 3 Tagen starke Kopfschmerzen, Fieber, Photophobie, Nackensteifigkeit
<b>EC11</b>	<b>Pneumonie</b>	Ein 65-jähriger Mann mit Bluthochdruck und einer degenerativen Gelenkerkrankung stellt sich in der Notaufnahme mit seit 3 Tagen	65 Jahre, m, Bluthochdruck und degenerative Gelenkerkrankungen, seit

	<p>bestehendem produktiven Husten und Fieber vor. Er hat eine Temperatur von 38,3 ° C, einen Blutdruck von 144/92 mmHg, eine Atemfrequenz von 22 Atemzügen pro Minute, eine Herzfrequenz von 90 Schlägen pro Minute und eine Sauerstoffsättigung von 92 Prozent bei Raumlufte. Die körperliche Untersuchung zeigt Rasselgeräusche und Egophonie im rechten unteren Lungenfeld. Die Leukozytenzahl beträgt 14.000 pro Kubikmillimeter. Alle weiteren Laborparameter sind normal. Eine Röntgenaufnahme des Thorax zeigt ein Infiltrat im rechten Unterlappen.</p>	<p>3 Tagen produktiver Husten und Fieber (38,3° C)</p>
<b>EC12 Lungenembolie</b>	<p>Ein 65-jähriger Mann stellt sich mit akuter Kurzatmigkeit über 30 Minuten in der Notaufnahme vor. Anfangs fühlte er sich schwach, verlor aber nicht das Bewusstsein. Er klagt über linksseitige Brustschmerzen, die sich bei tiefer Inspiration verschlimmern. Er hat keine bekannten Herz-Lungen-Erkrankungen. Vor einer Woche bekam er im Krankenhaus eine komplette Hüft-Endoprothese und lag nach der Entlassung aufgrund von schlecht kontrolliertem Schmerz für 3 Tage im Bett. Er bemerkte später eine Schwellung in seiner linken Wade. Die Wade ist im Rahmen der Untersuchung schmerzempfindlich. Er hat eine Temperatur von 38 ° C, eine Herzfrequenz von 112 Schlägen pro Minute, BP 95/65 und eine Sauerstoffsättigung bei Raumlufte von 91%.</p>	<p>65 Jahre, m, Kurzatmigkeit über 30 Minuten, Brustschmerz, der sich mit Inspiration verschlechtert, kürzlich zurückliegende Operation mit Bettruhe, Schwellung in der linken Wade, die schmerzempfindlich ist, Fieber</p>
<b>EC13 Rocky-Mountain-Fleckfieber</b>	<p>Ein 8-jähriger Junge aus Oklahoma wird am 4. Juli wegen Fieber, Schüttelfrost, Unwohlsein, Gelenkschmerzen und Kopfschmerzen in die Notaufnahme gebracht. Die körperliche Untersuchung zeigt einen makulopapulösen Ausschlag, der an seinen Handgelenken und Knöcheln am ausgeprägtesten ist.</p>	<p>8 Jahre, m, Fieber, Schüttelfrost, Gelenkschmerzen, Kopfschmerzen, Hautausschläge an Handgelenken und Knöcheln</p>
<b>EC14 Schlaganfall</b>	<p>Ein 70-jähriger Mann mit chronischem Bluthochdruck und Vorhofflimmern wird von einem Familienmitglied dabei beobachtet, wie er unter Übelkeit, Erbrechen und einer Kraftminderung der rechten Körperhälfte leidet. Zusätzlich bestehen Schwierigkeiten beim Sprechen und Verstehen von Sprache. Die Symptome begannen mit nur leicht verwaschener Sprache, bevor sie im Laufe weniger Minuten in eine schwere Aphasie und eine rechtsseitige Lähmung übergingen. Der Patient nimmt Warfarin ein.</p>	<p>70 Jahre, m, Übelkeit, Erbrechen, rechtsseitige Schwäche, Armlähmung, Schwierigkeiten beim Sprechen und Verstehen von Sprache</p>
<b>EC15 Tetanus</b>	<p>Ein 63-jähriger Mann hat sich während der Gartenarbeit einen Schnitt an der Hand zugezogen. Laut seinem Impfpass hat er keine komplette Tetanusimpfung. Er präsentiert sich mit den Symptomen eines generalisierten Tetanus: Trismus ("Kieferverschluss"), was zu einer Grimasse führt, die als "Risus sardonicus" (sardonisches Lächeln) bezeichnet wird. Intermittierende tonische Kontraktion seiner Skelettmuskeln verursachen intensive schmerzhafte Krämpfe, die Minuten andauern und während denen er das Bewusstsein behält. Die Krämpfe werden durch äußere Einflüsse (Lärm, Licht, Luftzug, physischer Kontakt) oder innere Reize ausgelöst. Durch die Krämpfe läuft er Gefahr Frakturen zu erleiden oder eine Rhabdomyolyse zu entwickeln. Die tetanischen Spasmen erzeugen auch Opisthotonus, bretartige Abdominalwandrigidität, Dysphagie und Apnoephasen aufgrund der Kontraktion der Thoraxmuskeln und / oder der Stimmritze oder Pharynxmuskeln. Während eines generalisierten Krampfes wölbt der Patient seinen Rücken, streckt</p>	<p>65 Jahre, m, kann den Mund nicht öffnen, Kontraktion der Muskeln verursacht schmerzhafte Krämpfe über Minuten, Schwitzen, Tachykardie, Schnitt in die Hand während der Gartenarbeit, nicht gegen Tetanus geimpft</p>

seine Beine aus, beugt seine Arme in Abduktion und ballt seine Fäuste. Im Rahmen mancher Krämpfe tritt Apnoe auf. Die autonome Überaktivität manifestiert sich zunächst als Reizbarkeit, Unruhe, Schwitzen und Tachykardie. Einige Tage später kann sich diese als Hyperpyrexie, Herzrhythmusstörungen, labile Hypertonie oder Hypotonie manifestieren.

<b>EC16 Sigmadivertikulitis</b>	Eine 85-jährige Frau stellt sich im Krankenhaus mit vor einigen Stunden aufgetretenen Schmerzen im linken Unterbauch vor. Sie gibt an, dass sie seit einigen Tagen unter Verstopfung leide. Seit gestern seien auch Übelkeit und Erbrechen hinzugekommen. Die körperliche Untersuchung zeigt eine lokalisierte Druckschmerzhaftigkeit im linken Unterbauch. Die Körpertemperatur liegt bei 38,5 ° C. Im Labor sind sowohl Leukozyten als auch CRP erhöht.	85 Jahre, w, akute Schmerzen im linken Unterbauch, Verstopfung, Übelkeit, Erbrechen, T = 38,5 ° C, CRP erhöht, Leukozyten erhöht
<b>EC17 Ileus</b>	Ein 50-jähriger Mann stellt sich mit starken Bauchschmerzen sowie schwallartigem Erbrechen im Krankenhaus vor. Die körperliche Untersuchung zeigt eine allgemeine abdominale Abwehrspannung. Die Auskultation ergibt klingende, gesteigerte Darmgeräusche. Im Labor sind sowohl Leukozyten als auch CRP erhöht.	50 Jahre, m, starke Bauchschmerzen, schwallartiges Erbrechen, abdominale Abwehrspannung, klingende, gesteigerte Darmgeräusche, CRP erhöht, Leukozyten erhöht

Mittlere Dringlichkeit (*non-emergent care = NE*)

<b>NE1 Akute Otitis media</b>	Ein 18 Monate altes Kleinkind präsentiert sich mit seit 1 Woche bestehender Rhinorrhoe, Husten und Verstopfung der Nase. Ihre Eltern berichten, dass sie gereizt ist, ruhelos schläft und nicht gut isst. Über Nacht bekam sie Fieber. Das Kind ist tagsüber in einer Kinderbetreuung. Beide Eltern rauchen. Die Befunde im Rahmen der körperlichen Untersuchung stehen im Einklang mit einer viralen Infektion der Atemwege einschließlich Rhinorrhoe und Stauung der Nase. Das Kleinkind wirkt reizbar und ängstlich und hat Fieber. Die Otoskopie zeigt eine wulstige, erythematöse Trommelfellmembran und fehlende Landmarken.	18 Monate, w, seit 1 Woche Rhinorrhoe, Husten, Verstopfung der Nase, Reizbarkeit, Mangel an Appetit, Fieber, besucht Kindertagesstätte
<b>NE2 Akute Pharyngitis</b>	Ein 7-jähriges Mädchen stellt sich mit plötzlichem Fieber, Übelkeit, Erbrechen und Halsschmerzen vor. Das Kind verneint Husten, Rhinorrhoe oder eine verstopfte Nase. Bei der körperlichen Untersuchung liegt die orale Temperatur bei 38,5 ° C. Im Rachenraum zeigt sich eine exsudative Pharyngitis mit vergrößerten zervikalen Lymphknoten. Ein schneller Antigentest ist positiv für Gruppe A Streptokokken (GAS).	7 Jahre, w, Fieber (38,5 ° C), Übelkeit, Erbrechen, Halsschmerzen, geschwollene Lymphknoten, Tonsilleneksudat; kein Husten, keine Rhinorrhoe und keine verstopfte Nase
<b>NE3 Akute Pharyngitis</b>	Herr A. ist ein 24-jähriger Mann, der sich wegen Beschwerden wie Halsschmerzen, Fieber und Kopfschmerzen in ihrer Klinik vorstellt. Seine Symptome begannen vor 2 Tagen mit akut einsetzenden Halsschmerzen und Fieber von 39 ° C. Dabei hatte er keinen Husten. Der körperliche Untersuchungsbefund ist, mit Ausnahme von Tonsilleneksudaten und einer leichten vorderen zervikalen Lymphadenopathie, normal. Er ist ansonsten bei guter Gesundheit und nimmt bis auf Ibuprofen gegen das Fieber keine Medikamente ein. Er hat keine Medikamentenallergien. (Centor Score = 4 - behandeln oder testen und behandeln)	24 Jahre, m, Halsschmerzen, Fieber (39 ° C), Kopfschmerzen, kein Husten, Tonsilleneksudate
<b>NE4 Akute Sinusitis</b>	Eine 35 Jahre alte Frau stellt sich mit seit 15 Tagen verstopfter Nase vor. Sie hat seit 12 Tagen Schmerzen im Gesicht und grünen	35 Jahre, w, seit 15 Tagen verstopfte Nase, Gesichtsschmerzen, grüner

	Nasenausfluss. Sie hat kein Fieber. Bei der körperlichen Untersuchung hat sie eine normale Temperatur. Der einzige auffällige Befund ist ein bei Druck schmerzhafter Oberkiefer. Sie ist ansonsten—bis auf eine leichte Fettleibigkeit—gesund. Abgesehen von einem frei verkäuflichen Nasenspray nimmt sie keine Medikamente ein. Sie hat keine Medikamentenallergien.	Nasenausfluss, kein Fieber
<b>NE5 Rückenschmerzen</b>	Ein 35 Jahre alter Mann stellt sich mit Rückenschmerzen nach Schneeschaukeln vor 3 Wochen vor. Bei der Untersuchung zeigt sich eine neu aufgetretene Fußheberschwäche am linken Fuß.	35 Jahre, m, Rückenschmerzen nach Schneeschaukeln, Fußheberschwäche links, Symptome seit 3 Wochen, Gefühlsverlust im Fuß
<b>NE6 Zellulitis</b>	Ein 45-jähriger Mann stellt sich mit akuten Schmerzen und Rötung der Haut seiner unteren Extremität vor. Er hat leichtes Fieber. Der prätibiale Bereich ist erythematös, ödematös und schmerzempfindlich.	45 Jahre, m, Schmerzen und Rötung der Haut, leichtes Fieber, Ödeme und Druckschmerz am Unterschenkel
<b>NE7 COPD-Anfall (mild)</b>	Eine 56-jährige Raucherin stellt sich mit seit einigen Tagen bestehender Kurzatmigkeit und Husten vor. Ihre Symptome begannen vor 3 Tagen mit Rhinorrhoe. Sie berichtet von einem chronischen Husten am Morgen, der zu weißem Sputum führt. Dieser habe in den letzten zwei Tagen zugenommen. Sie hat in den letzten 4 Jahren jeden Winter ähnliche Episoden gehabt. Sie raucht seit 40 Jahren 1 bis 2 Zigarettschachteln pro Tag. Sie bestreitet Bluthusten, Schüttelfrost oder Gewichtsverlust. Rezeptfreie Hustenpräparate brachten keine Besserung.	56 Jahre, w, Raucherin, Kurzatmigkeit und Husten seit mehreren Tagen, Rhinorrhoe vor 3 Tagen aufgetreten, weißer Auswurf, keine Schüttelfrost
<b>NE8 Influenza</b>	Eine 30-jährige Frau stellt sich im Januar mit seit 2 Tagen bestehendem Fieber, Husten, Kopfschmerzen und generalisierte Schwäche vor. Vor dem abrupten Auftreten der Symptome war sie gesund. Einige Viruserkrankungen hat sie in diesem Winter schon durchgemacht—aber nicht mit so schwerem Verlauf. Sie hatte im Rahmen der Arbeit Kontakt zu kranken Personen. In dieser Saison erhielt sie keinen saisonalen Grippeimpfstoff.	30 Jahre, w, seit 2 Tagen Fieber, Husten, Kopfschmerzen, Schwäche, bekam keine Gripeschutzimpfung
<b>NE9 Pfeiffer-Drüsenfieber</b>	Eine 16-jährige, weibliche Gymnasiastin stellt sich mit Fieber, Halsschmerzen und Müdigkeit vor. Sie begann sich vor einer Woche krank zu fühlen. Ihre Symptome wurden allmählich schlimmer und sie hat Schwierigkeiten beim Schlucken. Sie hat jeden Tag Fieber gehabt, und sie konnte heute Morgen kaum aufstehen. Sie erinnert sich nicht, dass sie in letzter Zeit jemandem mit einer ähnlichen Krankheit ausgesetzt war. Bei der körperlichen Untersuchung ist sie fieberhaft und sieht krank aus. Sie hat vergrößerte zervikale Lymphknoten, eine exsudative Pharyngitis mit Petechien des weichen Gaumens und einen schwach erythematösen, makulären Hautausschlag an Rumpf und Armen.	16 Jahre, w, seit 1 Woche Fieber, Halsschmerzen, Müdigkeit, Schluckbeschwerden, vergrößerte Lymphknoten, Exsudate, makulärer Hautausschlag an Rumpf und Armen
<b>NE10 Magengeschwür</b>	Ein 40-jähriger Mann stellt sich bei seinem Hausarzt mit seit 2 Monaten bestehenden, intermittierenden Oberbauchschmerzen vor. Er beschreibt den Schmerz als dumpf und nagend. Der Schmerz weckt ihn manchmal nachts auf. Der Schmerz wird durch Nahrung und Milch gelindert und bessert sich unter der Einnahme von Ranitidin. Er hatte vor etwa 5 Jahren eine ähnliche, aber mildere Episode, die mit Omeprazol behandelt wurde. Die körperliche Untersuchung zeigt einen fitten, scheinbar gesunden Mann. Der einzige auffällige Befund ist eine leichte Oberbauchempfindlichkeit	40 Jahre, m, seit 2 Monaten intermittierende Oberbauchschmerzen, stumpfer, nagender Schmerz, wacht nachts auf, Schmerz bessert sich durch Nahrung/Milch/Ranitidin, ähnliche Episode vor 5 Jahren

	bei Palpation des Abdomens.	
<b>NE11 Pneumonie</b>	Ein 6-jähriger Junge mit leichtem Asthma wird von seiner Mutter in die Klinik gebracht. Seit 5 Tagen leide er an Hustenanfällen. Seine Mutter berichtet, dass das Fieber des Kindes trotz Paracetamol-Therapie weiterhin erhöht ist. Er hat die Schule in den letzten 3 Tagen verpasst und er hat einen Klassenkameraden, der an einer Lungenentzündung erkrankt ist. Die Mutter berichtet, dass der Appetit des Kindes normal sei. Sein Husten zu Hause fördere gelbliches Sputum. Seine Vitalwerte in der Klinik sind: Atemfrequenz 19 Atemzüge/Minute, Herzfrequenz 80 Schläge/Minute und eine Temperatur von 38,7 ° C. Das Kind hat keine akute Atemnot. Die Auskultation der Lunge zeigt bilaterale Rasselgeräusche und gelegentliches Keuchen. Ein Röntgenbild der Lunge zeigt lobuläre Infiltrate ohne Pleuraergüsse.	6 Jahre, m, Asthma, seit 5 Tagen Husten, Fieber, Appetit gut, gelbes Sputum, Temperatur 38,7 ° C
<b>NE12 Salmonellen</b>	Ein 14-jähriger Junge stellt sich mit Übelkeit, Erbrechen und Durchfall vor. Achtzehn Stunden zuvor war er bei einem Picknick gewesen, wo er zusammen mit einer Reihe anderer Lebensmittel auch ungekochtes Hühnchen gegessen hat. Er berichtet 6-mal pro Tag mäßigvolumige, blutungsfreie Stühle abzuführen. Er hat leichte Bauchkrämpfe und leichtes Fieber. Er wird in einer Akutklinik untersucht und ist leicht tachykard (Herzfrequenz 105 Schläge pro Minute) mit einem normalen Blutdruck und einer leicht erhöhten Temperatur von 37,8 ° C. Die körperliche Untersuchung ist, abgesehen von leichten diffusen Schmerzen im Bauch und leicht erhöhten Darmgeräuschen, unauffällig. Er ist in der Lage, oral Flüssigkeiten aufzunehmen. Er wird angewiesen, genug Flüssigkeit und Elektrolyte zu sich zu nehmen.	14 Jahre, m, Übelkeit, Erbrechen, nicht blutiger Durchfall, leichte Bauchkrämpfe (T = 37,8), leichte Bauchschmerzen, Durchfall nach einem Picknick, rohes Hühnchen gegessen
<b>NE13 Gürtelrose</b>	Ein 77-jähriger Mann stellt sich mit seit 5 Tagen bestehenden brennenden Schmerzen auf der rechten Seite seiner Brust vor. Im weiteren Verlauf entwickelt sich ein Erythem sowie ein makulopapulöser Ausschlag in diesem schmerzhaften Bereich, begleitet von Kopfschmerzen und Unwohlsein. Der Ausschlag entwickelt sich weiter zu Klustern aus klaren Vesikeln, die in 3 bis 5 Tagen die Stadien Pusteln, Ulzera und Krustenbildung durchlaufen.	77 Jahre, m, seit 5 Tagen Brennen und Schmerzen auf der rechten Seite der Brust, Erythem, makulopapulöser Ausschlag, Kopfschmerzen, Unwohlsein, Hautausschlag der sich innerhalb 3 bis 5 Tagen zu klaren Vesikeln entwickelt
<b>NE14 Harnwegsinfekt</b>	Eine 26-jährige Frau, die frisch verheiratet ist, klagt seit 2 Tagen über schmerzhaftes Wasserlassen und das Gefühl, ständig urinieren zu müssen sowie häufigeres Wasserlassen. Sie leugnet Fieber, Schüttelfrost, Übelkeit, Erbrechen, Rückenschmerzen, Ausfluss oder vaginalen Pruritus.	26 Jahre, w, schmerzhaftes Wasserlassen, dringendes Wasserlassen, häufigeres Wasserlassen seit 2 Tagen, sexuell aktiv; kein Fieber, Schüttelfrost, Übelkeit, Erbrechen, Rückenschmerzen, Ausfluss, vaginaler Pruritus
<b>NE15 Schwindel</b>	Eine 65-jährige Frau stellt sich mit Schwindel vor. Sie beschreibt den Schwindel als ein plötzliches und schweres Drehgefühl, das dadurch verursacht wird, dass sie sich im Bett auf ihre rechte Seite rollt. Die Symptome dauern normalerweise <30 Sekunden. Die Symptome sind im Laufe des letzten Monats nächtlich und gelegentlich während des Tages aufgetreten, als sie ihren Kopf zurück neigte, um aufwärts zu schauen. Sie beschreibt kein auslösendes Ereignis vor dem Beginn und keinen damit verbundenen Hörverlust, Tinnitus oder andere neurologische Symptome. Otologische und neurologische	65 Jahre, w, Schwindel, plötzlicher Beginn, rezidivierend, dauert <30 Sek., Anhaltender Trigger, kein Hörverlust, Ohrgeräusche, Muskelschwäche, Gefühlsverlust

Untersuchungen sind normal, mit Ausnahme des Dix-Hallpike-Manövers, das auf der linken Seite negativ, auf der rechten Seite jedoch stark positiv ist.

<b>NE16 Morbus Basedow</b>	Eine 40-jährige Frau stellt sich mit Hitzewallung und diffusem Haarausfall beim Arzt vor. Seit ca. einem Jahr leide sie vermehrt unter innerer Unruhe sowie Zittern in den Händen und häufigem Schwitzen. Die körperliche Untersuchung zeigt eine vergrößerte Schilddrüse. Der Ruhepuls liegt bei 110 Schlägen pro Minute. Zusätzlich zeigt sich ein bilateraler Exophthalmus. Im Labor ist TSH erniedrigt. fT3 und fT4 sind im Normbereich.	40 Jahre, w, Hitzewallungen, diffuser Haarausfall, innere Unruhe, Zittern, häufiges Schwitzen, vergrößerte Schilddrüse, Ruhepuls 110 Schläge pro Minute, bilateraler Exophthalmus, TSH erniedrigt, fT3 u. fT4 normal
----------------------------	--	--

### Selbstbehandlung indiziert (self care = SC)

<b>SC1 Akute Bronchitis</b>	Eine 34-jährige Frau ohne bekannte Lungenerkrankungen stellt sich mit seit 12 Tagen bestehendem Husten vor. Sie hatte zunächst eine verstopfte Nase und leichte Halsschmerzen. Jetzt hat sie produktiven Husten ohne Hustenanfälle. Sie hatte keinen Kontakt zu kranken Personen. Bei der körperlichen Untersuchung hat sie keine Atemnot oder Fieber. Alle Vitalparameter sind normal. Es gibt keine Anzeichen für einen Infekt der oberen Atemwege. Bei der Auskultation der Lunge ist vereinzelt Keuchen zu hören.	34 Jahre, w, seit 12 Tagen Husten, anfängliche verstopfte Nase und Halsschmerzen, Husten, kein Fieber
<b>SC2 Akute Bronchitis</b>	Frau L ist eine 61-jährige Frau, die sich mit seit 4 Tagen bestehendem produktivem Husten mit gelben Sputum vorstellt. Ihre Symptome begannen vor 4 Tagen mit Rhinorrhoe und produktivem Husten. Sie hatte anfänglich für 2 Tage Fieber (38,3 ° C). Im Moment hat sie kein Fieber mehr. Im Rahmen der Untersuchung hat sie normale Vitalzeichen und einen unauffälligen Befund. Abgesehen von erhöhten Cholesterinwerten, für die sie mit Atorvastatin behandelt wird, ist sie gesund. Sie hat keine Medikamentenallergie.	61 Jahre, w, seit 4 Tagen Husten, gelbes Sputum, Rhinorrhoe, anfänglich Fieber, jetzt kein Fieber mehr
<b>SC3 Akute Konjunktivitis</b>	Ein 14-jähriger Junge ohne nennenswerte medizinische Vorgeschichte stellt sich mit geröteten Augen vor. Zuerst sei nur das rechte Auge betroffen gewesen. Nach drei Tagen sei aber dann auch das linke Auge betroffen gewesen. Er hat einen wässrigen Ausfluss aus beiden Augen. Beide Augen sind am Morgen verklebt. Er berichtet zusätzlich von neuen Symptomen der oberen Atemwege. Mehrere Kinder in seiner Freizeitgruppe hätten auch gerötete Augen gehabt. Er hat keine starken Schmerzen, keine Lichtempfindlichkeit und er trägt keine Kontaktlinsen. Bei der Untersuchung sind seine Pupillen gleich groß. Beide Pupillen reagieren gleich auf Lichtreize. Der rechte präaurikuläre Lymphknoten ist leicht schmerzempfindlich. Die Penlight-Untersuchung zeigt keine Hornhauttrübung.	14 Jahre, m, seit 3 Tagen rotes, gereiztes Auge (Ausbreitung von rechts nach links), Ausfluss, Symptome eines oberen Atemwegsinfektes, keine Schmerzen oder Lichtempfindlichkeit
<b>SC4 Akute Pharyngitis</b>	Ein 26-jähriger Mann stellt sich mit Halsschmerzen, Kopfschmerzen und nicht produktivem Husten vor. Seine Symptome begannen vor 2 Tagen mit akuten Halsschmerzen. Er hatte kein Fieber. Der körperliche Untersuchungsbefund ist normal—abgesehen von einer Rötung des Rachens. Er ist ansonsten bei guter Gesundheit. Außer Paracetamol gegen die Halsschmerzen nimmt er keine Medikamente ein. Er hat keine Medikamentenallergien.	26 Jahre, m, seit 2 Tagen Halsschmerzen, Kopfschmerzen, Husten, kein Fieber
<b>SC5 Allergischer Schnupfen</b>	Ein 22-jähriger Student gibt an, dass sich über die letzten 5 Jahre Symptome wie eine verstopfte Nase, Niesen und Nasenjucken stetig	22 Jahre, m, seit 5 Jahren stetige Verschlechterung von verstopfter Nase,



	<p>verschlechtert hätten. Die Symptome sind ganzjährig, aber während der Frühjahrssaison schlimmer. Im Laufe der Anamnese zeigt sich, dass er während der Frühjahrssaison erhebliches Augenjucken, Rötung und Tränenfluss sowie Gaumen- und Halsjucken hat. Er erinnert sich, dass seine Mutter ihm erzählt hat, dass er in der Kindheit Ekzeme hatte.</p>	<p>Niesen, Nasenjucken schlimmer im Frühjahr, Juckreiz im Auge, Rötung, Tränenfluss, Juckreiz im Gaumen und Hals, Ekzeme in der Kindheit</p>
<b>SC6 Rückenschmerzen</b>	<p>Ein 38-jähriger Mann, der bisher keine Rückenschmerzen hatte, entwickelte vor 2 Wochen akute Schmerzen im unteren Rücken beim Heben von Kisten. Der Schmerz ist im linken Lumbalbereich lokalisiert und ist mit Krämpfen verbunden. Er beschreibt frühere ähnliche Episoden vor einigen Jahren, die sich ohne Arztbesuch besserten. Er hat keine Schmerzen in den Beinen oder Lähmungen. Er leugnet auch Fieber, Schüttelfrost, Gewichtsverlust und kürzlich zurückliegende Infektionen. Ibuprofen hat etwas geholfen. Aber er hat dies nur zweimal am Tag für die letzten 3 Tage eingenommen, weil er nicht auf Schmerzmittel angewiesen sein will. Im Rahmen der Untersuchung sind Lumbalflexion und -extension als Folge von Schmerzen eingeschränkt. Die neurologische Untersuchung ist jedoch unauffällig.</p>	<p>38 Jahre, m, akute Rückenschmerzen nach dem Heben, keine Schmerzen oder Schwäche in den Beinen, kein Fieber, Schüttelfrost, Gewichtsverlust oder kürzlich zurückliegende Infektionen</p>
<b>SC7 Bienenstich ohne Anaphylaxie</b>	<p>Ein 9-jähriger Junge wird in die Notaufnahme gebracht, nachdem er bei einem Picknick von einer Biene gestochen wurde. Er weint hysterisch. Nach 15 Minuten Beruhigung zeigt die Untersuchung eine geschwollene, schmerzempfindliche Oberlippe, aber keine geschwollene Zunge, keinen Speichelfluss, keinen Stridor, keinen Hautausschlag und keine anderen Beschwerden.</p>	<p>9 Jahre, m, Bienenstich, geschwollene und schmerzempfindliche Oberlippe; keine Schwellung der Zunge, kein Speichelfluss, kein Stridor, kein Hautausschlag oder andere Beschwerden</p>
<b>SC8 Aphthe</b>	<p>Ein 17-jähriger männlicher Student stellt sich mit seit der frühen Schulzeit immer wieder rezidivierenden Ulzerationen im Mund vor. Er hat sonst keine respiratorischen, anogenitalen oder gastrointestinalen Hautläsionen. Seine Mutter litt unter ähnlichen Symptomen in der Jugend. Der Student ist Nichtraucher und trinkt kaum Alkohol. Er nimmt keine Medikamente ein. Die extraorale Untersuchung zeigt keine signifikanten Anomalien und keine Pyrexie; keine zervikale Lymphknotenvergrößerung; keine Anomalien des Nervus cranialis, der Speicheldrüsen oder des Kiefergelenks. Die orale Untersuchung zeigt ein normales Gebiss. Es gibt keine klinischen Anzeichen für einen Verlust oder eine Lockerung der Parodontalbefestigung. Er hat fünf 4 mm runde Geschwüre mit entzündlichen Halos in seinen bukkalen Schleimhäuten.</p>	<p>17 Jahre, m, mit rezidivierenden Ulzerationen, keine respiratorischen, anogenitalen, gastrointestinalen, Augen- oder Hautläsionen, Mutter hatte ähnliche Symptome in der Jugend, keine Medikamente</p>
<b>SC9 Candida-Hefe-Infektion</b>	<p>Eine 40-jährige, verheiratete Frau (monogam) klagt seit 2 Tagen über vaginalen Juckreiz mit dickem, weißem Ausfluss. Sie hat keine Bauchschmerzen oder Fieber.</p>	<p>40 Jahre, w, seit 2 Tagen vaginaler Juckreiz, dicker weißer Ausfluss, keine Bauchschmerzen oder Fieber</p>
<b>SC10 Verstopfung</b>	<p>Ein 5 Monate alter Junge stellt sich mit Schwierigkeiten beim Abführen von hartem Stuhl vor. Seine Mutter berichtet, dass er sich mehrere Stunden lang bemüht abzuführen. An manchen Tagen führe er gar nicht ab. Am folgenden Tag führe er dann unter Schreien Stuhl ab. In der Windel finden sich dann auch Blutflecken. Laut der Mutter wurde das Kind vor kurzem von Muttermilch auf Milchpulver umgestellt, welches er anfangs nur ungerne getrunken hat. Das Kind gedeiht ansonsten normal und ernährt sich ansonsten</p>	<p>5 Monate, m, Schwierigkeit / Verspätung beim Passieren von hartem Stuhl, stundenlang Stress, kann einen Tag verfehlen, schreit beim Stuhlgang und vereinzelte Blutflecken in der Windel, entwöhnt von Muttermilch zu Kuhmilch, ernährt sich normal</p>

	normal. Es gab keine neonatale Verzögerung der Defäkation und kein übermäßiges Erbrechen oder Blähungen.	
<b>SC11 Ekzem</b>	Ein 12-jähriges Mädchen stellt sich mit trockener, juckender Haut (an den Flexoren vor der Ellenbeuge, hinter ihren Knien und am Sprunggelenk) vor. Es zeigen sich auch Flecken an den Wangen. Die Haut ist trocken und schuppig. Sie hat Heuschnupfen und wurde kürzlich mit einer Eier- und Milchallergie diagnostiziert. Sie hat einen Bruder mit Asthma und einen Onkel und mehrere Cousins, bei denen Ekzeme diagnostiziert wurden.	12 Jahre, w, trockene, juckende Haut an den Flexoren an der Ellenbeuge, hinter den Knien, am Sprunggelenk, Wangen haben Flecken, trockene, schuppige Haut, Heuschnupfen, Ei und Milchallergie, Bruder hat Asthma und Onkel und Cousins haben Ekzeme
<b>SC12 Hordeolum</b>	Ein 30-jähriger Mann stellt sich mit einem seit gestern geschwellenem rechtem Auge vor. Er berichtet von leichten Schmerzen bei der Palpation des Augenlids. Trauma, Verkrustung oder eine Veränderung der Sehfähigkeit liegen nicht vor. Er hat keine Allergien oder Augenerkrankungen. Er hat in letzter Zeit keine neuen Seifen, Lotionen oder Cremes verwendet. In der Untersuchung zeigt sich eine Schmerzempfindlichkeit bei Palpation des Augenlids sowie ein Erythem in der Mittellinie des unteren Augenlids in der Nähe des Lidrandes. Der Rest der körperlichen Untersuchung, einschließlich des Globus, ist normal.	30 Jahre, m, seit dem Vortag schmerzhaftes, geschwellenes rechtes Auge, kein Trauma, keine Krustenbildung, keine Sehstörungen, keine Allergien oder Augenerkrankungen, lokalisierte Schmerzempfindlichkeit, Erythem (Rötung)
<b>SC13 Virale Infektion der oberen Atemwege</b>	Herr R. ist ein 56 Jahre alter Mann, der sich mit seit 6 Tagen bestehendem, unproduktiven Husten, verstopfter Nase und grünem Nasenausfluss vorstellt. Er hatte vorübergehend Fieber (38,2 ° C). Bis auf Rhinorrhoe ist der Untersuchungsbefund unauffällig. Abgesehen von einer Arthrose des rechten Knies ist er gesund. Er hat keine Medikamentenallergien.	56 Jahre, m, seit 6 Tagen Husten, verstopfte Nase, grüner Nasenausfluss, Fieber (38,2 ° C), Rhinorrhoe
<b>SC14 Virale Infektion der oberen Atemwege</b>	Ein 30-jähriger Mann stellt sich mit seit 2 Tagen bestehendem Schnupfen und Halsschmerzen vor. Er fühlt sich heiß und verschwitzt, hat leichte Kopfschmerzen, hustet klares Sputum und klagt über Muskelschmerzen. Er hätte gerne Antibiotika, wie sie ihm letztes Jahr verschrieben wurden, als er ähnliche Symptome hatte. Bei der Untersuchung hat er kein Fieber, er hat einen normalen Puls, sein Pharynx ist leicht entzündet. Zusätzlich hat er eine nicht-schmerzhafte Schwellung der zervikalen Lymphknoten. Es gibt keine Nackensteifigkeit und die Auskultation der Brust ist unauffällig. Er hat rezeptfreie Hustenmedikamente eingenommen. Diese haben ihm aber keine Linderung verschafft. Er raucht 10 Zigaretten pro Tag.	30 Jahre, m, seit 2 Tagen laufende Nase, Halsschmerzen, heiß, verschwitzt, leichte Kopfschmerzen, Husten mit klarem Auswurf, Muskelschmerzen, kein Fieber oder Nackensteifigkeit
<b>SC15 Erbrechen</b>	Ein 2 Jahre altes, männliches Kind hat Fieber. Das Kind hat sich bereits zweimal übergeben. Die Mutter sorgt sich, dass das Kind austrocknen könnte. Ein Glas Apfelsaft erbricht das Kind sofort wieder. Die Temperatur ist 38,0 ° C.	2 Jahre, m, leichtes Fieber (T = 38,0 ° C), zweimal erbrochen, erbricht Saft
<b>SC16 Refluxösophagitis</b>	Ein 55-jähriger Mann stellt sich bei seinem Arzt mit seit ca. einem Jahr bestehendem Sodbrennen vor. Vor allem nach dem Essen oder abends im Bett verspüre er ein retrosternales Brennen. Zusätzlich leide er an einem Druckgefühl in der Magenregion sowie Luftaufstoßen. Der Patient hat einen BMI von 30 und 25 pack years (py).	55 Jahre, m, postprandiales, retrosternales Brennen, Druckgefühl in der Magenregion, BMI 30, 25 pack years

## 7.4. Einverständniserklärung

### Einverständniserklärung

**Projekttitle:**

**Monozentrische prospektive Studie zur Beurteilungsfähigkeit von Medizinstudenten und medizinischen Applikationen in Bezug auf das Erkennen und Einschätzen von Erkrankungen**

.....  
Name und Vorname Studienteilnehmer/in in Druckbuchstaben

geb. am ..... Teilnehmer-Nr. ....

Im Rahmen des wissenschaftlichen Projektes werden personenbezogene Daten durch die beteiligten Wissenschaftler und Ärzte erhoben. Mir ist weiterhin bekannt, dass diese Daten zu Studienzwecken gesondert in elektronischer Form gespeichert werden und bin damit einverstanden. Die Weitergabe, Speicherung und Auswertung dieser projektbezogenen Daten erfolgt streng vertraulich und unter Einhaltung der gesetzlichen Bestimmungen.

Ich hatte Gelegenheit, alle die genannte Studie betreffenden Fragen zu klären und wurde im Rahmen der Initiierungspräsentation über das Ziel und den Verlauf der Studie aufgeklärt. Das Merkblatt zu den „Teilnahmeinformationen“ habe ich erhalten. Über das Recht, meine Teilnahme jederzeit ohne Angabe von Gründen beenden zu können, bin ich aufgeklärt worden.

**Ich erkläre hiermit meine Teilnahme an der oben genannten Studie.**

\_\_\_\_\_  
**Name (Druckbuchstaben) und Unterschrift**

\_\_\_\_\_  
**Ort, Datum**

Möglichkeit zur Dokumentation zusätzlicher Fragen seitens des Probanden oder sonstiger Aspekte des Aufklärungsgesprächs:

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

\_\_\_\_\_  
**Name des Studienarztes in Druckbuchstaben**

\_\_\_\_\_  
**Unterschrift des Studienarztes**

\_\_\_\_\_  
**Ort, Datum**